Momentum Turning Points

Christian L. Goulding
Research Affiliates, LLC, Newport Beach, CA 92660 USA

Campbell R. Harvey*
Duke University, Durham, NC 27708 USA
National Bureau of Economic Research, Cambridge, MA 02138 USA

Michele G. Mazzoleni
Research Affiliates, LLC, Newport Beach, CA 92660 USA

May 23, 2022

Turning points are the Achilles' heel of time-series momentum portfolios. Slow signals fail to react quickly to changes in trend while fast signals are often false alarms. We analyze how momentum portfolios of various intermediate speeds, formed by blending slow and fast strategies, handle turning points. We find that the intersection of slow and fast signal directions possesses predictive information, including predictably negative returns when both signals are negative. We propose a novel decomposition of momentum strategy alpha, highlighting the role of volatility timing; and a mean-variance optimal dynamic speed-selection strategy with efficient out-of-sample performance across international equity markets.

Keywords: time-series momentum, volatility timing, market timing, asset pricing, trend following, turning points, momentum speed, mean reversion, behavioral finance

JEL Classifications: G12, G13

*First posted on SSRN: November 18, 2019. Send correspondence to: Campbell R. Harvey, Fuqua School of Business, Duke University, Durham, NC 27708. Phone: +1 919 660 7768, E-mail: cam.harvey@duke.edu. We thank the editor, Nikolai Roussanov, as well as two anonymous referees for their comments. We appreciate the comments of Rob Arnott, Ravi Jagannathan, Juhani Linnainmaa, conference participants at the American Finance Association Annual Meeting, and seminar participants at Research Affiliates. We thank Kay Jaitly for editorial assistance. We also thank Ashish Garg, who was a collaborator on an early version of this paper.

1. Introduction

Time-series (TS) momentum strategies are based on two main premises. First, expected returns vary over time. Second, the *signs* of expected returns have some degree of persistence. If the expected return is positive (negative) this month, then it is more likely to remain positive (negative) next month than to flip sign. TS momentum strategies attempt to exploit such variation and persistence in expected returns by taking long positions in "uptrend" phases and short positions in "downtrend" phases, based on the signs of some lookback horizon of trailing returns. In uptrend or downtrend phases, such strategies will tend to place good bets. However, at momentum turning points, which mark reversals in trend from uptrend to downtrend or vice versa, TS momentum strategies are prone to place bad bets because they rely on observations of realized returns, which reflect a mixture of different trend regimes and noise.

The speed (or sensitivity to recent data) of the momentum signal balances the tension between reducing the impact of noise and reacting quickly to turning points. This tension plays out differently for different speeds. Either the momentum signal attempts to reduce the influence of noise by having a relatively long lookback window (e.g., 12 months) but thus is *slow* to react to a turning point (Type II error of missed detection), or the momentum signal attempts to be *fast* to react to a turning point by having a relatively short lookback window (e.g., 1 month) and therefore is more influenced by noise (Type I error of false alarm).

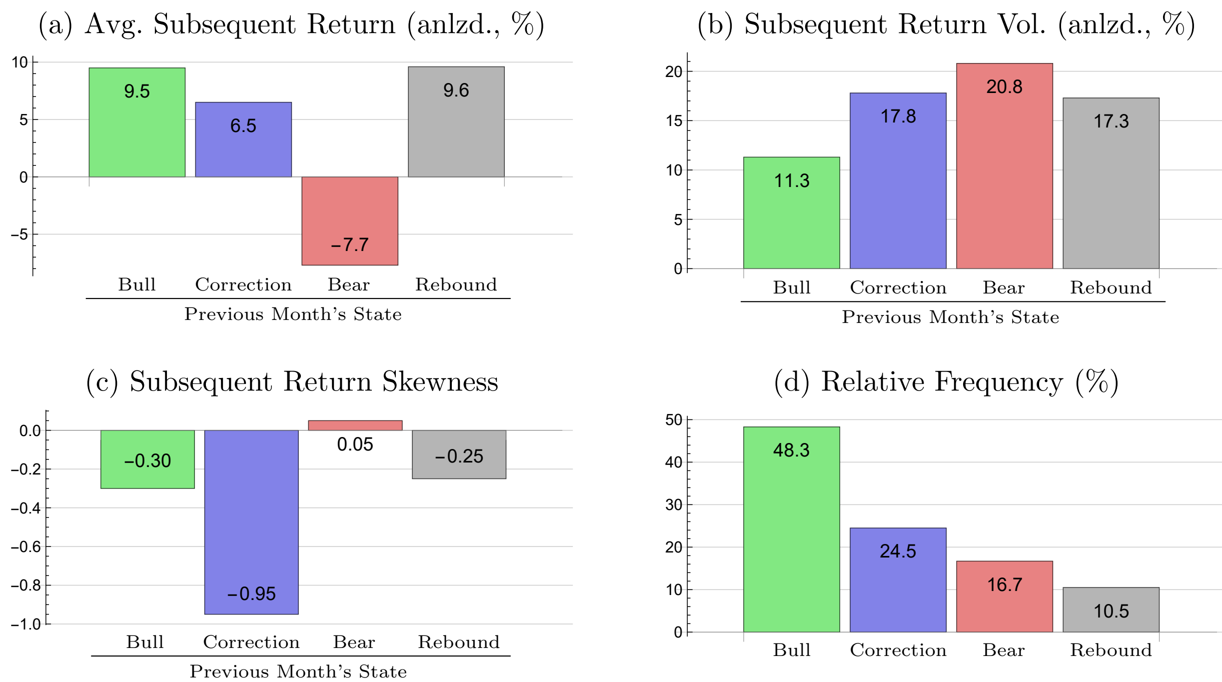
We develop a model of expected returns to examine connections between the performance of different speeds and the following unobservable variables: trend (persistence in expected returns), turning points, and noise levels in realized returns.[1] The literature has documented that TS momentum strategies based on slow momentum (i.e., the standard 12-month signal) tend to perform better than strategies based on fast momentum (e.g., the 1-month signal). Our model indicates that when the slow strategy (SLOW) outperforms the fast strategy (FAST) in the long run, it is because expected returns are relatively persistent and realized returns are relatively noisy. Yet in the short run after turning points, higher persistence translates into short-run average gains for FAST, while both higher persistence and higher noise translate into higher losses for SLOW. Thus, the same conditions that make SLOW attractive overall relative to FAST can also make it prone to suffer more around turning points and, at these times, FAST can be more effective.

These connections suggest that the union of information embedded in SLOW and FAST positions can be useful in detecting market turning points. When bets indicated by SLOW

[1]We thank an anonymous referee for suggesting this approach. See Section 2.2 for model details and analysis.

and FAST disagree, the intuition is that the market is more likely to be at a turning point. The agreement of SLOW and FAST to go long (short) is more likely to indicate the market is in the midst of an uptrend (downtrend). We find this intuition is not only supported by our model, but also consistent with the returns behavior of both the U.S. and international stock markets after each of the four different phases, or market cycles, defined by the up or down directions of each of the slow and fast momentum strategies.

Figure 1: **U.S. Stock Market Cycles**



Notes: This figure reports (a) the conditional average, (b) the conditional volatility and (c) the conditional skewness of monthly aggregate U.S. stock market excess returns, based on the market cycle state in the prior month, over the 50-year evaluation period from 1969-01 to 2018-12. A month ending at date t is classified as Bull if both the trailing 12-month return (arithmetic average monthly return), $r_{t-12,t}$, is nonnegative and the trailing 1-month return, $r_{t-1,t}$, is nonnegative. A month is classified as Correction if $r_{t-12,t} \geq 0$ but $r_{t-1,t} < 0$; as Bear if $r_{t-12,t} < 0$ and $r_{t-1,t} < 0$; and as Rebound if $r_{t-12,t} < 0$ but $r_{t-1,t} \geq 0$. The figure also reports (d) the relative frequency of these cycles over the same evaluation period. Market excess returns are U.S. excess value-weighted factor returns (Mkt-RF) from the Kenneth French Data Library.

Figure 1 summarizes, over a recent 50-year period of the U.S. stock market, the conditional behavior of the average, volatility, and skewness of returns in months *following* four market cycles and the monthly relative frequency of such states.[2] When both slow and fast momentum agree on the direction of trend, we call it a "Bull" or "Bear" state, depending on

[2]For most of the paper, we use a recent 50-year period for consistency with later analyses in which we use earlier data to warm up dynamic strategies for out-of-sample evaluation. In Internet Appendix F, we report results based on a longer evaluation period beginning in 1927, from which we draw similar inferences.

whether the agreement is to take a long or short position, respectively. These labels loosely map to phases of uptrend and downtrend: Bull states are followed by relatively high average returns with low volatility, and Bear states are followed by negative average returns with the highest relative volatility. When slow and fast momentum disagree, we call it a "Correction" state if slow momentum indicates a long position and a "Rebound" state if slow momentum indicates a short position. Similarly, these labels loosely map to the potential occurrences of turning points from uptrend to downtrend and vice-versa. Correction states are followed by deteriorating average returns, increased volatility, and severe downside outcomes—possibly a lead up to a Bear state. Rebound states are followed by average returns and skewness similar to Bull phases, but with higher volatility—possibly a lead up to a Bull state. Lastly, Corrections and Rebounds are significant in frequency and combined occur more than one-third of the time covered in our analysis.

Figure 1(a) indicates that the combination of both negative slow and negative fast momentum signals captures a state with negative expected returns (Bear states). To understand this phenomenon, we map the empirical properties of SLOW and FAST applied to the U.S. stock market into our model. We find that both persistence in expected returns and noise in realized returns are relatively high. Higher persistence in expected returns increases the potential for momentum signals to define predictive states. Because expected returns are unobservable and the noise content of observable realized returns is high, the extra observations used in slow momentum are necessary to enhance our inference of the trend. Slow momentum alone, however, is not sufficient for a clean inference because it relies on stale information.[3] The additional confirmation from fast momentum makes our combined inference more reliable, which explains the positive and negative expected returns following Bull and Bear states, respectively. Conflicting signals, associated with Correction and Rebound states, make our trend inference less reliable and, therefore, expected returns following these states align with the unconditional inference that stock market excess returns are positive.

Our market cycles also have close connections to the macroeconomy and the business cycle; in particular, Bear states, which have negative expected returns, are closely linked with high macroeconomic risk states. First, we study the behavior of 15 major macroeconomic indicators from the three broad categories of economic activity, risk, and survey-based variables. We find that surprises in these variables exhibit distinct behavior across our market cycles: bad news during Bear states, good news during Bull states, and insignificant news during Correction and Rebound states. Second, we document that Bear states are

[3]The sign of the slow signal by itself or the fast signal by itself is not enough to detect a negative expected return state. Annualized average monthly excess U.S. stock market returns following a negative average 12-month trailing return or negative 1-month trailing return, respectively, are 0.2% and 3.2% during the period 1927-07 to 2018-12.

the highest relative frequency state in the early months of recessions—occurring at twice the relative frequency of the other three states combined. Bull and Rebound states, which predict positive subsequent returns, increase to dominant relative frequencies late in periods of recession—over 90% combined relative frequency.

Momentum is one of the most pervasive factors—yet one of the most difficult to explain. In particular, negative expected returns after steep market declines in high-risk states are difficult to rationalize by most models of time-varying risk premia, which imply countercyclical expected market returns. The occurrence of Bear states early in recessions, when risk is arguably at its highest, makes the negative expected returns of these states even more puzzling. Behavioral models of slow information absorption have the potential to explain the predominance of negative expected return states, such as Bear states, early in recessions and the transition to dominance of positive expected return states, such as Bull and Rebound states, later in recessions.[4] None of these theories, however, tells us the optimal lookback horizon for a momentum strategy or how that horizon might change through time.

We analyze momentum strategies of various intermediate "speeds," which are fixed-proportion blends of SLOW and FAST, and highlight three findings. First, we show analytically and empirically that intermediate-speed momentum strategies have higher Sharpe ratios than the average Sharpe ratios of SLOW or FAST and link this behavior to the volatility of returns following turning-point states (Corrections and Rebounds). Second, we shed new light on the drivers of TS momentum alpha. Recent studies have argued that the profitability of TS momentum strategies is predominantly attributed to their static (average) tilts.[5] Empirically, we document that TS momentum strategies of all speeds on the U.S. stock market have positive average exposures (i.e., going long more often than short), yet their market betas are lower than suggested by these exposures—beta estimates nearer to zero for slow-to-intermediate speeds and negative for faster speeds. Our analysis reveals that this seeming disparity arises from the ability of these TS momentum portfolios to time volatility states. We derive a novel model-free decomposition of alpha into the sum of two components: market timing and volatility timing. Empirically, we estimate that market timing drives about two-thirds of the alpha of TS momentum strategies on the U.S. stock market over a recent 50-year period and volatility timing drives the remaining one-third. Third, we show analytically and empirically that the skewness of an intermediate-speed momentum strategy is scaled and shifted relative to the average skewness of SLOW and FAST, with the shift in the same direction as its Sharpe ratio: toward higher (more positive, less negative) skewness whenever the intermediate-speed Sharpe ratio is positive.

[4]See Gómez-Cram (2022).

[5]See Goyal and Jegadeesh (2017) and Huang et al. (2019).

We also study *dynamic* strategies with speeds that may vary each month based on observed market cycles.[6] We derive the state-dependent speed rule which yields the maximum Sharpe ratio. We estimate this strategy using historical returns and find its out-of-sample performance can efficiently track the best possible state-dependent performance.

Lastly, to test the external validity of our findings for the U.S. stock market, we examine the empirical performance of TS momentum strategies of various static and dynamic speeds across 20 international equity markets. Consistent with our analytical results, the Sharpe ratios of intermediate-speed strategies are higher than the average of the Sharpe ratios of SLOW and FAST uniformly across all markets. In addition, the Sharpe ratio of the dynamic-speed strategy is higher than the highest static-speed strategy for most countries.

1.1. The Setting

The premise of time-varying expected returns has extensive support in the literature (see, e.g., Fama and French, 1988, and Cochrane, 2011, among many). The idea of persistent trends also has support in the literature, which documents that asset returns measured over the recent past are positively correlated with future returns (Jegadeesh and Titman, 1993; 2001, Asness, 1994, Conrad and Kaul, 1998, Lee and Swaminathan, 2000, and Gutierrez and Kelley, 2008) and that these momentum effects are stable across assets and countries (Rouwenhorst, 1998, Griffin, Ji, and Martin, 2003, Israel and Moskowitz, 2012, and Asness, Moskowitz, and Pedersen, 2013).[7] Recent studies provide evidence that TS momentum strategies can successfully exploit these trends (Moskowitz, Ooi, and Pedersen, 2012; Georgopoulou and Wang, 2017; Ehsani and Linnainmaa, 2019).

Our examination of the economic linkages of market cycles contributes to the growing literature that suggests stronger connections between stock returns and fundamentals of

[6] Our approach is not to be confused with moving average crossovers. Levine and Pedersen (2016) show that moving average crossovers are essentially equivalent to *static* blends of time-series momentum strategies. Moreover, Hurst, Ooi, and Pedersen (2013) show that the returns of trend-following strategies such as Managed Futures funds and CTAs can be explained by *static* blends of time-series momentum strategies.

[7] Explanations of momentum rely on either behavioral or rational foundations. Studies on behavioral foundations include Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Odean (1998), Hong and Stein (1999, 2007), Gervais and Odean (2001), Grinblatt and Han (2002), Barberis and Shleifer (2003), Chan (2003), Cross et al. (2005), Grinblatt and Han (2005), Cross, Grinfeld, and Lamba (2006), Frazzini (2006), Shefrin (2008), Chui, Titman, and Wei (2010), Haldane (2010), Dasgupta, Prat, and Verardo (2011), Bandarchuk and Hilscher (2012), Antoniou, Doukas, and Subrahmanyam (2013), Avramov, Cheng, and Hameed (2013), and Chen and Lu (2013). Studies on rational foundations include Lo and MacKinlay (1990), Carhart (1997), Berk, Green, and Naik (1999), Ahn et al. (2002), Johnson (2002), Lewellen (2002), Allen, Morris, and Shin (2006), Watanabe (2008), Banerjee, Kaniel, and Kremer (2009), Verardo (2009), McLean (2010), Cespa and Vives (2012), Makarov and Rytchkov (2012), Vayanos and Woolley (2012, 2013), Aretz and Pope (2013), Jordan (2013), Liu and Zhang (2013), and Zhou and Zhu (2013).

the economy than previously understood. We document statistically positive and negative surprises across an array of macroeconomic indicators when conditioned on our Bull and Bear states, respectively. These macro news relationships are analogous to the findings of Albuquerque et al. (2015) that macroeconomic variables are correlated with stock market returns across their "long bull" and "long bear" episodes of the stock market, which they report last about 14.8 and 3.2 years on average, respectively. However, our results arise in shorter stock market episodes based on 12-month and 1-month momentum agreement, a novel finding. Cujean and Hasler (2017) document that TS momentum strengthens in bad economic times. We show that our Bull and Bear states, which proxy for uptrend and downtrend regimes of the stock market, concentrate in recessions, but at different periods within: Bears are more frequent early in recessions and Bulls are more frequent late in recessions. This distinction of strong uptrend and downtrend periods within bad times is consistent with the findings of Gómez-Cram (2022), which we discuss in detail in Section 3.

Huang et al. (2019) argue that predictability of 12-month TS momentum as captured by forecasting regressions does not appear to be statistically significant and, accordingly, the profitability of a diversified 12-month TS momentum portfolio can be largely attributed to its static tilt (net long positions). We arrive at a much different conclusion. Despite their positive static (average) tilts, TS momentum portfolios need not have positive betas to the underlying asset. Applying our decompositions of TS momentum beta and alpha, we document the role of innate volatility timing in TS momentum strategies, which can explain this disparity between static tilts and betas and contribute to meaningful alphas. Furthermore, appealing to static tilts as the primary explanation for TS momentum performance overlooks results from the conditional CAPM literature. Under the conditional CAPM, static tilts can be viewed as expected conditional betas, which differ from unconditional betas if any of expected returns, variances, or covariances change over time[8]—all active elements in TS momentum environments.

Goyal and Jegadeesh (2017) also argue that the net dollar exposure (i.e., static tilt) of TS momentum strategies is a key determinant of their profitability. Specifically, they propose that adding the net dollar exposure of a TS momentum portfolio to a cross-sectional (CS) momentum portfolio can reproduce TS momentum profits. This argument does not apply, however, to the case of TS momentum on a single asset, for which a CS momentum strategy is either not defined or trivial. In this paper, we focus on single-asset TS momentum.

Although our focus is on TS momentum, our approach shares some themes with the CS momentum literature. In a TS setting, we employ market cycles as Cooper, Gutierrez, and Hameed (2004), Daniel and Moskowitz (2016), and Daniel, Jagannathan, and Kim (2019)

[8]See Jagannathan and Wang (1996), Lewellen and Nagel (2006), and Boguth et al. (2011).

do in CS settings. Cooper, Gutierrez, and Hameed use a slow trailing three-year return to define market states, whereas we use slow and fast trailing returns signals.[9] Daniel and Moskowitz study CS momentum crashes and recoveries and propose a dynamic CS weighting strategy. Daniel, Jagannathan, and Kim use a two-state hidden Markov model of unobserved "turbulent" and "calm" states to predict crashes in CS momentum portfolios. We employ a finer cycle partition of observable states dictated by slow and fast momentum directions and develop a dynamic TS momentum strategy which blends slow and fast strategies using cycle-conditional information.

In a TS setting, we evaluate different horizons as Novy-Marx (2012) does in a CS setting. Novy-Marx points out that CS momentum works best at 7-to-12 month horizons. We demonstrate that shorter horizons are significant predictors of aggregate stock market returns and can augment slower signals. Liu and Zhang (2008) show in a CS momentum setting that recent winners load temporarily on the growth rate of industrial production and that macroeconomic risk can explain more than half of CS momentum profits. This work relates to our findings of distinct macroeconomic linkages to the market cycles determined by slow and fast TS momentum directions.

Our decompositions of beta and alpha for TS strategies have similarities in form to decompositions of the differences between unconditional and expected conditional versions of beta and alpha for CS strategies in the conditional CAPM literature.[10] For TS strategies, expected conditional alpha is not a useful benchmark because the conditional CAPM trivially holds for any TS strategy applied directly to excess market returns—conditional alpha is zero in every period. Moreover, decompositions in terms of conditional betas and conditional expected market returns are functions of an arbitrary conditioning information set, which is meant to represent what investors observe. In contrast, we decompose (unconditional) beta and alpha into direct functions of TS strategy weights and market returns, avoiding expressions dependent on conditioning information.

Our paper also bridges to the literature on volatility-managed (VOM) portfolios sparked by Moreira and Muir (2017) and further studied by Harvey et al. (2018). VOM portfolio alpha, by construction, is predominantly driven by volatility timing. Using our alpha decomposition, we show that TS momentum strategy alpha is driven by both market timing and innate volatility timing, without enhancements such as scaling by trailing volatility.

[9] Specifically, these authors rely on trailing three-year market returns to define two states—"up" and "down"—and forecast stock momentum consistent with the predictions of behavioral models (Daniel, Hirshleifer, and Subrahmanyam, 1998, and Hong and Stein, 1999).

[10] See Jagannathan and Wang (1996), Lewellen and Nagel (2006), and Boguth et al. (2011).

2. Momentum Speed and Market Cycles

In this section, we define a collection of TS momentum strategies of different speeds by blending elementary TS momentum strategies, SLOW and FAST, in various proportions. We also define market cycles stemming from the intersection SLOW and FAST.

2.1. Characterizing SLOW and FAST

We construct a simple framework. At date t, if the trailing 12-month excess return (arithmetic average monthly excess return), $r_{t-12,t}$, is nonnegative, then SLOW goes long one unit in the subsequent month, otherwise, it goes short one unit:

$$w_{\text{SLOW},t} := \begin{cases} +1 & \text{if } r_{t-12,t} \geq 0, \\ -1 & \text{if } r_{t-12,t} < 0. \end{cases} \tag{1}$$

If the prior 1-month return, $r_{t-1,t}$, is nonnegative then FAST goes long one unit in the subsequent month, otherwise, it goes short one unit:

$$w_{\text{FAST},t} := \begin{cases} +1 & \text{if } r_{t-1,t} \geq 0, \\ -1 & \text{if } r_{t-1,t} < 0. \end{cases} \tag{2}$$

The realized returns of SLOW and FAST for month $t+1$ are $r_{\text{SLOW},t+1} = w_{\text{SLOW},t} r_{t+1}$ and $r_{\text{FAST},t+1} = w_{\text{FAST},t} r_{t+1}$, respectively, where r_{t+1} is the realized underlying market excess return from date t to date $t+1$.

This strategy design intentionally omits more complex features to be consistent with recent TS momentum studies (Goyal and Jegadeesh, 2017, and Huang et al., 2019). We map trailing-return signals into binary weights rather than varying weights with signal magnitudes, the signal is not scaled by trailing volatility as in the work of Moskowitz, Ooi, and Pedersen (2012), and the signal does not exponentially weight past returns.[11] Furthermore,

[11] Note that volatility-targeting may constitute a phenomenon independent from TS momentum (e.g., Kim, Tse, and Wald, 2016, Moreira and Muir, 2017, and Harvey et al., 2018). In Internet Appendix C, we further compare our approach with that of Moskowitz, Ooi, and Pedersen (2012) and find that our medium-speed strategy (defined in Section 2.3) has a higher Sharpe ratio and generates alpha with respect to their 12-month strategy. In addition, trend following is often identified with moving-average crossover strategies, which exponentially weight past prices, whereas our SLOW and FAST portfolios average past returns. As shown by Levine and Pedersen (2016), moving-average crossovers and time-series momentum are close to equivalent. More generally, our focus is on blending strategies with different speeds rather than elaborating on a particular definition of a single speed. In Internet Appendix D, we further analyze TS momentum strategies with both binary and linear positions as functions of exponentially weighted moving averages of past returns.

we choose to operate on elementary strategies to facilitate our study in Section 5 of dynamic speed selection, which is a challenging and central problem in trend following.

The 12-month trailing return is a relatively slow-moving signal from month to month since 11 of 12 months of returns (92%) overlap in subsequent signals.[12] In contrast, the 1-month trailing return is a relatively fast-moving signal, having no overlapping returns between subsequent signals. We could have used horizons longer than 12 months to define slow momentum, but the signal overlap for such horizons in subsequent months would not differ materially from the high 92% overlap of the 12-month signal. Moreover, 12 months is the standard horizon analyzed in the TS momentum literature (Moskowitz, Ooi, and Pedersen, 2012, and Huang et al., 2019, among others). We could have used horizons longer than one month (e.g., 2 or 3 months) to define fast momentum. However, longer horizons have significant signal overlap (50% for the 2-month signal and 67% for the 3-month signal), which materially slows changes in their strategy positions. Moreover, the 1-month horizon is the shortest at the monthly level of data. For these reasons, we chose 12-month and 1-month horizons.[13]

2.2. A Model of Time-Varying Trend

We can view TS momentum strategies as maps of indirect estimates of expected returns (based on lagged realized returns) to portfolio positions. SLOW and FAST map signs of average trailing realized returns (as estimates of signs of expected returns) to long or short positions when such estimates are positive or negative, respectively. The accuracy of such estimates depends on the persistence of expected returns as well as the influence of noise. To examine the roles of persistence and noise, we consider a first-order autoregressive process, AR(1), for expected returns. In this model, realized returns are noisy observations of the AR(1) process. We apply SLOW and FAST to these realized returns and explore their performance based on different parameterizations of the persistence parameter and the amount of noise.

[12]Note that we do not skip the immediately lagged month as is done in some momentum strategies on individual securities in order to avoid short-term idiosyncratic reversals. Our empirical application is stock indices, for which the immediately lagged returns tend not to exhibit such reversals.

[13]The empirical evidence pertaining the U.S. stock market supports our rationale. Over our 50-year evaluation period, the 3-month TS momentum portfolio showed a higher correlation to 12-month momentum than to 1-month momentum (43% and 34%, respectively), whereas the correlation of the 24-month momentum portfolio to 12-month momentum was above 60%. In addition, the turnover of the 1-month signal was twice as large as that of the 3-month signal, another indication of the differences underlying two apparently similar strategies.

Specifically, the unobservable monthly expected excess return process, $\{\mu_t\}$, follows

$$\mu_t = c + \varphi\mu_{t-1} + \varepsilon_t, \tag{3}$$

where c is a scalar, φ is a persistence scalar, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed mean-zero normal random variables each having variance $\sigma_\varepsilon^2 > 0$. We assume that persistence is positive ($\varphi > 0$) to capture the notion of trend, consistent with our description in Section 1. For example, if last month's underlying mean return were positive ($\mu_{t-1} > 0$), then this month's mean return is more likely to also be positive ($\mu_t > 0$) than negative. We also assume $\varphi < 1$ so that the process is weak-sense stationary, i.e., the expected return process is time varying but its unconditional mean and autocovariances are time invariant. Under these assumptions, by standard AR(1) properties, μ_t has unconditional expected value $\mathbf{E}[\mu_t] = \frac{c}{1-\varphi}$ and unconditional variance $\mathbf{Var}[\mu_t] = \frac{\sigma_\varepsilon^2}{1-\varphi^2} > 0$. Let realized excess returns in each month, r_t, be random variables centered around μ_t, but observed with noise:

$$r_t = \mu_t + z_t, \tag{4}$$

where $\{z_t\}$ is a sequence of independent and identically distributed mean-zero normal random variables, and where each z_t has variance $\sigma_z^2 > 0$ and is jointly normal with ε_t having correlation coefficient $\lambda \in (-1, 1)$. Then, the unconditional average return is $\mathbf{E}[r_t] = \frac{c}{1-\varphi}$ and the unconditional return variance is $\mathbf{Var}[r_t] = \sigma_r^2$, where we define $\sigma_r = \sqrt{\frac{\sigma_\varepsilon^2}{1-\varphi^2} + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z}$.

We focus on the case in which $c = 0$ (unconditional mean excess return is zero) to ensure that our analysis is nontrivial for all choices of persistence and noise parameters. In contrast, if unconditional mean excess returns were positive (negative), then for sufficiently small noise σ_z, realized returns would be predominantly positive (negative), and therefore, a profitable TS strategy would be trivially almost always long (short). Moreover, for $\mathbf{E}[\mu_t] = 0$, excess returns on TS momentum strategies are entirely alpha. Therefore, their performance is exclusively a measure of their ability to time return opportunities because their average positions are unprofitable in expectation. Thus, expected returns are time varying but zero on average, deviating away from zero and remaining away with some persistence before eventually crossing zero to have the opposite sign, where they remain with some persistence until eventually crossing back, and so on.

Finally, define the noise ratio, θ, as the fraction of variance in realized return r_t contributed by noise z_t:

$$\theta = \frac{\mathbf{Cov}[z_t, r_t]}{\mathbf{Var}[r_t]} = \frac{\sigma_z^2 + \lambda\sigma_\varepsilon\sigma_z}{\sigma_r^2} = \frac{\sigma_z^2 + \lambda\sigma_\varepsilon\sigma_z}{\frac{\sigma_\varepsilon^2}{1-\varphi^2} + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z}. \tag{5}$$

If θ is small, then realized and expected returns are close to each other and, therefore, realized returns are more informative about the underlying trend (i.e., the sign of expected returns). If θ is large, then realized returns are dominated by noise and thereby less informative about the underlying trend.

Proposition 1 characterizes in closed-form the Sharpe ratios of our elementary trend-following strategies, SLOW and FAST (defined in Section 2.1), in terms of persistence φ and noise ratio θ.

Proposition 1. *The (annualized) Sharpe ratios of FAST and SLOW, respectively, can be expressed in terms of persistence φ and noise ratio θ as follows:*

$$\mathbf{Sharpe}[r_{FAST,t}] = \frac{\sqrt{\frac{24}{\pi}}(1-\theta)\varphi}{\sqrt{1-\frac{2}{\pi}(1-\theta)^2\varphi^2}}, \tag{6}$$

$$\mathbf{Sharpe}[r_{SLOW,t}] = \frac{\sqrt{\frac{24}{\pi}}(1-\theta)\varphi}{\sqrt{\frac{(1-\theta)[12(1-\varphi^2)-2\varphi(1-\varphi^{12})]+12\theta(1-\varphi)^2}{(1-\varphi^{12})^2}-\frac{2}{\pi}(1-\theta)^2\varphi^2}}. \tag{7}$$

Moreover, the denominators in (6) and (7) are proportional to the standard deviations of the respective strategies and, therefore, are positive.

Proof. See Internet Appendix A. □

Proposition 1 indicates that both strategies have positive Sharpe ratios if there is some persistence ($\varphi > 0$) and if $\theta < 1$. It is possible in general, however, that θ is above one for certain values of model parameters (φ, σ_ε, σ_z and λ) if realized return noise and expected return innovations are negatively correlated (i.e., $\lambda < 0$). Because we are studying momentum strategies, we are interested in those combinations of parameters for which conditional expected returns (given all past return observations) are positively related to past returns. Under this condition, Proposition 2 shows that we can focus our attention on values of $\theta < 1$. Furthermore, if we impose the mild condition that the variance of noise in realized returns is at least half as large as the variance of innovations in expected returns ($\sigma_z^2 \geq \frac{1}{2}\sigma_\varepsilon^2$), then Proposition 2 indicates that we can focus our attention on values of θ in the unit interval. In Section 4.1.3, we find that the inferred value of the noise ratio for U.S. stock market returns is consistent with this implication that θ lies in the unit interval.

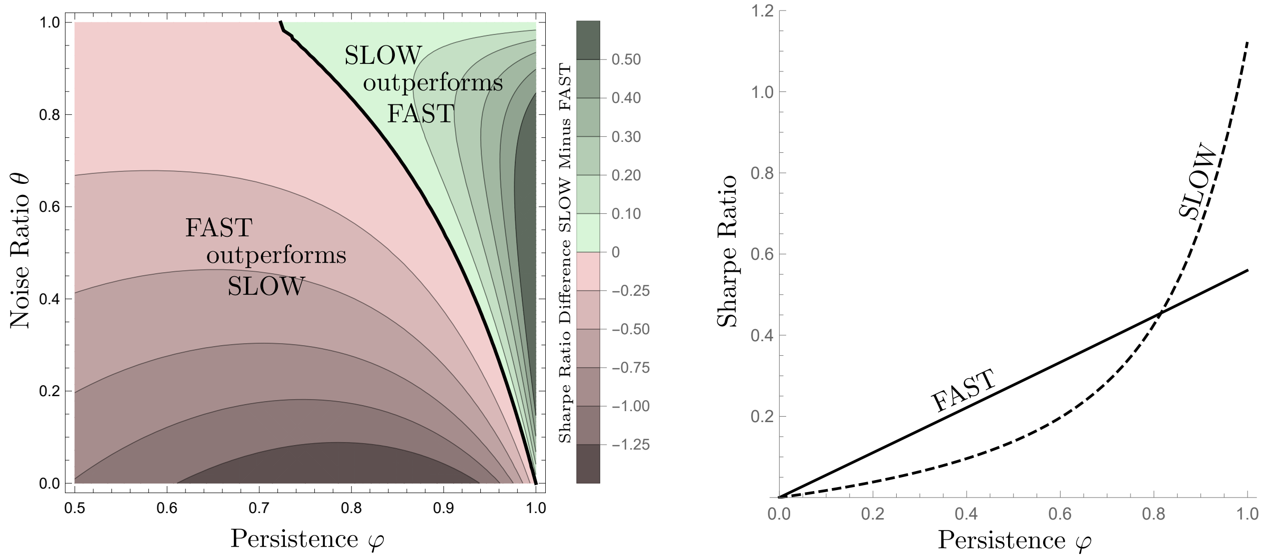
Proposition 2. *If $\mathbf{E}[r_t|r_{t-1}, r_{t-2}, \ldots]$ is positively related to past returns $r_{t-1}, r_{t-2}, \ldots$ (i.e., $\frac{\partial \mathbf{E}[r_t|r_{t-1}, r_{t-2}, \ldots]}{\partial r_{t-i}} > 0$ for all $i = 1, 2, \ldots$), then $\theta < 1$. If, in addition, we have $\sigma_z^2 > \frac{1}{2}\sigma_\varepsilon^2$, then $\theta > 0$.*

Proof of Proposition 2. See Internet Appendix A. □

Figure 2: **SLOW and FAST Absolute and Relative Performance**

(a) Relative Performance as Function of Persistence and Noise

(b) Performance as Function of Persistence with Relatively High Noise ($\theta = 0.8$)



Notes: Figure (a) plots contour levels of the difference between the Sharpe ratios of SLOW and FAST for various combination of persistence levels and noise ratios. Combinations of persistence and noise ratios in the upper-right region above the bold curve correspond to a higher Sharpe ratio for SLOW than FAST. Figure (b) plots Sharpe ratios of SLOW (dashed) and FAST (solid) for various persistence levels and at noise ratio $\theta = 0.8$.

Under the conditions of Proposition 2, the Sharpe ratios of SLOW and FAST are continuous functions over the bounded set of persistence ($0 < \varphi < 1$) and noise ($0 < \theta < 1$) parameters, so we can analyze these Sharpe ratios graphically on the unit square: $(\varphi, \theta) \in [0, 1] \times [0, 1]$. In Figure 2, we compute and plot the absolute and relative levels of the Sharpe ratios expressed in Proposition 1 for a range of persistence and noise ratios. Figure 2(a) is a contour plot of the difference between the Sharpe ratios of SLOW and FAST for all noise levels and a range of persistence levels.[14] This figure reinforces the intuition that for sufficiently high persistence, SLOW outperforms FAST, and for sufficiently low noise, FAST outperforms SLOW. The literature has documented that SLOW tends to work better than FAST for many different asset classes when measured by overall risk-adjusted performance. From Figure 2(a), we infer that such relative performance would be consistent with a mar-

[14] We plot persistence above 0.5, however, the patterns evident in the plot extend continuously for lower persistence levels. We focus on the higher persistence levels to highlight the region where relative performance switches.

ket in which the excess returns exhibit both relatively high mean persistence and realized noise.[15] Figure 2(b) plots the Sharpe ratios of SLOW and FAST for various persistence levels at a relatively high noise ratio. We infer from this figure that for noisy realized returns the outperformance of SLOW over FAST accelerates as mean persistence increases.

2.2.1. Model Turning Points

How do turning points affect the performance of SLOW and FAST? Although expected returns are unobservable in practice, we can use our model to analyze the hypothetical impact of (model) turning points (i.e., changes in signs of model conditional means). Uptrend corresponds to periods of positive expected returns, $\{\mu_t : \mu_t > 0\}$, and downtrend corresponds to periods of negative expected returns, $\{\mu_t : \mu_t < 0\}$. Turning points, therefore, occur whenever successive values of $\{\mu_t\}$ change sign.[16] For developing intuition from this hypothetical exercise, we set $\lambda = 0$, which implies that innovations (ε_t, z_t) are uncorrelated and that $\theta \in (0, 1)$.

Proposition 3. *Let $\lambda = 0$. The conditional expected returns of FAST and SLOW, respectively, given the most recent two (unobservable) consecutive return means are as follows:*

$$\mathbf{E}[r_{FAST,t}|\mu_{t-1}, \mu_{t-2}] = \varphi\mu_{t-1} \cdot \left[2\Phi\left(\frac{\mu_{t-1}}{\sigma_r\sqrt{\theta}}\right) - 1\right], \tag{8}$$

$$\mathbf{E}[r_{SLOW,t}|\mu_{t-1}, \mu_{t-2}] = \varphi\mu_{t-1} \cdot \left[2\Phi\left(\frac{\frac{1}{12}\left(\mu_{t-1} + \frac{1-\varphi^{11}}{1-\varphi}\mu_{t-2}\right)}{\sigma_r\sqrt{\frac{(1-\theta)}{12}\left[\frac{1+\varphi}{1-\varphi} + \frac{1}{12} - \frac{2(1-\varphi^{12})+(1-\varphi^{11})^2}{12(1-\varphi)^2}\right] + \frac{\theta}{12}}}\right) - 1\right], \tag{9}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. If $\mu_{t-1} \neq 0$, then $\mathbf{E}[r_{FAST,t}|\mu_{t-1}, \mu_{t-2}] > 0$. If μ_{t-1} and μ_{t-2} are each nonzero with different signs and $|\mu_{t-1}| < \frac{1-\varphi^{11}}{1-\varphi}|\mu_{t-2}|$, then

$$\mathbf{E}[r_{FAST,t}|\mu_{t-1}, \mu_{t-2}] > 0 > \mathbf{E}[r_{SLOW,t}|\mu_{t-1}, \mu_{t-2}]. \tag{10}$$

Proof. See Internet Appendix A. □

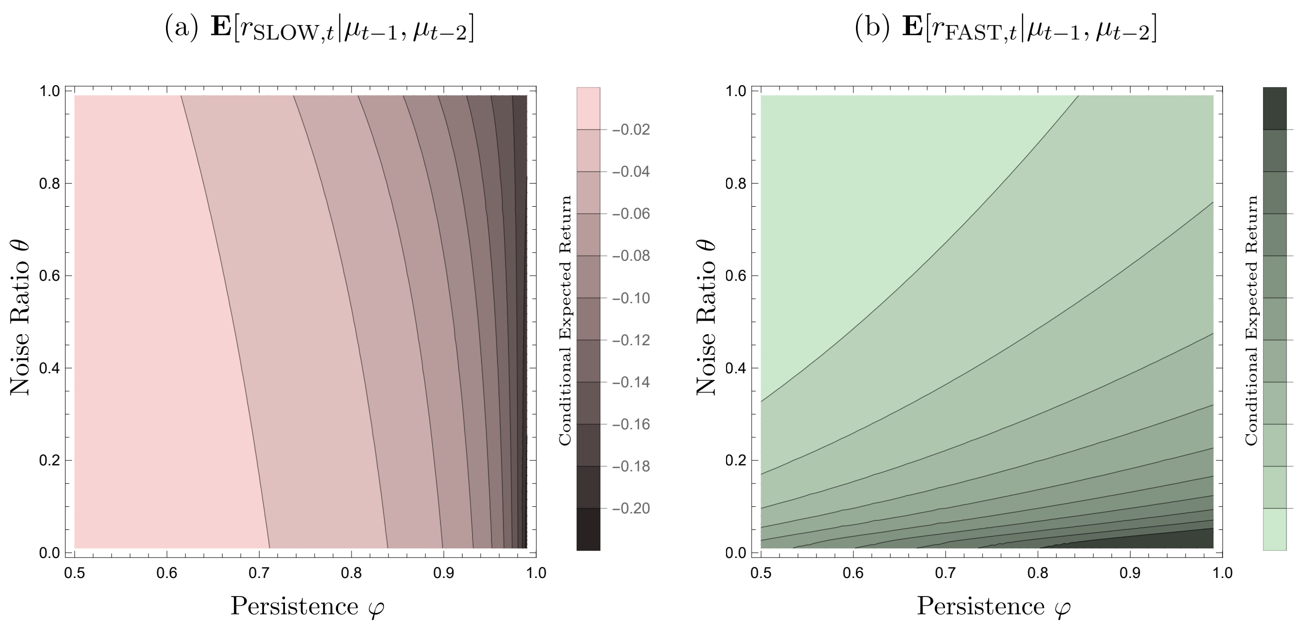
Proposition 3 provides closed-form expressions for conditional expected excess returns of SLOW and FAST as functions of the expected returns of the previous two periods. The

[15] In Section 4.1.3, we estimate the model-implied values of mean persistence and realized noise for U.S. stock market returns and find that both are relatively high.

[16] In subsequent sections, we develop a simple measure of turning points based on observable realized returns, consistent with our analysis in this section.

condition that μ_{t-1} and μ_{t-2} are each nonzero with different signs means we have encountered a (model) turning point. Moreover, (10) indicates that, after turning points, FAST tends to be profitable and SLOW tends to incur losses, on average. The condition $|\mu_{t-1}| < \frac{1-\varphi^{11}}{1-\varphi}|\mu_{t-2}|$ derives from the conditional expected value of the 12-month signal, $\frac{1}{12}\left(\mu_{t-1} + \frac{1-\varphi^{11}}{1-\varphi}\mu_{t-2}\right)$, which appears in (9). This condition requires that the magnitude of the conditional expected value of the 1-month signal, μ_{t-1}, was not excessively large compared to μ_{t-2}, so that a single month's return was not expected to be large enough to flip the sign of the slow signal. For example, if $\varphi = 0.8$, and if two month's ago the expected return was half-a-standard-deviation above zero, $\mu_{t-2} = 0.5\sigma_r$, then these conditions require that the most recent expected return turned negative but not more than about two-standard-deviations below zero, $\mu_{t-1} \geq -2.3\sigma_r$. Any turning point with opposite-sign expected returns of equal magnitude (e.g., $-\mu_{t-1} = \mu_{t-2} = 0.5\sigma_r$) satisfies this condition. Again, we can analyze the results of Proposition 3 graphically.

Figure 3: **Expected Returns After Model Turning Points**



Notes: This figure shows contour plots of the conditional expected returns of (a) $\mathbf{E}[r_{\text{SLOW},t}|\mu_{t-1}, \mu_{t-2}]$, and (b) $\mathbf{E}[r_{\text{FAST},t}|\mu_{t-1}, \mu_{t-2}]$, following a (model) turning point ($\text{sign}(\mu_{t-1}) \neq \text{sign}(\mu_{t-2})$) over various values of mean persistence φ and noise ratios θ for the parameters $-\mu_{t-1} = \mu_{t-2} = 0.3\sigma_r$.

Figure 3 plots the conditional expected returns of (a) SLOW and (b) FAST after a turning point as given in Proposition 3 for various values of mean persistence and noise, with opposite-sign expected returns at 0.3-standard-deviation magnitude each, $-\mu_{t-1} = \mu_{t-2} = 0.3\sigma_r$. After the turning point, expected returns are negative for SLOW and positive for FAST under all noise and persistence levels.[17] Figure 3(a) shows that the conditions where SLOW performs well overall correspond with the conditions where SLOW struggles most after turning points: relatively high mean persistence and realized noise. Therefore, the acceleration of overall performance of SLOW with higher persistence under relatively high noise (as shown in Figure 2(b)) tends to come at the cost of exacerbating momentum crashes. Figure 3(b) shows that with lower noise levels or higher persistence levels, FAST delivers increasingly positive returns after turning points.

Opposite signs of expected returns of SLOW and FAST after a turning point imply that these strategies tend to take opposite positions after turning points. This observation indicates that the information in the union of SLOW and FAST strategy positions can be useful in detecting turning points. We further explore these connections in subsequent sections.

We fully recognize that the model assumptions we have employed do not hold in the real world. For example, returns are generally not normally distributed and expected returns need not follow AR(1). Throughout the remainder of the paper, we do not maintain the distributional assumptions introduced in this section unless explicitly referring back to this model. Instead, we employ the mild standing assumption that moments in analytical expressions and derivations exist and second moments are non-zero. In this sense, we characterize our subsequent results as being model free.

2.3. Characterizing Intermediate Speeds and Market Cycles

We define a continuum of intermediate strategies with signal speeds between SLOW and FAST by

$$w_t(a) := (1-a)w_{\mathrm{SLOW},t} + a\, w_{\mathrm{FAST},t}, \tag{11}$$

$$r_{t+1}(a) := w_t(a)r_{t+1} = (1-a)r_{\mathrm{SLOW},t+1} + a\, r_{\mathrm{FAST},t+1}, \tag{12}$$

where the speed parameter $a \in [0,1]$ is a scalar. At $a=0$, the speed is slow: $w_t(0) = w_{\mathrm{SLOW},t}$. Here, if the preceding 12-month return is positive, the strategy is long one unit even when the most recent month's return is negative. At $a=1$, the speed is fast: $w_t(1) = w_{\mathrm{FAST},t}$. Here, any change of sign in the most recent month's return results in a strategy position change. For $a \in (0,1)$, the speed is intermediate and, in particular, at $a = \frac{1}{2}$ we call the

[17] We plot persistence above 0.5, however, the signs of expected returns in the plot remain the same for lower persistence levels.

speed "medium" (MED):

$$w_{\text{MED},t} := w_t\left(\tfrac{1}{2}\right) = \tfrac{1}{2}w_{\text{SLOW},t} + \tfrac{1}{2}w_{\text{FAST},t}, \tag{13}$$

with $r_{\text{MED},t+1} := w_{\text{MED},t}r_{t+1}$.

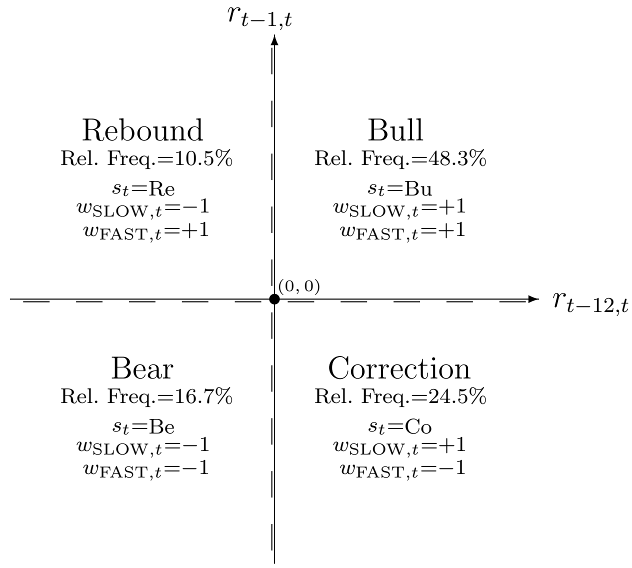
Blending slow and fast momentum portfolios to form intermediate-speed strategies is a key analytical choice. By doing so, we embed sensitivity to periods of disagreement between the slow and fast strategies, which potentially signal turning points. For example, when SLOW indicates a long position (+1) and FAST indicates a short position (−1), then the intermediate-speed strategy with speed $a = \frac{3}{4}$ takes a lower magnitude short position ($-\frac{1}{2} = \frac{1}{4}(+1) + \frac{3}{4}(-1)$). The MED portfolio is out of the market altogether ($0 = \frac{1}{2}(+1) + \frac{1}{2}(-1)$). In contrast, strategies that go long or short based on the sign of some horizon of trailing returns $r_{t-k,t}$ do not scale down their positions when signals of longer and shorter horizons disagree. For example, the sign of the six-month signal $r_{t-6,t}$ could take several months to reflect a turning point in trend, during which time the strategy is fully long or short.[18]

Figure 4 maps the intersection of slow and fast signals to one of four observable *market cycles* or states. We use s_t to denote the state at date t. To capture the notion of a sustained upward trend, we label a month ending at date t as Bull if both the 12-month and 1-month trailing returns are nonnegative ($r_{t-12,t} \geq 0$ and $r_{t-1,t} \geq 0$, on the upper-right quadrant of the diagram). We label the other three quadrants of the diagram (going clockwise) of trailing 12-month and 1-month returns as Correction, Bear, and Rebound, respectively. We will also use the abbreviations Bu, Co, Be, and Re to denote the respective market states (e.g., $s_t = \text{Re}$ for a Rebound). Admittedly, the labels assigned to these cycles have a relatively loose meaning. In particular, short-lived market gyrations could lead to brief and unintuitive correction or rebound classifications, which could be addressed by more sophisticated classification rules. Yet, because our simple cycle definitions are one to one with the four possible slow/fast strategy position pairs for the subsequent month, these cycles provide an exact ex-ante specification of the slow and fast momentum strategies, and by extension, all intermediate speed strategies in (11) and (12).

Over our 50-year evaluation period, Bull months have been the most common with a relative monthly frequency of 48.3%, reflecting the average positive risk premium offered by the U.S. stock market. Bear months are relatively uncommon—approximately one-sixth of the time (16.8%)—whereas Correction and Rebound months amount to the remaining

[18]In Internet Appendix E, we compare the performance of the 6-month TS momentum strategy, MOM6, with our intermediate-speed strategies and find that across different metrics MOM6 performs similarly to the SLOW strategy (i.e., MOM12). This result is not surprising because 5 of 6 months (83%) in the 6-month signal $r_{t-6,t}$ overlap in consecutive months. This overlap is close to that of the 12-month signal $r_{t-12,t}$ (92%).

Figure 4: **Stock Market Cycles as a Function of Momentum**



Notes: This diagram defines the four cycles of an asset's price trajectory as a function of SLOW and FAST momentum positions. A month ending at date t is classified as Bull if both the trailing 12-month return (arithmetic average monthly return), $r_{t-12,t}$, is nonnegative, and the trailing 1-month return, $r_{t-1,t}$, is nonnegative. A month is classified as Correction if $r_{t-12,t} \geq 0$ but $r_{t-1,t} < 0$; as Bear if $r_{t-12,t} < 0$ and $r_{t-1,t} < 0$; and as Rebound if $r_{t-12,t} < 0$ but $r_{t-1,t} \geq 0$. Also, the diagram reports the relative frequency of monthly aggregate U.S. stock market cycles over the 50-year period from 1969-01 to 2018-12. Market returns are U.S. excess value-weighted factor returns (Mkt-RF) from the Kenneth French Data Library.

34.8% of the months. In other words, about once every three months, on average, SLOW and FAST suggest a different position in the stock market. These phases will be the focus of much of our analysis.

Market cycles conveniently map to the properties of SLOW and FAST for analysis. For instance, consider the strategy in (13) that employs the medium-speed signal, which is the equally-weighted average of the slow and fast strategies. Conditioning on the various market cycles as of date t, the equally-weighted strategy return at date $t+1$, $r_{\text{MED},t+1} = w_{\text{MED},t} r_{t+1}$, has long-run expected value:

$$\begin{aligned}
\mathbf{E}[r_{\text{MED},t+1}] &= \mathbf{E}[\tfrac{1}{2}\left(w_{\text{SLOW},t} + w_{\text{FAST},t}\right) r_{t+1}] \\
&= \mathbf{E}[\tfrac{1}{2}\left(1+1\right) r_{t+1} | s_t = \text{Bu}]\mathbf{P}[s_t = \text{Bu}] + \mathbf{E}[\tfrac{1}{2}\left(1-1\right) r_{t+1} | s_t = \text{Co}]\mathbf{P}[s_t = \text{Co}] \\
&\quad + \mathbf{E}[\tfrac{1}{2}\left(-1-1\right) r_{t+1} | s_t = \text{Be}]\mathbf{P}[s_t = \text{Be}] + \mathbf{E}[\tfrac{1}{2}\left(-1+1\right) r_{t+1} | s_t = \text{Re}]\mathbf{P}[s_t = \text{Re}] \\
&= \mathbf{E}[r_{t+1} | \text{Bu}]\mathbf{P}[\text{Bu}] - \mathbf{E}[r_{t+1} | \text{Be}]\mathbf{P}[\text{Be}], \qquad (14)
\end{aligned}$$

where $\mathbf{P}$ represents the probability measure. This framework can be extended beyond the calculation of expected returns. In particular, the nature of market cycles also has an impact on the Sharpe ratio of a trend strategy (see Section 4.1) and its skewness (see Section 4.3). Moreover, it opens the intriguing possibility of "speed timing" by strategically adjusting the speed parameter a across market states, a challenge so far not tackled, to the best of our knowledge, by any existing paper (see Section 5).

3. The Economic Linkages of Market Cycles

We further motivate the relevance of our cycles by evaluating their economic linkages.

3.1. Market Cycles and Macroeconomic News

We study three groups of macro indicators. The first group includes four indices assembled by the Federal Reserve Bank of Chicago that cover major components of the U.S. economy—production, consumption, employment, and sales—and monetary policy shocks estimated by Gertler and Karadi (2015). The second group includes five risk-related indicators: (i) the National Financial Conditions Index (NFCI) by the Chicago Fed; (ii) the Pastor and Stambaugh (2003) liquidity innovation measure (PS Liquidity); (iii) the Treasury–Eurodollar (TED) spread; (iv) the liquidity metric for the Treasury bond market (Noise) developed by Hu, Pan, and Wang (2013); and (v) the high-volatility/low-volatility valuation spread (Vol Spread) of Pflueger, Siriwardane, and Sunderam (2018). The last group of variables, which are based on media or survey data, includes: (i) the daily news-based Economic Policy Uncertainty Index (Policy) by Baker, Bloom, and Davis (2016); (ii) the University of Michigan Consumer Sentiment Index; (iii) the Purchasing Manager's Index (PMI) compiled by the Institute for Supply Management; (iv) the current-quarter recession probabilities by the Survey of Professional Forecasters (SPF); and (v) the next-quarter corporate profit expectations also from the SPF. In total, we collected 15 time series of macro indicators.

In Table 1, we map the four market cycles contemporaneously to innovations in the 15 macro indicators. These innovations are the residuals from individual autoregressive models, whose order is determined by the Bayesian Information Criterion.[19]

Our evidence reveals a common theme. During Bull and Bear markets, when SLOW and FAST agree, we observe significant positive and negative macroeconomic shocks, respectively. During Corrections and Rebounds, innovations are statistically insignificant. A potential shift or stall in the macro environment, captured by neither significant positive nor negative average innovations, tends to coincide with turning points in the cycles of the stock

[19]The only two exceptions to this approach consist of the liquidity innovation metric (PS Liquidity) constructed by Pastor and Stambaugh (2003) and the monetary policy shocks (MP Shock) estimated by Gertler and Karadi (2015).

Table 1: **Macro Innovations t-Statistics by Market Cycle**

Economy News					
Market Cycle	Production	Consumption	Employment	Sales	MP Shock
Bull	1.53	2.25	2.30	2.63	0.41
Correction	1.46	−0.29	1.06	−0.13	1.72
Bear	−3.15	−3.24	−4.06	−3.73	0.58
Rebound	−1.41	−0.23	−1.30	−0.65	−3.96
Risk News					
Market Cycle	NFCI*	PS Liquidity	TED*	Noise*	Vol Spread
Bull	1.72	4.63	1.67	1.53	3.80
Correction	−1.21	−1.67	−1.67	−0.52	−1.02
Bear	−3.41	−6.22	−3.03	−4.29	−4.19
Rebound	2.32	0.43	2.00	2.15	−0.86
Survey News					
Market Cycle	Policy*	Consumer Sentiment	PMI	SPF Recession*	SPF Corporate Profits
Bull	2.09	1.64	3.48	2.30	1.80
Correction	0.06	−0.10	−0.77	−1.21	−0.61
Bear	−4.04	−3.30	−3.21	−3.00	−2.83
Rebound	−0.28	0.09	−2.39	0.84	0.76

Notes: This table reports the t-statistics of the innovations in 15 macro series conditional on four contemporaneous states of the U.S. stock market (Bull, Correction, Bear, and Rebound). These innovations are the residuals from individual autoregressive models, whose order is determined by the Bayesian Information Criterion. The only two exceptions to this approach consist of the Pastor and Stambaugh (2003) liquidity innovation metric (PS Liquidity) and the monetary policy shocks (MP Shock) estimated by Gertler and Karadi (2015). The symbol * indicates that for the selected series the sign is reversed, so that a negative shock can be interpreted as a negative outcome. A month ending at date t is classified as Bull if both the trailing 12-month return (arithmetic average monthly return), $r_{t-12,t}$, is nonnegative, and the trailing one-month return, $r_{t-1,t}$, is nonnegative. A month is classified as Correction if $r_{t-12,t} \geq 0$ but $r_{t-1,t} < 0$; as Bear if $r_{t-12,t} < 0$ and $r_{t-1,t} < 0$; and as Rebound if $r_{t-12,t} < 0$ but $r_{t-1,t} \geq 0$. The starting and ending dates of the macro series vary according to availability from the original sources, and they are listed in Internet Appendix B. The stock-market returns are U.S. excess value-weighted factor returns (Mkt-RF) from the Kenneth French Data Library.

market. Only monetary policy shocks paint a somewhat different picture, with surprise cuts associated with Rebound phases and hikes with Correction phases.

3.2. Market Cycles and Business Cycles

In Figure 5, we plot the relative frequencies of Bull, Correction, Bear, and Rebound months in early and late periods of NBER recessions and expansions over the most recent

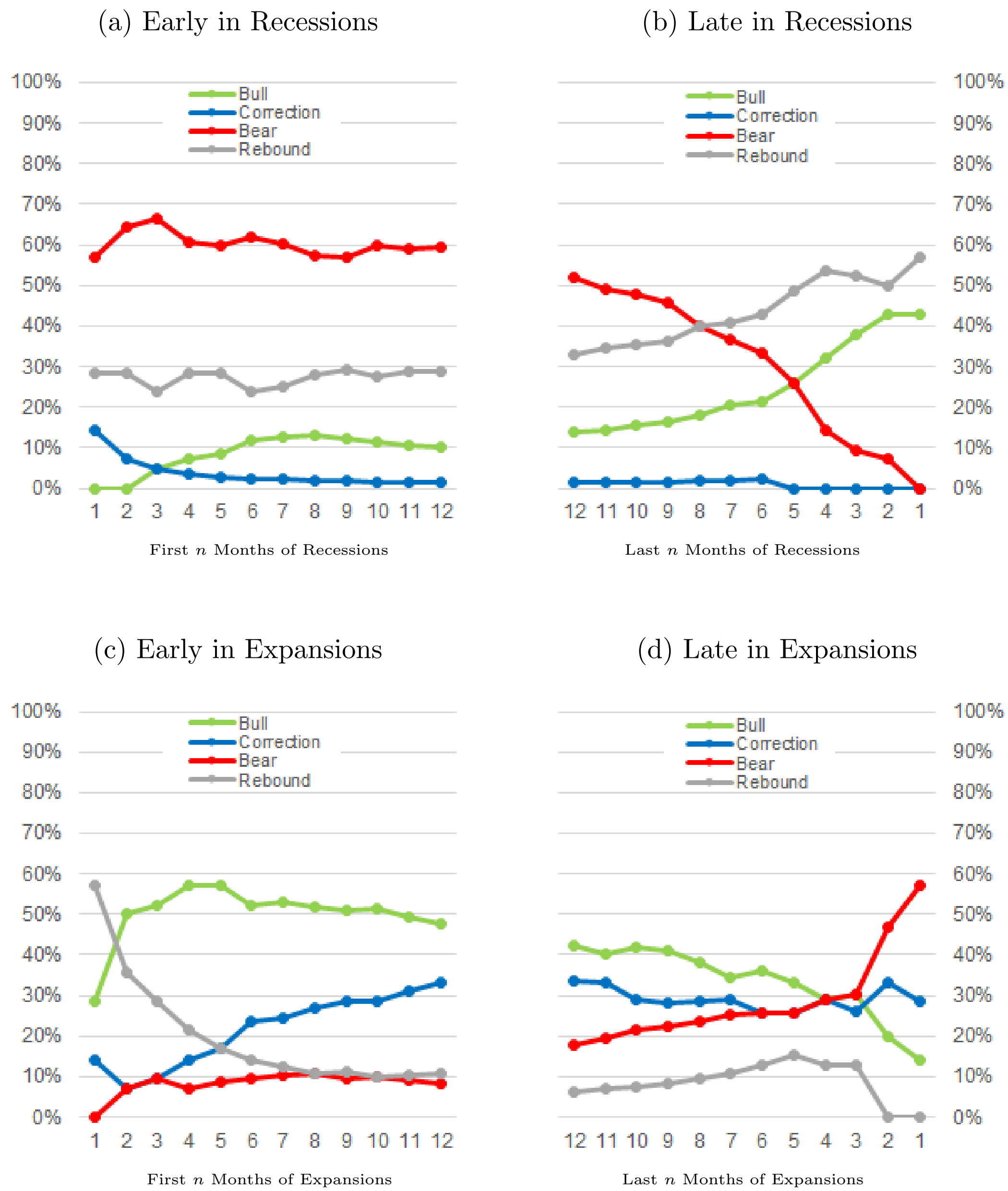
50-year period of our sample.[20] Bear states are most frequent early in recessions. Their frequency plummets toward the end of recessions, stays low during early periods of expansions, and ramps up again at the tail end of expansions. Bull states are most infrequent early in recessions. Their frequency ramps up toward the end of recessions, stays high and continues to rise early in expansions, and gradually decreases toward the tail ends of expansions. Correction states are rare during recessions. Their frequency increases during expansions through to the cusp of the next recession. Rebound states are second-most frequent early in recessions and ramp up toward the end of recessions. Rebound state frequency sharply drops early in expansions and remains infrequent through to the cusp of the next recession.

This evidence connects the empirical fact we documented earlier in Figure 1, of predictably negative returns following Bear market states and predictably positive returns following Bull (and Rebound) market states, to the central finding of Gómez-Cram (2022) that returns are predictably negative early in recessions and rise to relatively high positive levels late in recessions. First, consider Figure 5(a), which shows that Bear market states are most frequent early in recessions, whereas Bull market states are least frequent. In particular, Bears account for two-thirds of market states in the first three months of recessions, 14 times the relative frequency of Bulls during the same early recession period. Because Bear states, by definition, mean that the most recent month's return is negative, a high relative frequency of Bear states concentrated early in recession periods implies sequences of predictably negative returns during these early periods. Second, consider Figure 5(b), which shows that toward the end of recessions the relative frequency of Bear market states plummets while the relative frequencies of Bull and Rebound states ramp up to dominate the other states—over 90% relatively frequency in the last three months. Because both Bull and Rebound states, by definition, mean the the most recent month's return is nonnegative, the dominance of these states concentrated late in recessions implies the strongest periods of predictably positive returns occur late in recessions.

This evidence is consistent with the findings of Cujean and Hasler (2017) and Gómez-Cram (2022) that time-series momentum is concentrated in bad economic times. Consider the relative frequencies of negative- and positive-return months. Negative-return months correspond to the occurrence of Bear or Correction states and positive-return months correspond to the occurrence of Bull or Rebound states. In recessions, Bears and Corrections are at their highest combined frequency (over 70% early in recessions) and Bulls and Rebounds are at their highest combined frequency (over 90% late in recessions). The concentrated sequences of these pairs of states occurring in recessions suggest that negative returns clus-

[20]We obtain qualitatively similar results (unreported) over the longer sample based on data beginning in 1926.

Figure 5: **Relative Frequencies of U.S. Market Cycles in Early and Late Stages of NBER Recessions and Expansions**



Notes: This figure plots the relative frequencies of market-cycle months during the (a) first and (b) last n months of NBER recessions and (c) the first and (d) last n months of NBER expansions from 1969-01 to 2018-12. A month ending at date t is classified as either an NBER recession or expansion month based on the NBER recession-indicator time series from FRED. A month ending at date t is classified as Bull if both the trailing 12-month return (arithmetic average monthly return), $r_{t-12,t}$, is nonnegative and the trailing 1-month return, $r_{t-1,t}$, is nonnegative. A month is classified as Correction if $r_{t-12,t} \geq 0$ but $r_{t-1,t} < 0$; as Bear if $r_{t-12,t} < 0$ and $r_{t-1,t} < 0$; and as Rebound if $r_{t-12,t} < 0$ but $r_{t-1,t} \geq 0$. Market excess returns are U.S. excess value-weighted factor returns (Mkt-RF) from the Kenneth French Data Library.

ter the most early in recessions and positive returns cluster the most late in recessions. This evidence suggests that recessions tend to possess the strongest periods of both negative momentum (early) and positive momentum (late).

3.3. Economic Implications

Given the close connections between market cycles and macroeconomic risk states, predictably negative returns after Bear states pose a puzzle for traditional macrofinance asset-pricing models. These models generally predict countercyclical risk premia: compensation for risk taking is positive and relatively high in bad times. As documented in Section 3.1, macro news is generally negative in Bear states, indicating the tendency of Bear states to coincide with bad times. Moreover, as discussed in Section 3.2, the earliest periods in recessions, in which the depth of the recession has not yet been realized, are arguably the periods of greatest economic uncertainty and, therefore, should be the periods of highest expected returns under these models. Thus, the high relative frequency of Bear states in early recession periods is especially difficult for these models to rationalize.

These findings parallel those of Gómez-Cram (2022), who constructs a state-space model estimated from a collection of macroeconomic variables to show that returns are predictably negative in the early periods of recessions and rise to high positive expectations in later periods of recessions. We arrive at these puzzling facts in a simpler way using only trailing stock market returns, but the central implications for models of time-varying risk premia are shared with those explored by Gómez-Cram. Specifically, Gómez-Cram analyzes rational models including the conditional CAPM, several popular macrofinance models that relate economic state variables to the equity premium through the main asset-pricing equation, and several models that impose economic restrictions based on the present-value identity.[21] Implications of these models include: stock market returns are countercyclical, early periods of recessions support uptrend rather than downtrend, and strategies that reduce exposure to risk during bad times do not generate positive unconditional alpha. These implications are incompatible with our evidence. In addition, in subsequent sections we develop momentum strategies formed using our market cycles that generate positive unconditional alpha, which such models cannot account for.

Behavioral models of slow information updating, however, have the potential to explain the evidence. In particular, Gómez-Cram (2022) develops a model in which investors use the present-value identity to value stocks but discount future cash flows using subjective expectations that are slow to reflect new information, such as worsening economic conditions.

[21] See Gómez-Cram (2022) and the references therein.

In this behavioral framework, the subjective cash-flow component of expected returns can offset the rational risk-premia component to induce negative return predictability when the economy is transitioning to recession. Moreover, later in recessions as economic conditions are more fully reflected in beliefs, the impact of the subjective component reduces and expected returns revert to the countercyclical behavior inherent in rational frameworks.

Furthermore, our results suggest a role for the sense of cyclicality itself to vary over time in models of time-varying risk premia. Bull, Correction, Bear, and Rebound states can each occur in any period of the business cycle. As shown in Figure 5, there are periods within both the high-risk state of recessions and the low-risk state of expansions in which Bear states are most likely or Bull and Rebound states are most likely. For example, we see a ramp up in the relative frequency of Bear states late in expansions, which is not inconsistent with countercyclical risk premia. In addition, the spread between the predicted expected returns of Bull and Bear states can be economically meaningful in any period of the business cycle.[22] We conjecture that a model of subjective beliefs with reaction speeds to macro news that vary as a function of the market cycle could help reconcile variation in cyclicality with the prevailing view of wholly countercyclical risk premia.

4. Static Speed Selection

In this section, we analyze momentum strategies of various intermediate speeds defined as in (11), (12), and (13). Our empirical analyses focus on U.S. aggregate stock market returns over a recent 50-year period.[23] For most of the paper, we evaluate performance over a recent 50-year period for consistency with later analyses in which we use earlier data to warm up our dynamic strategies for out-of-sample evaluation. In Internet Appendix F, we report results for U.S. equities based on a longer evaluation period beginning in 1927, from which we draw similar inferences as in our main analyses.

Table 2 summarizes estimates of basic unconditional moments for TS momentum strategies of various speeds alongside the long-only market strategy, whose strategy weights equal to one every month. Results are gross of transaction costs.[24] We highlight these results in

[22]Over the period 1927-07 to 2018-12, the annualized Bull–Bear spread is 16.0% unconditionally, 23.7% in recessions, and 7.5% in expansions.

[23]Recent studies focus on applications of momentum to *aggregate* factors. Ehsani and Linnainmaa (2019) show evidence of TS momentum across equity factors that appears to subsume the CS momentum risk factor. A companion paper by Arnott et al. (2019) shows that CS factor momentum is a pervasive phenomenon that, moreover, drives industry momentum. Gupta and Kelly (2019) show that portfolios of TS factor momentum can add value to a wide array of investment strategies.

[24]We estimate transaction costs to be modest under conservative assumptions. Let the bid–ask spread on S&P 500 Index futures be 2 basis points (bps), roll costs 0.5 bps (with rolls happening every quarter), and commissions 0.1 bps; then the ratio of alpha to transaction costs does not surpass 3% for the strategies

three parts: return and risk, market timing, and tail behavior.

Table 2: **Performance Summary by Speed**

	Market	$a=0$ SLOW	$a=\frac{1}{4}$	$a=\frac{1}{2}$ MED	$a=\frac{3}{4}$	$a=1$ FAST
Return and Risk						
Average (%) (anlzd.)	5.91	6.46	6.17	5.88	5.59	5.30
Volatility (%) (anlzd.)	15.64	15.62	12.72	11.60	12.74	15.66
Sharpe Ratio (anlzd.)	0.38	0.41	0.48	0.51	0.44	0.34
Market Timing						
Average Position	1.00	0.46	0.39	0.32	0.25	0.18
Market Beta	1.00	0.15	0.05	−0.04	−0.13	−0.23
Alpha (%) (anlzd.)	0.00	5.58	5.85	6.12	6.39	6.66
Alpha t-statistic	—	2.54	3.24	3.71	3.57	3.07
Tail Behavior						
Skewness	−0.55	−0.43	−0.13	0.02	0.03	0.15
Max. Drawdown (%)	−54.36	−43.43	−37.96	−34.43	−34.07	−44.53
Average (anlzd.)/|Max. DD|	0.11	0.14	0.16	0.17	0.17	0.12

Notes: This table reports the sample average, volatility, Sharpe ratio, average position, market beta and alpha, alpha t-statistic, skewness, maximum drawdown, and ratio of average return to absolute maximum drawdown for monthly returns of the long-only market strategy (Market) and of TS momentum strategies of various speeds. The slow strategy weight applied to the market return in month $t+1$, $w_{\mathrm{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\mathrm{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\mathrm{SLOW},t} + aw_{\mathrm{FAST},t}$, for $a \in [0,1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return (Mkt-RF) from the Kenneth French Data Library. The evaluation period is 1969-01 to 2018-12.

Return and risk. First, momentum strategies of every speed exhibit positive average historical returns in our backtest. Although they all report average returns similar in magnitude to the long-only market strategy, intermediate-speed strategies achieve their returns with lower volatilities, and therefore report the highest average return per unit of risk as measured by the Sharpe ratio (see Section 4.1).

Market timing. Second, momentum strategies of every speed have positive average positions in the underlying market (ranging from 0.18 to 0.46 long, on average), yet their market betas are relatively small in magnitude with negative point estimates for intermediate-to-fast speeds. Intermediate-speed strategies are approximately market-beta neutral and deliver alphas having the largest t-statistics (see Section 4.2).

reported in Table 2.

Tail behavior. Third, momentum strategies of every speed exhibit more desirable skewness than the market strategy. Higher positive (or less negative) skewness is a desirable property for most investors, and skewness of intermediate-to-fast strategies performs best in this regard. Moreover, intermediate-speed strategies have less severe maximum drawdowns and record the highest average returns per unit of drawdown risk (see Section 4.3).

These three categories of results emphasize several merits of *intermediate*-speed momentum strategies. In particular, MED delivers the highest Sharpe ratio (0.51) and it does so with approximate market beta neutrality (beta of -0.04), delivering the most statistically significant alpha (t-statistic of 3.71), without exhibiting the negative skewness behavior of the underlying market nor the severe drawdown behavior of the market and the slow and fast strategies. In the following three subsections, we identify properties of the stock market's return distribution and develop related theory to help explain these findings.

4.1. Return and Risk

4.1.1. Conditional Return Contributions

Table 3 (Panel A) decomposes the unconditional average returns in the first row of Table 2 into their conditional contributions following each market cycle. Specifically, we employ the following decomposition of returns using zero-one indicators of the market cycle s_t at date t:

$$r_{t+1} = r_{t+1}1_{\{s_t=\mathrm{Bu}\}} + r_{t+1}1_{\{s_t=\mathrm{Be}\}} + r_{t+1}1_{\{s_t=\mathrm{Co}\}} + r_{t+1}1_{\{s_t=\mathrm{Re}\}}. \tag{15}$$

For example, the conditional contribution to speed-a TS momentum strategy after Bull cycles is the sample estimate of $\mathbf{E}[r_{t+1}(a)1_{\{s_t=\mathrm{Bu}\}}]$. For the market excess return, which has unconditional expectation 5.91%, its conditional contribution after Bull cycles is 4.59%, which is the product of its conditional expected value after Bull cycles, 9.5%, and the relative frequency of Bull cycles, 0.483 (cf. Figure 1). The first row of Panel A repeats the unconditional values reported in Table 2 for ease of reference. The middle four values in each column (and the last two values in each column) sum to the value in the first row (subject to rounding).

First, consider the decomposition of average returns to the slow strategy as compared with the long-only market strategy. After Bulls and Corrections, SLOW is long one unit, such as the market strategy, and so gets the same contribution as the market after these phases, namely, 4.59% and 1.59%, respectively. After Bears and Rebounds, SLOW shorts the market and flips the sign of the market contribution from -1.29% to 1.29% and from 1.01% to -1.01% after Bears and Rebounds, respectively.

Table 3: **Market-Cycle Decomposition of Returns by Speed**

Panel A: Average Returns

Average (%) (anlzd.)	Market	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Unconditional	5.91	6.46	6.17	5.88	5.59	5.30
Cycle-Conditional Decomposition						
Bull	4.59	4.59	4.59	4.59	4.59	4.59
Correction	1.59	1.59	0.79	0.00	−0.79	−1.59
Bear	−1.29	1.29	1.29	1.29	1.29	1.29
Rebound	1.01	−1.01	−0.51	0.00	0.51	1.01
Bull + Bear	3.31	5.88	5.88	5.88	5.88	5.88
Correction + Rebound	2.60	0.58	0.29	0.00	−0.29	−0.58

Panel B: Variance of Returns

Variance (%) (anlzd.)	Market	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Unconditional	2.44	2.44	1.61	1.34	1.62	2.45
Cycle-Conditional Decomposition						
Bull	0.63	0.63	0.63	0.63	0.63	0.63
Correction	0.77	0.77	0.19	0.00	0.20	0.79
Bear	0.73	0.71	0.71	0.71	0.71	0.71
Rebound	0.31	0.32	0.08	0.00	0.08	0.31
Bull + Bear	1.36	1.34	1.34	1.34	1.34	1.35
Correction + Rebound	1.09	1.10	0.27	0.00	0.28	1.10

Notes: This table reports the unconditional sample average (Panel A) and sample variance (Panel B) of monthly returns of the long-only market strategy and of TS momentum strategies of various speeds and their contribution in months immediately following various market states. Contributions in each column sum to their corresponding unconditional values in the first row of each panel (subject to rounding) either across the four individual states in the middle rows or across the two state pairs in the last two rows. The long-only market strategy (Market) and TS momentum strategies of various speeds are the same as those described in Table 2. Market states, s_t, are defined as follows. Bull: $w_{\text{SLOW},t} = w_{\text{FAST},t} = +1$; Correction: $w_{\text{SLOW},t} = +1, w_{\text{FAST},t} = -1$; Bear: $w_{\text{SLOW},t} = w_{\text{FAST},t} = -1$; and Rebound: $w_{\text{SLOW},t} = -1, w_{\text{FAST},t} = +1$. The evaluation period is 1969-01 to 2018-12.

Second, consider the average return contributions across different speeds. TS momentum strategies of all speeds get the same contribution after Bull and Bear cycles because each is long after Bulls and short after Bears. These states account for the bulk of the average returns to all strategies. Contributions after Corrections and Rebounds explain the differences in average returns by speed. SLOW's loss after Rebounds is similar in magnitude to FAST's loss after Corrections. The net contribution to SLOW after Corrections and Rebounds is 0.58%. In contrast, FAST flips the sign relative to SLOW after Correction and Rebound cycles, yielding a net contribution of −0.58%. MED is simply out of the market after Corrections and Rebounds. These facts explain why the unconditional average returns are close to each other and to that of the long-only market strategy.

Table 3 (Panel B) decomposes the unconditional variance of returns (squares of volatilities reported in second row of Table 2) into their cycle-conditional contributions following each observed market cycle based on the same cycle decomposition of returns as in (15). The cycle-conditional variance contributions are computed as the covariances of each of the terms on the right-hand side of (15) with the overall return. For example, the conditional variance contribution to speed-a TS momentum strategy after Bull cycles is the sample estimate of $\mathbf{Cov}[r_{t+1}(a)1_{\{\text{Bu}\}}, r_{t+1}(a)]$. The middle four values in each column (and the last two values in each column) sum to the value in the first row (subject to rounding).

The majority of return variance for all speeds of TS momentum strategies and for the long-only market strategy are generated after Bull and Bear markets. However, intermediate-speed strategies scale down their market exposure relative to SLOW and FAST after Corrections and Rebounds and, therefore, experience lower exposure to variance risk emanating from these states. In particular, after these periods of disagreement, MED exits the market altogether and avoids any exposure to the variance risk of these states. These facts explain why the unconditional return variances (and volatilities) are lower for intermediate-speed strategies compared to SLOW and FAST and compared to the long-only market strategy.

4.1.2. Mapping SLOW/FAST Disagreement to Sharpe Ratios

When one blends two strategies that are relatively uncorrelated, such as SLOW and FAST, it is not that surprising to see the variance decrease and the Sharpe ratio increase. In this subsection, however, we provide an alternative characterization of the Sharpe ratio in terms of a disagreement multiplier, which isolates the volatility contribution from states of disagreement between SLOW and FAST after Corrections and Rebounds. This multiplier is also useful for understanding the skewness of a blend, a topic we discuss in Section 4.3.

The role of such disagreement in determining Sharpe ratios manifests analytically as a *disagreement multiplier*, $D(a)$, which appears in (19) of Proposition 4. $D(a)$ captures the ratio of the volatility of the unconditional market return to the volatility of the momentum strategy, in terms of volatility contributions from Bull or Bear states and from Correction or Rebound states, respectively.[25]

[25]As before, let r_{t+1} denote the return on an underlying security and recall that the speed-a strategy return, $r_{t+1}(a)$, is just the weighted average of SLOW and FAST with weight a applied to FAST. Also, let

Proposition 4 (Sharpe ratio decomposition)**.** *The Sharpe ratio of $r_{t+1}(a)$ can be expressed in terms of the Sharpe ratios of $r_{SLOW,t+1}$ and $r_{FAST,t+1}$, respectively, and the market cycles, as follows:*

$$\begin{aligned}\mathbf{Sharpe}[r_{t+1}(a)] = (1-a)\,\mathbf{Sharpe}[r_{SLOW,t+1}]D(a,\mathbf{E}[r_{SLOW,t+1}])\\ + a\,\mathbf{Sharpe}[r_{FAST,t+1}]D(a,\mathbf{E}[r_{FAST,t+1}]),\end{aligned} \tag{16}$$

where

$$D(a,\mu) := \sqrt{\frac{\mathbf{E}[r_{t+1}^2]-\mu^2}{\mathbf{E}[r_{t+1}^2|{}^{Bu}_{Be}]\mathbf{P}[{}^{Bu}_{Be}] + (2a-1)^2\mathbf{E}[r_{t+1}^2|{}^{Co}_{Re}]\mathbf{P}[{}^{Co}_{Re}] - (\mathbf{E}[r_{t+1}(a)])^2}}. \tag{17}$$

Approximating squared average strategy returns by $(\mathbf{E}[r_{t+1}(a)])^2 \approx 0$ for $a \in [0,1]$, we have

$$\mathbf{Sharpe}[r_{t+1}(a)] \approx ((1-a)\,\mathbf{Sharpe}[r_{SLOW,t+1}] + a\,\mathbf{Sharpe}[r_{FAST,t+1}])\,D(a), \tag{18}$$

where the term $D(a)$, multiplying the weighted average of Sharpe ratios in (18), is

$$D(a) := D(a,0) = \sqrt{\frac{\mathbf{E}[r_{t+1}^2]}{\mathbf{E}[r_{t+1}^2|{}^{Bu}_{Be}]\mathbf{P}[{}^{Bu}_{Be}] + (2a-1)^2\mathbf{E}[r_{t+1}^2|{}^{Co}_{Re}]\mathbf{P}[{}^{Co}_{Re}]}}. \tag{19}$$

$D(a)$ is greater than or equal to one and is maximized at $a = \frac{1}{2}$ on $a \in [0,1]$:

$$D\left(\tfrac{1}{2}\right) \geq D(a) \geq 1, \quad a \in [0,1]. \tag{20}$$

Moreover, $a = \frac{1}{2}$ is the unique maximizer with $D(\frac{1}{2}) > 1$ if the relative frequency of such states is not zero so that $\mathbf{E}[r_{t+1}^2|{}^{Co}_{Re}]\mathbf{P}[{}^{Co}_{Re}] > 0$.

Proof. See Internet Appendix A. □

Proposition 4 indicates that the risk-adjusted performance of intermediate-speed momentum strategies ($0 < a < 1$) is greater than the a-weighted average risk-adjusted performances of the slow and fast momentum strategies, taken separately, as long as squared expected returns are relatively small and can be approximated by zero.[26] In particular, (18) states that the Sharpe ratio of a strategy with speed a is approximated by the product of two components: (A) the a-weighted blend of the separate Sharpe ratios of SLOW and FAST,

$\{{}^{\mathrm{Bu}}_{\mathrm{Be}}\}$ denote the union of events Bull or Bear, and $\{{}^{\mathrm{Co}}_{\mathrm{Re}}\}$ the union of events Correction or Rebound. We apply similar notation throughout the paper for the union of any of the four market states.

[26]For example, if the average return is 5%, then its square is only 0.25%. Note that the approximation is exact at the endpoints of the interval $[0,1]$.

and (B) the disagreement multiplier, $D(a)$. In general, the speed a^* of the strategy that maximizes the Sharpe ratio depends on both of these components.

Proposition 4 states that component (B) is maximized at $a = \frac{1}{2}$. Intermediate-speed strategies reduce volatility originating from Correction and Rebound states. Such volatility exposure is largely captured by the conditional average squared returns following these states, $\mathbf{E}[r_{t+1}^2|_{\mathrm{Re}}^{\mathrm{Co}}]$, times the relative frequency of these states, $\mathbf{P}[_{\mathrm{Re}}^{\mathrm{Co}}]$. Hence, $\mathbf{E}[r_{t+1}^2|_{\mathrm{Re}}^{\mathrm{Co}}]\mathbf{P}[_{\mathrm{Re}}^{\mathrm{Co}}]$ represents the frequency-weighted contribution of these states to the unconditional volatility of the strategy. Intermediate-speed strategies with parameter a in the vicinity of $\frac{1}{2}$ tend to scale down positions following such periods, boosting the overall risk-adjusted performance by reducing exposure to risk. The factor $(2a-1)^2$ multiplying $\mathbf{E}[r_{t+1}^2|_{\mathrm{Re}}^{\mathrm{Co}}]\mathbf{P}[_{\mathrm{Re}}^{\mathrm{Co}}]$ captures this scaling-down, and this reduction in the denominator boosts the disagreement multiplier $D(a)$ above one. At $a = \frac{1}{2}$, volatility contributions from Correction and Rebound states are completely eliminated ($(2(\frac{1}{2}) - 1)^2 = 0$).

If the Sharpe ratios of SLOW and FAST are sufficiently different from each other, then component (A) can play a dominant role. For example, suppose the Sharpe ratio of SLOW ($a = 0$) is much larger than that of FAST. Although component (B) is maximized at $a = \frac{1}{2}$, it might not be large enough to overcome component (A)'s reduced contribution at $a = \frac{1}{2}$ to the product in (18). In this example, a^* could be 0.

However, as a corollary, if the Sharpe ratios of SLOW and FAST are sufficiently close to each other, then component (A) is relatively insensitive to a, and so a^* depends primarily on the disagreement multiplier. In this case, $a^* \approx \frac{1}{2}$. This corollary helps explain MED's relatively high Sharpe ratio in Table 2 because average returns are roughly the same magnitude across all speeds. Moreover, we can apply this corollary to the situation in which one does not have conviction about which of SLOW and FAST will have the higher Sharpe ratio ex ante. Compared to a randomized experiment of choosing to invest in either SLOW or FAST, this corollary informs us that MED has a higher ex-ante Sharpe ratio than the expected Sharpe ratio of the experiment.

Furthermore, because Proposition 4 is a model-free result, it extends beyond our running empirical example of the U.S. stock market and beyond our model of Section 2.2 to momentum strategies of various speeds applied in any market. We examine the performance of international equity markets in Section 6 and of equity factors in Internet Appendix G. Because equity factors take offsetting long and short positions in equity subportfolios, the economic linkages to equity returns discussed in Section 3 may not apply to factor returns. However, the general statistical properties outlined in Proposition 4 still apply to equity factors: under mild assumptions, the Sharpe ratio of MED is higher than the average of the Sharpe ratios of SLOW and FAST. In addition, for each of the equity factors in Internet

Appendix G, there is an intermediate-speed strategy that has both Sharpe ratio and alpha t-statistic higher than those of both SLOW and FAST.

4.1.3. Mapping SLOW/FAST Sharpe Ratios to Mean Persistence and Realized Noise Levels in Market Returns

We now use the empirical results of the momentum strategies reported in Table 2 together with the model of Section 2.2 to compute implications for the persistence behavior of expected returns and the noise-level behavior of realized returns. We also discuss the optimal static blend under the model.

Using (6) and (7) of Proposition 1, we map our empirical evidence into Figure 2(a). Specifically, we have two equations of the model Sharpe ratios of FAST and SLOW (respectively, (6) and (7)) in terms of two unknowns: persistence in expected returns, φ, and the noise ratio θ. We jointly estimate $(\widehat{\varphi}, \widehat{\theta})$ as the root of these two Sharpe ratio equations each set equal to their corresponding empirical Sharpe ratio values:

$$0.34 = \frac{\sqrt{\frac{24}{\pi}}\left(1-\widehat{\theta}\right)\widehat{\varphi}}{\sqrt{1-\frac{2}{\pi}\left(1-\widehat{\theta}\right)^2(\widehat{\varphi})^2}}, \tag{21}$$

$$0.41 = \frac{\sqrt{\frac{24}{\pi}}\left(1-\widehat{\theta}\right)\widehat{\varphi}}{\sqrt{\frac{\left(1-\widehat{\theta}\right)[12\left(1-(\widehat{\varphi})^2\right)-2\widehat{\varphi}\left(1-(\widehat{\varphi})^{12}\right)]+12\widehat{\theta}(1-(\widehat{\varphi}))^2}{(1-(\widehat{\varphi})^{12})^2}-\frac{2}{\pi}\left(1-\widehat{\theta}\right)^2(\widehat{\varphi})^2}}, \tag{22}$$

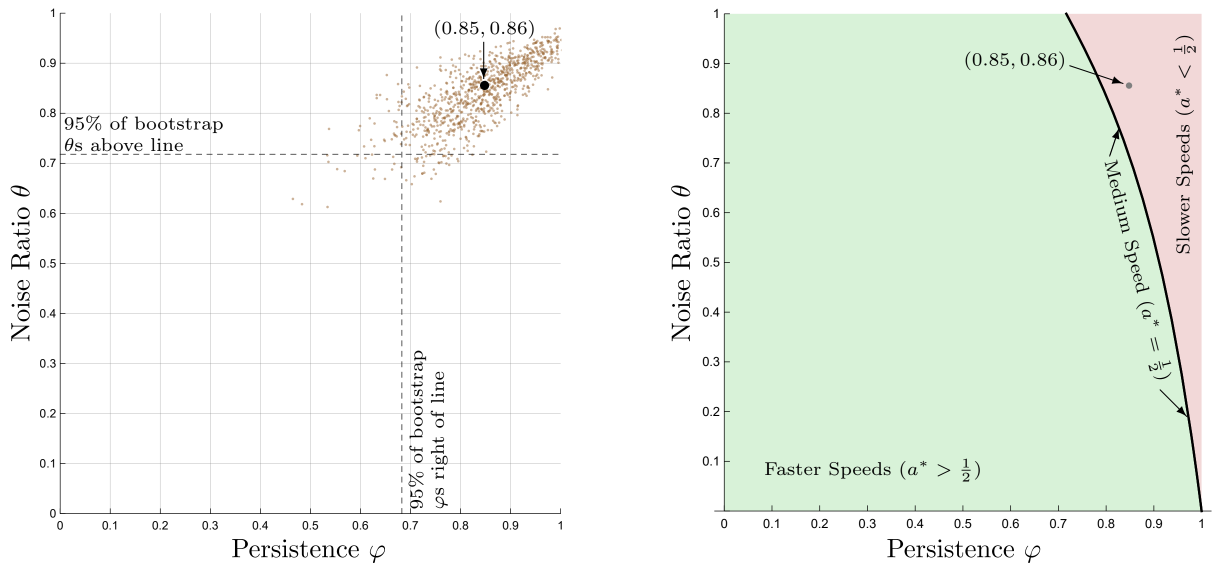
where 0.34 and 0.41 are the empirical Sharpe ratios of FAST and SLOW, respectively, reported in Table 2. We obtain $(\widehat{\varphi}, \widehat{\theta}) = (0.85, 0.86)$, rounded to the nearest percent. As a robustness check to potential estimation error, we perform 1,000 block-bootstrap estimates of $(\widehat{\varphi}, \widehat{\theta})$ based on FAST/SLOW Sharpe ratio pairs computed from 10 non-overlapping 5-year blocks from our 50-year history, sampled with replacement, plugged into (21) and (22), respectively.

Figure 6(a) plots these parameter pair estimates on the unit square. Figure 6(a) also plots the 5th percentiles (dashed lines) for the bootstrapped estimates of each of $\widehat{\varphi}$ and $\widehat{\theta}$. These results suggest that expected returns have relatively high persistence (0.85 point estimate, above 0.68 in 95% of bootstrap estimates) and that realized returns have relatively high noise (0.86 point estimate, above 0.72 in 95% of bootstrap estimates).

The unconditional Sharpe ratio of a static blend (12) of FAST and SLOW under the model of Section 2.2 cannot be analyzed in general in closed form, so we do not provide a general solution for the speed parameter $a^* \in [0,1]$ that maximizes the unconditional Sharpe

Figure 6: **Implied Mean Persistence and Realized Noise Level**

(a) Estimated φ and θ (b) Model Optimal Speed Zones



Notes: Figure (a) plots the model-implied mean persistence and realized noise ratio pair $(\widehat{\varphi}, \widehat{\theta}) = (0.85, 0.86)$ inferred from the empirical Sharpe ratios of FAST and SLOW and model equations (6) and (7). The small brown dots represent model-implied estimates of $(\widehat{\varphi}, \widehat{\theta})$ corresponding to 1,000 block-bootstrap estimates of Sharpe ratios of FAST and SLOW. The vertical dashed line ($\varphi = 0.68$) represents the 5th percentile of bootstrap $\widehat{\varphi}$ estimates. The horizontal dashed line ($\theta = 0.72$) represents the 5th percentile of bootstrap $\widehat{\theta}$ estimates. Figure (b) plots the model-implied Sharpe-ratio-maximizing speeds a^* for various persistence and noise ratios. The solid curve represents combinations of persistence and noise ratios such that the Sharpe-ratio-maximizing speed is the equal blend $a^* = \frac{1}{2}$. Combinations in the zone below this curve (green zone) are associated with faster optimal speeds, $a^* > \frac{1}{2}$. Combinations above (red zone) are associated with slower optimal speeds, $a^* < \frac{1}{2}$.

ratio.[27] However, when the Sharpe ratios of FAST and SLOW are equal, Proposition 5 indicates that the equal-blend MED generates the maximum Sharpe ratio among all static blends under the model.

Proposition 5 (Maximization when Sharpe ratios are equal)**.** *Under the model of Section 2.2, if SLOW and FAST have the same Sharpe ratio, then $a^* = \frac{1}{2}$ gives the optimal Sharpe ratio among all static blends, $a \in [0, 1]$, of SLOW and FAST.*

Proof of Proposition 5. See Internet Appendix A. □

By Proposition 5, for all parameter pairs (φ, θ) on the solid curve in Figure 6(b), the optimal blend that maximizes the unconditional Sharpe ratio is $a^* = \frac{1}{2}$. Our estimate of

[27]The expected return of a static blend is available in closed form. To compute the volatility, however, we need the expected value of the square of a normal random variable times the signs of two jointly normal correlated random variables: $\mathbf{E}[r_t^2 \cdot \text{sign}(r_{t-1}) \cdot \text{sign}\left(\frac{1}{12}(r_{t-1} + \cdots + r_{t-12})\right)]$. This expectation does not admit a closed-form expression.

$(\widehat{\varphi}, \widehat{\theta}) = (0.85, 0.86)$ lies close to this curve, consistent with $a = \frac{1}{2}$ generating a higher Sharpe ratio than other static speeds reported in Table 2 ($a = 0, \frac{1}{4}, \frac{3}{4}, 1$). Because FAST has a higher Sharpe ratio than SLOW for combinations of persistence and noise below the curve in Figure 6(b), we can infer from (18) that such combinations are associated with faster optimal speeds ($a^* > \frac{1}{2}$). Similarly, combinations of persistence and noise above this curve are associated with slower optimal speeds ($a^* < \frac{1}{2}$). Note that Proposition 5 highlights a nontrivial property of our static-speed momentum strategies; it does not generally hold for any two portfolios that if those portfolios have the same Sharpe ratios then an equal blend of those portfolios has the maximum Sharpe ratio of all blends.

4.2. Market Timing

4.2.1. The Determinants of Market Beta and Alpha

Given the preponderance of Bull cycles—around half of all months—as reported in Figure 4, it should not come as a surprise that trend following of the U.S. stock market implies a positive static tilt. Indeed, as detailed in Table 2, the average position of the momentum strategies ranged from 46% to 18% from slow to fast speeds. As also detailed in Table 2, however, the betas of the momentum strategy returns are relatively low in magnitude and range from 0.15 to -0.23 from slow to fast speeds, with negative point estimates for intermediate-to-fast speeds—something we might not expect given their positive tilts.

To understand this evidence, we first disentangle the timing bets and static bets in expected returns with a widely-used decomposition:

$$\begin{aligned} r_{t+1}(a) &= w_t(a) r_{t+1} \\ &= \underbrace{(w_t(a) - \mathbf{E}[w_t(a)])}_{\text{dynamic}} r_{t+1} + \underbrace{\mathbf{E}[w_t(a)]}_{\text{static}} r_{t+1}, \end{aligned} \tag{23}$$

where the first equality above matches the first equality in (12). Taking expectations, we have

$$\mathbf{E}[r_{t+1}(a)] = \underbrace{\mathbf{Cov}[w_t(a), r_{t+1}]}_{\text{market timing}} + \underbrace{\mathbf{E}[w_t(a)]}_{\substack{\text{static}\\ \text{dollar}\\ \text{exposure}}} \mathbf{E}[r_{t+1}], \tag{24}$$

because $(w_t(a) - \mathbf{E}[w_t(a)])$ is mean zero. The covariance term is also known as the market-timing component, and it represents the share of expected returns generated by the dynamic bets of the signal as reflected by the strategy weights $w_t(a)$. In contrast, the average strategy weight summarizes the static dollar exposure of the strategy.

Huang et al. (2019) argue that static average tilts are the primary source of TS momentum performance. To support their argument, they construct two portfolios: (1) a TS momentum portfolio that goes long or short based on the sign of trailing 12-month returns, similar to our SLOW, and (2) a portfolio that goes long or short based on the sign of the entire history of returns. This second strategy is essentially a buy-and-hold market strategy when inception-to-date market returns are positive. They show that these two strategies perform similarly with respect to average and risk-adjusted returns, thereby calling into question the role of timing in the 12-month signal.

However, if their argument were correct, we would have a beta puzzle. To see the problem, consider the first two columns of Table 2. The "Market" column reflects a buy-and-hold portfolio and the "SLOW" column reflects a 12-month TS momentum portfolio. Indeed, both portfolios exhibit similar performance in average returns (approximately 6%) and Sharpe ratios (approximately 0.40). Yet, if SLOW is inheriting its performance from its static tilt, then why is its market beta so low—only 0.15? Small betas imply meaningful alphas, which map to the alphas estimated in Table 2. Moreover, small betas imply low correlations, so we could blend SLOW with the market buy-and-hold strategy to achieve a better Sharpe ratio—a contradiction to their argument.

What explains this apparent beta puzzle? The market-timing component must be adding negative beta with respect to the underlying market, which offsets the beta of the static allocation. As reported in Panel A of Table H.1 in Internet Appendix H, the timing share of the portfolios displays insignificant or only marginally significant excess returns, a result that echoes the results of Huang et al. (2019). Yet, as reported in Panel B of the same table, these portfolios recorded negative betas with respect to the underlying market. To better understand this mechanism, we next disentangle the timing and static components in the market covariance, beta, and alpha:

Proposition 6 (Covariance with underlying market)**.** *The (contemporaneous) covariance between the speed strategy returns and the long-only market strategy returns can be decomposed* *as follows:*[28]

$$\begin{aligned}&\mathbf{Cov}[r_{t+1}(a), r_{t+1}] \\ &\quad = \mathbf{E}[w_t(a)]\mathbf{Var}[r_{t+1}] + \mathbf{Cov}[w_t(a), r_{t+1}]\mathbf{E}[r_{t+1}] + \mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1})^2],\end{aligned} \tag{25}$$

for $a \in [0, 1]$*. The market beta and alpha, respectively, can be decomposed as follows:*

$$\begin{aligned}&\mathbf{Beta}[r_{t+1}(a)] \\ &\quad = \underbrace{\mathbf{E}[w_t(a)]}_{\substack{\textit{static}\\ \textit{component}}} + \underbrace{\frac{\mathbf{Cov}[w_t(a), r_{t+1}]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}]}_{\substack{\textit{market timing}\\ \textit{component}}} + \underbrace{\frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}}_{\substack{\textit{volatility timing}\\ \textit{component}}},\end{aligned} \tag{26}$$

$$\begin{aligned}&\mathbf{Alpha}[r_{t+1}(a)] \\ &\quad = \mathbf{Cov}[w_t(a), r_{t+1}]\left(1 - \frac{(\mathbf{E}[r_{t+1}])^2}{\mathbf{Var}[r_{t+1}]}\right) - \frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}],\end{aligned} \tag{27}$$

for $a \in [0, 1]$*. Approximating squared average market returns by* $(\mathbf{E}[r_{t+1}])^2 \approx 0$*, we have*

$$\mathbf{Alpha}[r_{t+1}(a)] \approx \mathbf{Cov}[w_t(a), r_{t+1}] - \frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}]. \tag{28}$$

Proof. See Internet Appendix A. □

Equation (26) of Proposition 6 indicates that the market beta of any of our momentum strategies can be decomposed into a static component and a market-timing component, similar to the decomposition of the expected return in (24); however, an important additional volatility-timing component arises, which reflects the predictability of strategy weights for subsequent return volatility.[29] Therefore, if the momentum weights significantly covary with

[28] Our decompositions have similarities in form to the relationships established between unconditional beta (alpha) and expected conditional beta (alpha) in the conditional CAPM literature. Jagannathan and Wang (1996) were the first to establish the exact relation between unconditional beta and expected conditional beta (equations A10–A14 in the appendix of their paper). This relationship is also developed and examined by Lewellen and Nagel (2006) and Boguth et al. (2011). Conditional beta and expected conditional beta in these studies can be mapped to our strategy weight and static beta component, respectively. However, the other components in these studies include covariances between conditional beta and each of *conditional* market returns, *conditional* squared market returns, and *conditional* market variance, *given a conditioning information set*. Boguth et al. (2011) organize these terms into two groups and use the labels "market timing" and "volatility timing" (equations (5) and (6)). Because our expressions use market excess returns directly rather than conditional expected returns given a conditioning information set, there is no direct counterpart to the market and volatility timing components in our decompositions.

[29] Previous studies focus almost exclusively on market timing. An exception is Boguth et al. (2011), who argue that volatility timing can be an important component, if not the dominant driver of unconditional alpha. Our results are consistent and we show that volatility timing comprises approximately one- to two-

the subsequent squared return deviations from their mean, then the beta of the momentum portfolio is not well approximated by the beta of the average momentum position. Thus, even if the market-timing component is relatively small compared to a larger positive static component, the volatility-timing component could be relatively large, but of opposite sign, and enough to offset the static component of the market beta. As indicated in Table 4, this is indeed the case for the U.S. stock market. Such TS momentum predictability is not taken into account by Huang et al. (2019).

Table 4: **Beta and Alpha Decompositions by Speed**

Market Beta and Alpha	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Beta	0.15	0.05	-0.04	-0.13	-0.23
Alpha (%) (anlzd.)	5.58	5.85	6.12	6.39	6.66
Alpha t-stat.	2.54	3.24	3.71	3.57	3.07
Beta Components					
Static	0.457	0.387	0.317	0.247	0.177
Market Timing	0.008	0.008	0.008	0.008	0.009
Volatility Timing	-0.315	-0.339	-0.364	-0.389	-0.414
Alpha Components (%) (anlzd.)					
Market Timing	3.72	3.84	3.96	4.09	4.21
Volatility Timing	1.86	2.00	2.15	2.30	2.44

Notes: This table reports the sample market beta, alpha, and alpha t-statistic of monthly returns of momentum strategies of various speeds repeated from Table 2. The table also reports the (additive) decomposition of beta into static, market-timing, and volatility-timing components according to estimates of the terms in (26) and the (additive) decomposition of alpha into market timing and volatility timing according to estimates of the terms in (27). TS momentum strategies of various speeds are the same as those described in Table 2. The evaluation period is 1969-01 to 2018-12.

Table 4 repeats the market beta, alpha, and alpha t-statistic estimates reported in Table 2 for ease of reference and also shows the market beta breakdown into the three (additive) components corresponding to the terms in (26). The static beta component equals the average position as reported in Table 2. The market-timing component of beta is small and near zero for all speeds. The volatility-timing component of beta is roughly as large in magnitude as the static component, but of opposite sign. Together, these terms sum to the market betas in the top row.

Table 4 also reports the market alpha breakdown into its two (additive) components corresponding to the terms in (27). The market-timing component in the alpha approxima-

thirds of the alpha of our TS momentum strategies.

tion of (28) is identical to the market-timing component in the widely-used expected return decomposition of (24). The alpha decomposition reveals, however, that volatility timing, in addition to market timing, can be a driver of alpha. Indeed, Table 4 reports that the volatility-timing component composes a large portion—over 33%—of each strategy's overall alpha estimated over the evaluation period. Volatility timing accounts for over 50% of alpha over the last 15 years (see Internet Appendix I).

Why do we find near-zero market timing in beta yet a large relative market-timing component in alpha? The answer is fairly straightforward. When beta is zero, alpha is just the expected strategy excess return, which is composed of market timing and static components only, as in (24). However, the static part of beta (which can be non-zero even when beta is zero) perfectly offsets any static component in alpha coming from expected strategy returns. Therefore, there is no static component in the decomposition of alpha in (27). If there were no volatility timing, then we would expect 100% of alpha to come from market timing—via expected strategy returns or market timing in beta. The reason that alpha market timing is only about 65% of alpha rather than 100% is that volatility timing explains the remaining portion.

To summarize, the average position of a TS momentum portfolio does not need to reflect the beta exposure of the strategy. If weights predict volatility, then the market timing and static dollar exposure terms offer an incomplete picture of the alpha generated by TS momentum.

4.2.2. Relation to Volatility-Managed Portfolios

A potential overlap between TS momentum strategies and the volatility-managed (VOM) portfolios of Moreira and Muir (2017) should not be surprising. VOM portfolios increase exposure to an underlying portfolio following low-volatility states and decrease exposure following high-volatility states. Moreira and Muir document that managing the leverage of a strategy in this manner can increase Sharpe ratios and deliver alpha with respect to the underlying portfolio. The contemporaneous correlation between stock market returns and their monthly volatility has been about -28% over our 50-year evaluation period.[30] This negative correlation implies that trailing volatility tends to be lower (higher) when trailing returns are positive (negative). Therefore, we expect some positive comovement between the weights of a TS momentum strategy and a VOM strategy applied to a common underlying portfolio.

[30] This negative association is often referred to as the leverage effect in reference to an economic explanation of the phenomenon offered by Black (1976) and Christie (1982). According to this explanation, negative equity returns lead to higher firm leverage, which in turn should translate into higher equity volatility as firms become riskier. See also Harvey et al. (2018).

Is TS momentum alpha primarily a repackaging of volatility management alpha? Pertaining to the U.S. stock market, Moreira and Muir highlight two key ingredients in VOM portfolios: (1) trailing volatility tends to be uncorrelated to subsequent returns; and (2), volatility tends to be persistent at short horizons. These ingredients entail little-to-no market timing and mostly volatility timing in VOM portfolio alpha, by construction. Moreira and Muir's theoretical results predict positive VOM alpha when there is no correlation between volatility and expected returns, highlighting that the positive empirical VOM alpha they report can be explained solely by volatility timing. In contrast, TS momentum alpha can have significant market timing in addition to volatility timing, as illustrated in Table 4. Therefore, TS momentum strategies appear to extract a distinct source of alpha relative to VOM strategies. To get a broader answer, one can apply our alpha decomposition to any collection of TS strategies on a common underlying portfolio—including more complex versions of momentum and VOM strategies—to measure the degree of overlap in their market timing and volatility timing. See also Internet Appendix C, in which we compare MED with the TS momentum strategy of Moskowitz, Ooi, and Pedersen (2012), which explicitly uses volatility targeting.

4.3. Tail behavior

4.3.1. Market-cycle decomposition

Table 5 reports various percentiles of monthly market returns in months following each of the four market cycles. Corrections introduce extreme outcomes and volatility despite most outcomes being positive (median monthly return of 1.07%).[31] Yet, extreme outcomes tend to be more severe on the downside than the upside. The fast strategy tends to flip deeper Correction losses into gains by going short after Corrections, which can help explain its slightly positive point estimate for skewness in Table 2. However, FAST also has full exposure to volatility from both Correction and Rebound states, in which the spread of returns is larger on both the positive and negative sides relative to Bull states. Intermediate-speed strategies reduce exposure to both volatility and extreme events associated with these states. In particular, MED avoided this exposure altogether with a zero position in months after these states, which can help explain its approximately zero point estimate for skewness in Table 2.

Downside risk exposure after each market cycle can help explain the pattern of maximum

[31] Five of the 10 worst months in our 50-year evaluation period were after Correction phases: −23.24% (1987-10); −16.08% (1998-08); −12.90% (1980-03); −11.91% (1978-10); and −10.72% (2000-11). Four of the 10 best months in the evaluation period were after Correction phases: 12.47% (1987-01); 12.16% (1976-01); 11.35% (2011-10); and 10.84% (1991-12).

Table 5: **Cycle-Conditional Market Return Distributions**

Return Percentiles (%)	Bull	Correction	Bear	Rebound
MIN	−9.55	−23.24	−17.23	−10.35
P01	−7.85	−14.62	−12.79	−10.16
P05	−4.64	−7.14	−10.10	−8.41
P10	−3.37	−5.64	−8.06	−5.51
P25	−1.51	−2.08	−4.83	−2.44
P50	1.05	1.07	−0.89	1.15
P75	3.07	3.82	3.98	4.59
P90	4.68	5.84	6.82	7.24
P95	6.13	7.15	7.99	7.98
P99	7.21	11.79	13.68	10.61
MAX	9.59	12.47	16.10	11.30

Notes: This table reports various percentiles of monthly market excess returns in months immediately following each of four market states—Bull, Correction, Bear, and Rebound—defined in terms of slow and fast momentum strategy positions as in Table 3. PX is the X-th percentile. For example, P10 is the 10th percentile of monthly returns. The evaluation period is 1969-01 to 2018-12.

drawdowns across strategies reported in Table 2. After Bear states, the magnitudes of lower percentile (negative) returns are higher in most cases than the magnitudes of symmetrically higher percentile (positive) returns. Momentum strategies of all speeds are short after Bear markets, turning this downside into upside. Consistent with this evidence, Table 2 reports lower maximum drawdowns than the long-only market strategy for all speeds. Intermediate speeds further reduce exposure to extreme downside events by scaling down after Corrections and Rebounds. Accordingly, intermediate speeds report lower maximum drawdowns and higher average returns per unit of absolute maximum drawdown in Table 2.

4.3.2. Skewness

We formalize the relationship between skewness of MED relative to the skewness of SLOW and FAST in terms of MED's Sharpe ratio and the disagreement multiplier, $D(a)$, introduced in Section 4.1.2. The connection between the Sharpe ratio and the skewness of a random variable is illuminated in Lemma 7, which we apply to obtain Proposition 8.

Lemma 7. *For any Y such that its first three moments are defined and $\mathbf{SD}[Y] > 0$, then* $\mathbf{Skew}[Y] = \frac{\mathbf{E}[Y^3]}{(\mathbf{SD}[Y])^3} - \mathbf{Sharpe}[Y]\,(3 + (\mathbf{Sharpe}[Y])^2)$, *where* $\mathbf{Sharpe}[Y] = \frac{\mathbf{E}[Y]}{\mathbf{SD}[Y]}$.

Proof. See Internet Appendix A. □

Proposition 8 (Skewness decomposition)**.** *The skewness of $r_{t+1}(a)$ can be expressed in terms of the skewness of $r_{SLOW,t+1}$ and $r_{FAST,t+1}$, respectively, and a disagreement multiplier. An* *exact expression as well as an approximation for all $a \in [0,1]$ based on $(\mathbf{E}[r_{t+1}(a)])^2 \approx 0$ for $a \in [0,1]$ are given in* (A49) *and* (A50)*, respectively, in Internet Appendix* A*. In the special case of $a = \frac{1}{2}$, we have*

$$\begin{aligned}\mathbf{Skew}[r_{t+1}(\tfrac{1}{2})] &\approx \frac{1}{2}(\mathbf{Skew}[r_{SLOW,t+1}] + \mathbf{Skew}[r_{FAST,t+1}])\left(D(\tfrac{1}{2})\right)^3 \\ &+ 3\,\mathbf{Sharpe}[r_{t+1}(\tfrac{1}{2})]\left(\left(D(\tfrac{1}{2})\right)^2\left[1 + \left(\frac{\mathbf{Sharpe}[r_{FAST,t+1}] - \mathbf{Sharpe}[r_{SLOW,t+1}]}{2}\right)^2\right] - 1\right),\end{aligned} \tag{29}$$

where $D(\frac{1}{2})$ is as defined in (19) *for $a = \frac{1}{2}$.*

Proof. See Internet Appendix A. □

Proposition 8 indicates that skewness of MED is scaled up relative to the average skewness of SLOW and FAST—since $\left(D(\frac{1}{2})\right)^3 > 1$ per Proposition 4—and shifted to the right when MED has a positive Sharpe ratio—since $\left(D(\frac{1}{2})\right)^2 > 1$ and because the term in the outer parentheses is always positive. The multiplier $D(\frac{1}{2})$ is the same disagreement multiplier which appears in Proposition 4, $D(a)$, evaluated at $a = \frac{1}{2}$. As reported in Table 2, the slow strategy has negative skewness of -0.43, but the fast strategy has positive skewness of 0.15. Their average skewness is negative at -0.14. The disagreement multiplier $D(\frac{1}{2}) = 1.34$ amplifies the first term in (29) by a factor of $\left(D(\frac{1}{2})\right)^3 = 2.42$, bringing its contribution to -0.34. The second term in (29) shifts this value to the right by 0.36, yielding the slight positive skewness of 0.02 for MED as reported in Table 2.

As before, because Proposition 8 is a model-free result, it extends beyond our running empirical example and beyond our the model of Section 2.2 to momentum strategies of various speeds applied in any market. Moreover, as a corollary to this result, if SLOW and FAST both have nonnegative skewness and Sharpe ratios when applied in some market, then the skewness of MED is going to be positive and higher than the maximum skewness of both SLOW and FAST.

5. Dynamic Speed Selection

The differences in the conditional return distributions following Correction and Rebound cycles (as reported in Table 5 and discussed in Section 4.3.1) raise the question of whether a dynamic momentum strategy could offer improved performance over a static strategy. For example, how might a strategy perform if its weights following Corrections and Rebounds were individually specified instead of implied by the weighted average of slow and fast strategies? Would strategy weights of 0 and $\frac{1}{2}$ following Corrections and Rebounds perform better than the zero weights implied by the MED strategy, or the $-\frac{1}{2}$ and $\frac{1}{2}$ weights of the $a = \frac{3}{4}$ strategy? We address questions of this nature in this section.

Instead of using a static speed parameter a, we let it vary dynamically with the four market cycles. The dynamic speed $a_{s(t)}$ at date t is a function of the observable market state $s(t)$ at that date, which is one of $\{\text{Bu}, \text{Co}, \text{Be}, \text{Re}\}$. For example, if in a Correction cycle at date t ($s(t) = \text{Co}$), then a_{Co} is the parameter which will govern the blending between slow and fast strategy weights in the subsequent month. If the cycle remains in Correction at date $t+1$, then we apply the same a_{Co} for the next month. If the cycle shifts to Bear at date $t+2$, then we apply a_{Be} for the next month, and so on.

The corresponding dynamic strategy return is

$$r_{t+1}(a_{s(t)}) = w_t(a_{s(t)})\, r_{t+1} = [(1 - a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t}]r_{t+1}. \tag{30}$$

Note that since $w_{\text{SLOW},t} = w_{\text{FAST},t}$ with magnitude one after Bull or Bear cycles, the dynamic weight in (30) is invariant to the values of a_{Bu} and a_{Be}. That is, $w_t(a_{\text{Bu}}) = 1$ after Bull for all a_{Bu} and $w_t(a_{\text{Be}}) = -1$ after Bear for all a_{Be}. Nevertheless, the dynamic weight is sensitive to the values of a_{Co} and a_{Re} following Correction and Rebound cycles, respectively. In Proposition 9, we establish the values of these state-conditional speed parameters that maximize the steady-state Sharpe ratio of the dynamic returns.

Proposition 9 (Sharpe ratio maximizing dynamic speed rule)**.** *Consider the problem of choosing the state-conditional speeds $a_{s(t)}$ which will be applied following every occurrence of state $s(t)$ in order to achieve the highest steady-state Sharpe ratio:*

$$\max_{a_{s(t)}:s(t)\in\{Co,Re\}} \mathbf{Sharpe}[r_{t+1}(a_{s(t)})]. \tag{31}$$

If $\mathbf{E}[r_{t+1}|Bu]\mathbf{P}[Bu] > \mathbf{E}[r_{t+1}|Be]\mathbf{P}[Be]$, then

$$a_{Co} = \frac{1}{2}\left(1 - \frac{\mathbf{E}[r_{t+1}^2|{}^{Bu}_{Be}]\mathbf{P}\left[{}^{Bu}_{Be}\right]}{\mathbf{E}[r_{t+1}|Bu]\mathbf{P}[Bu] - \mathbf{E}[r_{t+1}|Be]\mathbf{P}[Be]}\frac{\mathbf{E}[r_{t+1}|Co]}{\mathbf{E}[r_{t+1}^2|Co]}\right), \tag{32}$$

$$a_{Re} = \frac{1}{2}\left(1 + \frac{\mathbf{E}[r_{t+1}^2|{}^{Bu}_{Be}]\mathbf{P}\left[{}^{Bu}_{Be}\right]}{\mathbf{E}[r_{t+1}|Bu]\mathbf{P}[Bu] - \mathbf{E}[r_{t+1}|Be]\mathbf{P}[Be]}\frac{\mathbf{E}[r_{t+1}|Re]}{\mathbf{E}[r_{t+1}^2|Re]}\right), \tag{33}$$

is the unique state-conditional speed pair that maximizes (31).

Proof. See Internet Appendix A. $\square$

Proposition 9 specifies the dynamic speed selections that maximize the steady-state Sharpe ratio in terms of cycle-conditional first- and second-population moments of market returns. The condition $\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] > \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]$ ensures that the weights are maximizers and not minimizers. This condition is typically satisfied because expected returns are typically positive following Bull cycles and negative following Bear cycles.

Because population values of the first and second moments in (32) and (33) are not observable, we use historical estimates of these moments to approximate their values. We use the label "DYN" to denote the investable strategy that employs state-dependent speeds based on estimated versions of (32) and (33), using only data prior to strategy implementation (i.e., no look-ahead bias). Specifically, DYN weights are defined as follows: $w_{\text{DYN},t} := w_t(\widehat{a}_{s(t)}) = (1 - \widehat{a}_{s(t)})w_{\text{SLOW},t} + \widehat{a}_{s(t)}w_{\text{FAST},t}$, where $s(t) \in \{\text{Bu}, \text{Co}, \text{Be}, \text{Re}\}$; $\widehat{a}_{\text{Bu}} = \widehat{a}_{\text{Be}} = \frac{1}{2}$; and $\widehat{a}_{\text{Co}}$ and $\widehat{a}_{\text{Re}}$ are estimated using sample averages $\widehat{\mathbf{E}}[\cdot]$ and $\widehat{\mathbf{P}}[\cdot]$, which denote estimates of the average and frequency of their given argument, respectively, based on data prior to strategy implementation.[32] For example, $\widehat{\mathbf{E}}[r^2_{t+1}|{}^{\text{Bu}}_{\text{Be}}]$ is the historical sample average of squared returns in months after Bull or Bear states within the historical estimation window, and $\widehat{\mathbf{P}}\left[{}^{\text{Bu}}_{\text{Be}}\right]$ is the frequency of Bull or Bear states within that historical sample. If these cycle-conditional return moments are relatively stable over time, then DYN may offer improved performance out of sample.

5.1. Performance of Dynamic Strategy

Table 6 reports the out-of-sample performance of DYN over various evaluation windows in our 50-year evaluation period—from 50 years ago forward in 5 year increments to the most recent 15 years. State-dependent speeds $\widehat{a}_{\text{Co}}$ and $\widehat{a}_{\text{Re}}$ of DYN are fixed over each evaluation window using data from the corresponding fixed training window from 1926-07 to the month before the evaluation start month.[33] We define the efficiency of the DYN strategy over each evaluation window as the ratio of its Sharpe ratio to that of the highest performing state-conditional speed strategy ex post, which we label the "OPT" strategy. OPT uses fixed speeds a^*_{Co} and a^*_{Re} over each evaluation window that give the highest in-sample Sharpe ratio rather than speeds estimated from trailing data. OPT's Sharpe ratios vary over different evaluation windows from 0.570 to 0.721, indicating that *different periods expose the strategies to different performance opportunities.* We can measure the performance of any state-conditional-speed strategy relative to the ceiling set by OPT, including DYN and

[32] If either of the estimates $\widehat{a}_{\text{Co}}$ or $\widehat{a}_{\text{Re}}$ fall outside the unit interval $[0, 1]$, then we set its value to the nearest endpoint, 0 or 1.

[33] Results are similar if we update speed estimates $\widehat{a}_{\text{Co}}$ and $\widehat{a}_{\text{Re}}$ at end of each month $t-1$ during the evaluation period using data from expanding rather than fixed training windows from 1926-07 to month $t-1$ to form strategy weights for each month t.

Table 6: **DYN Strategy Performance Over a 50-Year Period**

DYN Strategy Training Window			Evaluation					
			Evaluation Window			Sharpe Ratio (anlzd.)		
From (yr-mo)	To (yr-mo)	Length (yrs)	From (yr-mo)	To (yr-mo)	Length (yrs)	DYN $(\widehat{a}_{\text{Co}}, \widehat{a}_{\text{Re}})$	OPT	Efficiency DYN/OPT
1926-07	1968-12	42.5	1969-01	2018-12	50.0	0.524 $(0.00, 0.58)$	0.570	0.920
1926-07	1973-12	47.5	1974-01	2018-12	45.0	0.547 $(0.07, 0.59)$	0.572	0.956
1926-07	1978-12	52.5	1979-01	2018-12	40.0	0.611 $(0.08, 0.65)$	0.626	0.977
1926-07	1983-12	57.5	1984-01	2018-12	35.0	0.614 $(0.22, 0.67)$	0.623	0.985
1926-07	1988-12	62.5	1989-01	2018-12	30.0	0.688 $(0.26, 0.69)$	0.721	0.954
1926-07	1993-12	67.5	1994-01	2018-12	25.0	0.675 $(0.11, 0.71)$	0.684	0.988
1926-07	1998-12	72.5	1999-01	2018-12	20.0	0.564 $(0.17, 0.69)$	0.579	0.975
1926-07	2003-12	77.5	2004-01	2018-12	15.0	0.611 $(0.16, 0.69)$	0.621	0.984

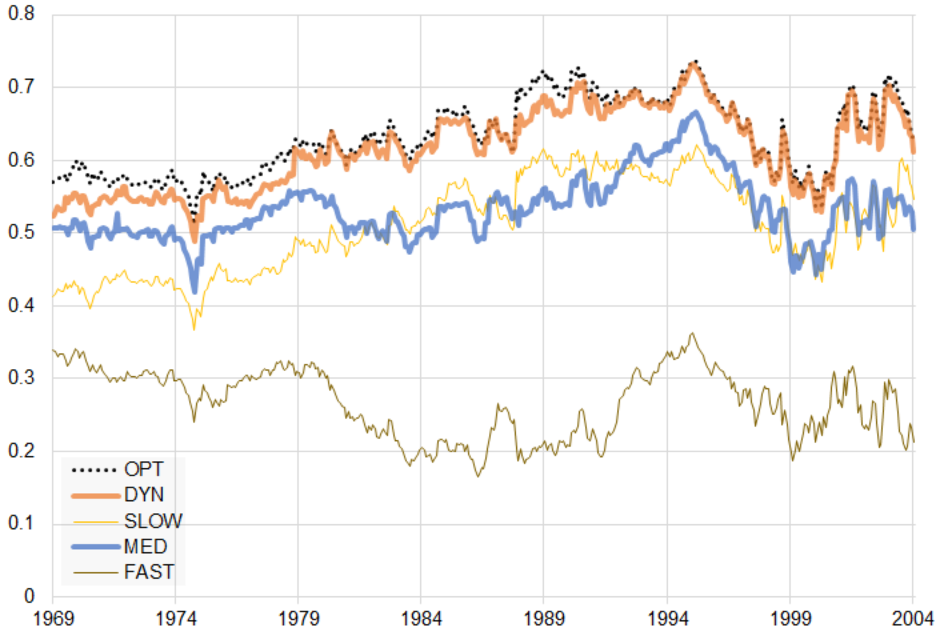
Notes: This table reports the DYN momentum strategy's Sharpe ratio and its efficiency for various evaluation windows within the 50-year evaluation period. DYN is the dynamic (state-dependent) speeds strategy whose weights, $w_t(a_{s(t)}) = (1 - a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t}$, are based on point estimates of the mean-variance optimal state-dependent speeds, $a_{s(t)}$, using $\widehat{a}_{\text{Bu}} = \widehat{a}_{\text{Be}} = \frac{1}{2}$ and sample versions of (32) and (33) for $\widehat{a}_{\text{Co}}$ and $\widehat{a}_{\text{Re}}$, respectively, based on data in the training window; $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative and otherwise equals -1; $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative and otherwise equals -1; and the four observable market states at date t are defined as: Bu: $w_{\text{SLOW},t} = w_{\text{FAST},t} = +1$; Be: $w_{\text{SLOW},t} = w_{\text{FAST},t} = -1$; Co: $w_{\text{SLOW},t} = +1, w_{\text{FAST},t} = -1$; and Re: $w_{\text{SLOW},t} = -1, w_{\text{FAST},t} = +1$. If either of the estimates $\widehat{a}_{\text{Co}}$ or $\widehat{a}_{\text{Re}}$ fall outside the unit interval $[0, 1]$, then we set its value to the nearest endpoint, 0 or 1. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted market factor return (Mkt-RF) from the Kenneth French Data Library. OPT is the dynamic strategy that would have achieved the maximum Sharpe ratio, ex post. State-dependent speeds of both strategies are fixed over the evaluation window. For example, using 57.5 years of data from 1926-07 through 1983-12 to estimate speeds for DYN, its Sharpe ratio over the subsequent 35-year evaluation period from 1981-07 through 2018-12 was 0.614. Relative to the best possible ex-post state-dependent strategy, OPT, which had a Sharpe ratio of 0.623, DYN was 98.5% = 0.614/0.623 efficient.

the special cases of static speeds. DYN consistently exhibits efficiency of 92% or higher across different evaluation windows, reflecting its capacity to exploit performance opportunities out of sample.

The out-of-sample performance of DYN also exhibits modest gains compared to static-speed strategies. Figure 7 compares the Sharpe ratios of DYN and OPT with those of various static-speed strategies over evaluation windows beginning from each date on the horizontal axis through the end of the sample period. The time series values of DYN and OPT at 5-year-interval timestamps correspond to the values reported in Table 6. Note that the static-speed strategy with highest Sharpe ratio is not always MED.[34] Regardless of which

[34]Figure J.1 in Internet Appendix J is the same as Figure 7 but also includes Sharpe ratios for static speeds $a = \frac{1}{4}$ and $a = \frac{3}{4}$, highlighting that other static speeds can generate higher Sharpe ratios than MED.

Figure 7: **Sharpe Ratios (Month t to End of Sample): DYN vs. Static Strategies**



Beginning Month t to 2018-12

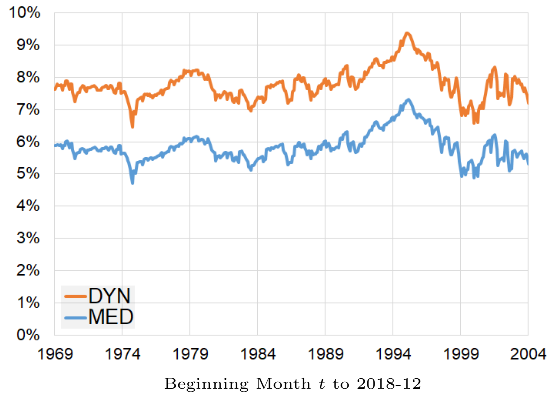
Notes: This figure plots, for each month t in 1969-01 to 2004-01, the Sharpe ratio of various strategies over the evaluation period beginning month t and ending December 2018. Strategies OPT and DYN are the same as those described in Table 6. Static-speed strategies are the same as those described in Table 2.

static-speed strategy performs best, DYN edges out the best static-speed strategy, including MED, uniformly across all evaluation periods. Although its Sharpe ratio outperformance on an absolute scale is marginal in some periods (e.g., DYN's 0.52 to MED's 0.51 over the 50-year evaluation window beginning 1969), DYN consistently attains nearly all of the performance opportunities available (captured by OPT). For example, the gap between DYN and OPT is consistently slim—efficiency never below 92%. In contrast, the gap between MED and OPT is sometimes relatively large—MED as low as 75% efficiency. Therefore, from an ex-ante point of view not knowing the best static speed, DYN offers a systematic approach to adjust the speed.

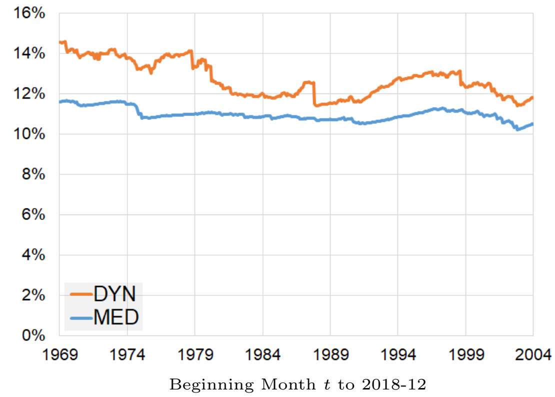
Furthermore, DYN generates alpha with respect to MED. Figure 8 reports several time series of performance metrics for DYN and MED over evaluation windows beginning from each date on the horizontal axis through the end of the sample period. DYN uniformly generates higher excess returns than MED (Figure 8(a)), but also uniformly generates higher volatility (Figure 8(b)). Recall that MED completely exits the market after Corrections and Rebounds, which eliminates its exposure to volatility as well as to return opportunities after these states. In contrast, DYN takes (scaled-down) positions after these states, allowing

Figure 8: **Performance of DYN vs. MED (Month t to End of Sample)**

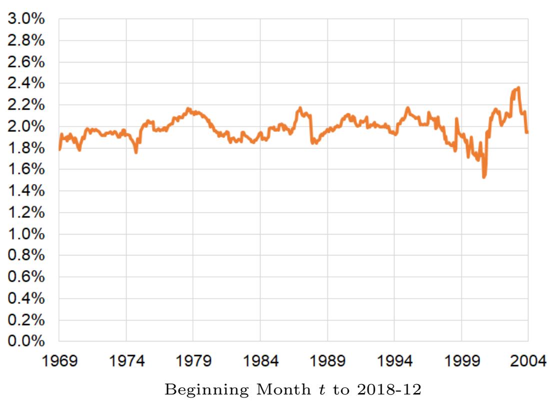
(a) Average Returns (anlzd.)



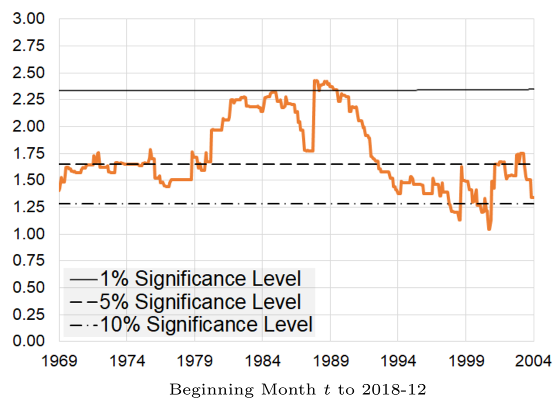
(b) Return Volatilities (anlzd.)



(c) Alphas of DYN (anlzd., benchmarked against MED)



(d) Alpha t-Statistics of DYN (benchmarked against MED)



Notes: This figure plots, for each month t in 1969-01 to 2004-01, (a) the average returns and (b) volatilities of DYN and MED, (c) the alphas of DYN with respect to MED, and (d) the alpha t-statistics of DYN with respect to MED with one-sided significance reference levels over the evaluation period beginning month t and ending December 2018. Strategy MED is the same as in Table 2 and strategy DYN is the same as in Table 6.

DYN to take advantage of potentially favorable return/volatility trade-offs reflected in past market-cycle return patterns. Figure 8(c) plots the time series of alphas of DYN with respect to MED, showing an average of about 2.0% per annum across evaluation windows. Figure 8(d) plots the t-statistics corresponding to these alphas and shows an average one-sided significance level of about 5%. Despite the fact that their Sharpe ratios are relatively close over the 50-year evaluation period beginning 1969, DYN exhibits statistically significant alpha relative to MED at the one-sided 10% significance level.

5.2. DYN Sensitivity to Estimation Error

We now analyze the sensitivity of the state-conditional speed estimates of DYN to estimation error via block-bootstrap resampling of historical returns. We find that state-dependent speed pairs that yield the highest performance for DYN tend to be those in the region having relatively slow speed after Corrections and relatively fast speed after Rebounds. Moreover, Sharpe ratio performance is relatively insensitive to variation of speeds within this region.

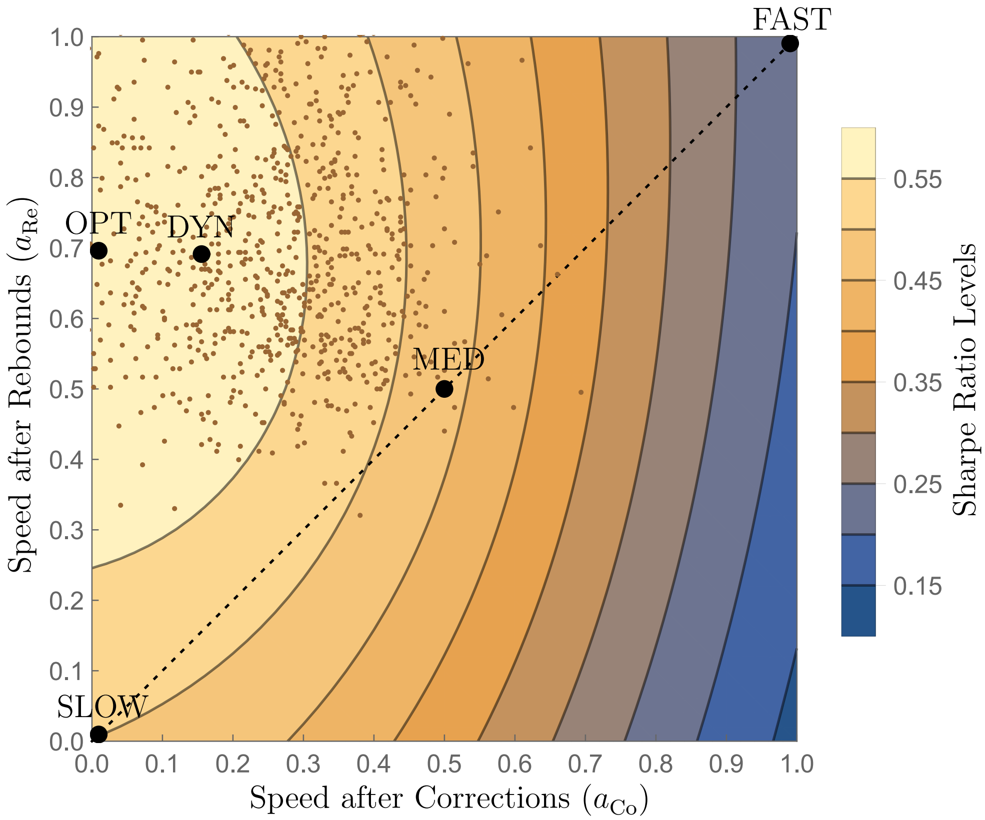
Figure 9 graphs Sharpe ratio contour levels for momentum strategies with state-dependent speed pairs $(a_{\mathrm{Co}}, a_{\mathrm{Re}})$ on the unit square over the most recent 15-year evaluation period 2004-01 to 2018-12. Each speed pair is set at the beginning of the evaluation period. The concentric curves represent speed pairs with equal Sharpe ratio levels. Lighter-shaded regions reflect higher Sharpe ratios and darker-shaded regions reflect lower Sharpe ratios. Static strategies—strategies whose speeds do not vary by state—are represented by the points on the dotted diagonal line. OPT is the state-dependent speed pair that would have achieved the highest Sharpe ratio ex post, and DYN indicates the state-dependent pair based on estimates of state-conditional first and second moments using monthly data from 1926-07 to 2003-12 prior to the evaluation period.

The figure illustrates that the Sharpe ratio is less sensitive to deviations in speed pair values near the in-sample mean-variance optimal speed pair, OPT, than it is farther away. The DYN speed pair based on the historical sample is close to OPT and is approximately as good a performer. The small dots represent the state-dependent speed pairs from 1,000 block-bootstrap estimates of state-conditional first and second moments using monthly data from 1926-07 to 2003-12.[35] The bootstrap DYN speed pairs tend to fall into the upper-left quadrant of the graph in Figure 9, which corresponds to strategies that tilt to SLOW after Corrections and to FAST after Rebounds.

To summarize, state-dependent speed pairs that yield the highest performance tend to be those that are relatively slow after Corrections and relatively fast after Rebounds. Moreover, speed pairs in this region are closer to optimal, and Sharpe ratio performance in this region is less sensitive to estimation error.

[35]The block bootstrap is performed as follows. We partition the 930 historical monthly returns on the U.S. stock market from 1926-07 to 2003-12 into 10 equal-sized nonoverlapping blocks of 93 consecutive months. We randomly sample 10 of these blocks with replacement to form one bootstrap historical sample. For each bootstrap sample, we determine cycle states and estimate state-conditional first and second moments. We plug these estimates into (32) and (33) to generate a speed pair for the sample.

Figure 9: **Sharpe Ratio Performance of Different State-Conditional Speed Pairs Over the Most Recent 15 Years**



Notes: This figure reports Sharpe ratio levels of dynamic-speed momentum strategies over the evaluation period 2004-01 to 2018-12 for speed pairs $(a_{\mathrm{Co}}, a_{\mathrm{Re}})$ where strategy weights take the form: $(1-a_{s(t)})w_{\mathrm{SLOW,t}}+a_{s(t)}w_{\mathrm{FAST,t}}$, where $w_{\mathrm{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative and otherwise equals -1; $w_{\mathrm{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative and otherwise equals -1; and the four observable market states at date t are defined as: Bu: $w_{\mathrm{SLOW},t} = w_{\mathrm{FAST},t} = +1$; Be: $w_{\mathrm{SLOW},t} = w_{\mathrm{FAST},t} = -1$; Co: $w_{\mathrm{SLOW},t} = +1, w_{\mathrm{FAST},t} = -1$; and Re: $w_{\mathrm{SLOW},t} = -1, w_{\mathrm{FAST},t} = +1$. Speeds $a_{\mathrm{Bu}} = a_{\mathrm{Be}} = 0.5$ for all strategies in the figure. The dotted diagonal highlights the continuum of static strategies, which apply the same speeds after all states—SLOW ($a_{s(t)} = 0$), MED ($a_{s(t)} = 0.5$), and FAST ($a_{s(t)} = 1$) are labeled. OPT represents the dynamic strategy pair that would have achieved the best Sharpe ratio ex post. DYN is the dynamic (state-dependent) speeds strategy whose weights are based on point estimates of the mean-variance optimal state-dependent speeds, $a_{s(t)}$, using $\widehat{a}_{\mathrm{Bu}} = \widehat{a}_{\mathrm{Be}} = \frac{1}{2}$ and sample versions of (32) and (33) for $\widehat{a}_{\mathrm{Co}}$ and $\widehat{a}_{\mathrm{Re}}$, respectively, based on data before the evaluation window beginning 1926-07. If either $\widehat{a}_{\mathrm{Co}}$ or $\widehat{a}_{\mathrm{Re}}$ fall outside the unit interval $[0, 1]$, then we set its value to the nearest endpoint, 0 or 1. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted market factor return (Mkt-RF) from the Kenneth French Data Library. The small dots represent block bootstrap point estimates for DYN using monthly data from 1926-07 to 2003-12. Note that the intermediate-speed static strategy with $a = \frac{1}{4}$ achieves approximately the highest Sharpe ratio of all static strategies over the measurement period.

5.3. Market-Cycle Patterns

The conclusion from our state-dependent speed analysis—elect slower-speed momentum after Correction months and faster-speed momentum after Rebound months—is reinforced by market cycle patterns, as those reported in Table 7.

Table 7: **Market Cycle Transitions**

	Monthly Transition Probability (%)					
	Bull_{t+1}	Correction_{t+1}	Bear_{t+1}	Rebound_{t+1}	Up_{t+1}	Down_{t+1}
Bull_t	62.8	34.8	2.1	0.3	63.3	36.7
Correction_t	61.2	29.9	8.8	0.0	61.2	38.8
Bear_t	9.0	0.0	55.0	36.0	45.0	55.0
Rebound_t	14.3	1.6	42.9	41.3	55.6	44.4

Notes: This table reports the relative frequency of transitions from one market state (row) in month t to another market state (column) in month $t+1$ (first four columns) and the relative frequency of positive (Up) and negative (Down) returns next month (last two columns) over the 50-year period from 1969-01 to 2018-12. Market state classifications are the same as those described in Table 3.

Table 7 reports for the 50-year evaluation period the relative frequency of each market state in the subsequent month (first four columns) given the market state in the preceding month (rows). It also reports the relative frequency of positive and negative returns in the subsequent month (Up and Down states, respectively, in the last two columns) given the market state in the preceding month (rows). First, Corrections tend to revert to Bulls. Given the current month is a Correction state, then most of the time (61.2%) the next month is a Bull month and an Up month. If we view a Correction state as an alarm that uptrend could be turning to downtrend, then this alarm tends to be a false alarm. The fast strategy takes a short position after Corrections, which is a bad bet. Therefore, adopting a slower-speed momentum position after Corrections is a better bet. Second, the market is more likely to go up after Rebounds. Given the current month is a Rebound state, then most of the time (55.6%) the next month is an Up month. If we view a Rebound state as an alarm that downtrend could be turning to uptrend, then regardless of the accuracy of this alarm, FAST takes a long position, which is a good bet. Therefore, adopting a faster-speed momentum position after Rebounds is more effective than adopting slower speeds after Rebounds.

6. Evidence from Other Equity Markets

The conclusion to elect intermediate-speed momentum strategies from our static-speed analysis using U.S. equity data in Section 4 carries over to international equity markets.

Table 8: **Sharpe Ratios of Static-Speed Strategies in International Markets**

	Sharpe Ratio (anlzd.)			
Country	$a = 0.00$ SLOW	$a = 1.00$ FAST	Average of SLOW, FAST	$a = 0.50$ MED
AUS	0.121	0.076	0.098	0.123
AUT	0.435	0.600	0.518	0.627
BEL	0.476	0.460	0.468	0.619
CAN	0.234	0.498	0.366	0.486
CHE	0.619	0.468	0.543	0.687
DEU	0.498	0.211	0.354	0.436
DNK	0.454	0.410	0.432	0.569
ESP	0.281	0.169	0.225	0.280
FIN	0.720	0.649	0.685	0.869
FRA	0.382	0.355	0.369	0.466
GBR	0.227	−0.111	0.058	0.072
HKG	0.274	0.451	0.363	0.443
IRL	0.339	0.309	0.324	0.413
ITA	0.539	0.016	0.277	0.351
JPN	0.411	0.127	0.269	0.333
NLD	0.488	0.090	0.289	0.391
NOR	0.376	0.621	0.498	0.669
SGP	0.184	0.502	0.343	0.462
SWE	0.440	0.303	0.371	0.475
USA	0.511	0.155	0.333	0.462

Notes: This table reports the Sharpe ratios (anlzd.) for various strategies applied to different country equity markets evaluated over the period from 1981-01 to 2018-12 for most countries—first year of data (1980) is used to define the SLOW signal prior to the evaluation period. Static-speed strategy weights are formed according to $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$ for speed parameter $a \in [0, 1]$, where $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative and otherwise equals -1; and $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative and otherwise equals -1. MED is the strategy for which $a = \frac{1}{2}$. Monthly strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the excess equity return for the country. Excess returns are obtained via Datastream. See Internet Appendix B for data details. Green-highlighted values are larger than the average of SLOW and FAST values.

Table 8 reports the Sharpe ratios for static-speed momentum strategies applied in 20 country equity markets and evaluated since 1981.[36] For all countries in the table, as predicted by Proposition 4, the Sharpe ratio for MED is higher than the average of the Sharpe ratios of SLOW and FAST.

Table 9 reports the Sharpe ratios for static-speed and dynamic-speed momentum strategies applied in these 20 equity markets and evaluated over the last 15 years. Here, we use

[36]Our data sample begins in 1980 for most countries. See Internet Appendix B for data details. Data in the first year (1980) is used to define the SLOW (12-month) signal for the first strategy return in 1981-01.

Table 9: **Sharpe Ratios of the Dynamic-Speed Strategy in International Markets**

	Sharpe Ratio (anlzd.)					
Country	$a = 0.00$ SLOW	0.25	0.50 MED	0.75	1.00 FAST	DYN $(\widehat{a}_{\text{Co}}, \widehat{a}_{\text{Re}})$
AUS	0.363	0.412	0.415	0.346	0.256	0.516 $(0.00, 1.00)$
AUT	0.422	0.560	0.670	0.694	0.647	0.729 $(0.60, 0.76)$
BEL	0.678	0.718	0.679	0.547	0.392	0.815 $(0.28, 1.00)$
CAN	0.405	0.545	0.652	0.644	0.564	0.650 $(0.61, 1.00)$
CHE	0.644	0.774	0.832	0.750	0.605	0.956 $(0.00, 0.94)$
DEU	0.459	0.479	0.437	0.326	0.208	0.535 $(0.00, 0.61)$
DNK	0.599	0.662	0.650	0.531	0.385	0.557 $(0.11, 1.00)$
ESP	0.147	0.140	0.112	0.066	0.024	0.202 $(0.00, 0.80)$
FIN	0.604	0.703	0.699	0.513	0.320	0.787 $(0.23, 0.75)$
FRA	0.424	0.495	0.505	0.405	0.284	0.622 $(0.39, 0.87)$
GBR	0.303	0.219	0.067	−0.097	−0.205	0.477 $(0.00, 1.00)$
HKG	0.297	0.453	0.597	0.632	0.585	0.545 $(0.25, 1.00)$
IRL	0.624	0.694	0.697	0.604	0.474	0.721 $(0.49, 0.68)$
ITA	0.363	0.339	0.272	0.170	0.073	0.363 $(0.00, 0.00)$
JPN	0.519	0.566	0.559	0.478	0.370	0.519 $(0.00, 0.00)$
NLD	0.397	0.356	0.222	0.036	−0.095	0.461 $(0.00, 0.80)$
NOR	0.394	0.527	0.624	0.599	0.508	0.580 $(0.59, 1.00)$
SGP	0.288	0.392	0.473	0.470	0.415	0.531 $(0.61, 1.00)$
SWE	0.587	0.685	0.688	0.519	0.336	0.782 $(0.14, 0.58)$
USA	0.595	0.624	0.570	0.421	0.265	0.670 $(0.00, 0.83)$

Notes: This table reports the Sharpe ratios (anlzd.) for various strategies applied to different country equity markets evaluated over the 15-year period from 2004-01 to 2018-12. Static-speed strategy weights are formed according to $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$ for speed parameter $a \in [0,1]$, where $w_{\text{SLOW},t}$, equals +1 if the trailing 12-month return (arithmetic average monthly return) is nonnegative and otherwise equals −1; and $w_{\text{FAST},t}$, equals +1 if the trailing 1-month return is nonnegative and otherwise equals −1. DYN strategy weights take the form $(1 - a_{s(t)})w_{\text{SLOW,t}} + a_{s(t)}w_{\text{FAST,t}}$, where speed $a_{s(t)}$ depends on the four observable market states at date t: Bu: $w_{\text{SLOW},t} = w_{\text{FAST},t} = +1$; Be: $w_{\text{SLOW},t} = w_{\text{FAST},t} = -1$; Co: $w_{\text{SLOW},t} = +1, w_{\text{FAST},t} = -1$; and Re: $w_{\text{SLOW},t} = -1, w_{\text{FAST},t} = +1$. DYN speeds are based on point estimates of the mean-variance optimal state-dependent speeds, $a_{s(t)}$, using $\widehat{a}_{\text{Bu}} = \widehat{a}_{\text{Be}} = \frac{1}{2}$ and sample versions of (32) and (33) for $\widehat{a}_{\text{Co}}$ and $\widehat{a}_{\text{Re}}$, respectively, based on data before the evaluation window beginning 1980-01. If either $\widehat{a}_{\text{Co}}$ or $\widehat{a}_{\text{Re}}$ fall outside the unit interval $[0,1]$, then we set its value to the nearest endpoint, 0 or 1. Monthly strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the excess equity market factor return for the country. Excess returns are obtained via Datastream. See Internet Appendix B for data details. Green-highlighted values are at least as large as the maximum static-speed values and yellow-highlighted values are at least as large as the maximum of SLOW and FAST values.

the earlier part of the data sample before 2004 to estimate the state-conditional speeds of the dynamic strategy used in the evaluation period (similarly to Section 5) for each country. As before, if either $\widehat{a}_{\text{Co}}$ or $\widehat{a}_{\text{Re}}$ fall outside the unit interval $[0,1]$, then we set its value to the nearest endpoint, 0 or 1. Because the estimation period for four states (Bull, Correction, Bear, Rebound) is relatively short (20 years of data split four ways), the estimates are relatively noisy, and the above restriction comes into play for approximately half of the speed estimates in Table 9.

With few exceptions, the recommendation from our state-dependent speed analysis on U.S. equity in Section 5 carries over to international equity markets: employ slower-speed momentum after Corrections and faster-speed momentum after Rebounds. For most countries, the Sharpe ratio of the DYN strategy is higher than the highest static-speed strategy (highlighted green in Table 9). The countries for which their DYN Sharpe ratios are not greater than or equal to the highest static speed are Canada (CAN), Denmark (DNK), Hong Kong (HKG), Japan (JNP), and Norway (NOR). Of these, all but Denmark and Hong Kong have DYN Sharpe ratios at least as high as the maximum of SLOW and FAST. Nevertheless, even for Denmark and Hong Kong, their DYN Sharpe ratios are higher than the average Sharpe ratios of the SLOW and FAST strategies and higher than the Sharpe ratios of some static-speed strategies.

7. Conclusion

We develop a model of time-varying trend to study the dynamics of elementary slow (12-month) and fast (1-month) monthly TS momentum strategies. We use its predictions for the performance of these strategies to estimate the level of persistence in expected returns and noise in realized returns of the U.S. stock market. We find relatively high persistence and high noise, properties that make the union of information in slow and fast momentum signals relevant for detecting trends and momentum turning points.

Accordingly, we use the agreement or disagreement between the uptrend or downtrend indications of slow and fast momentum to characterize four market cycles: Bull, Correction, Bear, and Rebound. Bulls and Bears proxy for uptrend and downtrend, respectively, while Corrections and Rebounds proxy for turning points in trend. We find that these market cycles, despite being simple in construction, possess predictive information for stock market returns and have close connections to the macroeconomy and the business cycle. Bear states, which predict negative expected returns, tend to concentrate early in recessions and coincide with bad macro news, which is difficult to rationalize under the prevailing view of countercyclical risk premia. Bull and Rebound states, which predict positive expected returns, tend to concentrate late in recessions, consistent with a countercyclical view. All four market states can occur in any phase of the business cycle, however, which suggests a role for the sense of cyclicality itself to vary over time in models of time-varying risk premia.

We apply the predictive potential of our market cycles to develop investable strategies with positive unconditional alpha. Intermediate-speed strategies, formed by blending slow and fast TS momentum, vary the exposure to the good bets associated with uptrend (Bull) or downtrend (Bear) phases and the bad bets associated with turning points (Correction or Rebound). For the U.S. stock market, we find that intermediate-speed strategies empirically exhibit many advantages over slow and fast strategies, including higher Sharpe ratios, less severe drawdowns, more positive skewness, and higher significance of alphas. We provide analytical results that generalize our findings for Sharpe ratios and skewness to any market.

We find that both market timing and volatility timing play important but overlooked roles for TS momentum. In particular, we provide a novel model-free decomposition of momentum's alpha into the sum of market-timing and volatility-timing components. Market timing, the focus of recent TS momentum studies, reflects the covariance between strategy positions and subsequent market returns; in contrast, volatility timing reflects the covariance between strategy positions and subsequent squared market return deviations from their mean. For the U.S. stock market, we estimate about two-thirds of TS momentum's alpha over our 50-year evaluation period is attributable to market timing, and the remaining one-third is attributable to volatility timing.

We derive a dynamic momentum strategy that varies its speeds based on market cycles so to maximize its Sharpe ratio. We implement an investable version and find modest, but consistent improvements, when to compared to static strategies, in out-of-sample risk-adjusted performance. We document that these results, with few exceptions, hold up across international equity markets.

Our framework is purposely simple. Based on the sign of past returns, our elementary strategies will either buy or sell. That is, we intentionally use binary signals that do not take into account the magnitude of past returns. Not only does it allow us to appropriately relate our work to previous literature, it allows us to look at four market states and analytically derive the moments of various trading strategies. Further, we are able to tackle a heretofore unanswered question: is it possible to dynamically adjust momentum speed to improve performance characteristics over static strategies? Building a richer model that takes return magnitudes into account is a topic of on-going research.

References

Ahn, D. H., Boudoukh, J., Richardson, M., Whitelaw, R. F., 2002. Partial adjustment or stale prices? implications from stock index and futures return autocorrelations. Review of Financial Studies 15, 655–689.

Albuquerque, R., Eichenbaum, M., Papanikolaou, Rebelo, S., 2015. Beauty contests and iterated expectations in asset markets. Journal of Monetary Economics 76, Supplement, S21–S36.

Allen, F., Morris, S., Shin, H. S., 2006. Beauty contests and iterated expectations in asset markets. Review of Financial Studies 19, 719–752.

Antoniou, C., Doukas, J. A., Subrahmanyam, A., 2013. Cognitive dissonance, sentiment, and momentum. Journal of Financial and Quantitative Analysis 48, 245–275.

Aretz, K., Pope, P. F., 2013. Capacity choice, momentum and long-term reversals. Tech. rep., working Paper.

Arnott, R. D., Clements, M., Kalesnik, V., Linnainmaa, J. T., 2019. Factor momentum, available at SSRN 3116974.

Asness, C. S., 1994. Variables that explain stock returns. Ph.D. thesis, University of Chicago.

Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and momentum everywhere. Journal of Finance 68, 929–985.

Avramov, D., Cheng, S., Hameed, A., 2013. Time-varying momentum payoffs and illiquidity. Tech. rep., working Paper.

Baker, S. R., Bloom, N., Davis, S. J., 2016. Measuring economic policy uncertainty. Quarterly Journal of Economics 131, 1593–1636.

Bandarchuk, P., Hilscher, J., 2012. Sources of momentum profits: Evidence on the irrelevance of characteristics. Review of Finance 17, 809–845.

Banerjee, S., Kaniel, R., Kremer, I., 2009. Price drift as an outcome of differences in higher-order beliefs. Review of Financial Studies 22, 3707–3734.

Barberis, N., Shleifer, A., 2003. Style investing. Journal of Financial Economics 49, 161–199.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Berk, J. B., Green, R. C., Naik, V., 1999. Optimal investment, growth options, and security returns. Journal of Finance 54, 1553–1607.

Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association pp. 171–181.

Boguth, O., Carlson, M., Fisher, A., Simutin, M., 2011. Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas. Journal of Financial Economics 102, 363–389.

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

Cespa, G., Vives, X., 2012. Expectations, liquidity, and short-term trading. Tech. rep., cESifo Working Paper Series.

Chan, W. S., 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. Journal of Financial Economics 70, 223–260.

Chen, Z., Lu, A., 2013. Slow diffusion of information and price momentum in stocks: Evidence from options markets, available at SSRN 2339711.

Christie, A. A., 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. Journal of Financial Economics 10, 407–432.

Chui, A. C., Titman, S., Wei, K. J., 2010. Individualism and momentum around the world. Journal of Finance 65, 361–392.

Cochrane, J. H., 2011. Presidential address: Discount rates. Journal of Finance 66, 1047–1108.

Conrad, J., Kaul, G., 1998. An anatomy of trading strategies. Review of Financial Studies 11, 489–519.

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

Cross, R., Grinfeld, M., Lamba, H., 2006. A mean-field model of investor behaviour. J. Phys. Conf. Ser 55, 55–62.

Cross, R., Grinfeld, M., Lamba, H., Seaman, T., 2005. A threshold model of investor psychology. Physica A: Statistical Mechanics and its Applications 354, 463–478.

Cujean, J., Hasler, M., 2017. Why does return predictability concentrate in bad times? Journal of Finance 72, 2717–2758.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and investor security market under-and overreactions. Journal of Finance 53, 1839–1886.

Daniel, K., Jagannathan, R., Kim, S., 2019. A hidden markov model of momentum. Tech. rep., national Bureau of Economic Research.

Daniel, K., Moskowitz, T. J., 2016. Momentum crashes. Journal of Financial Economics 122, 221–247.

Dasgupta, A., Prat, A., Verardo, M., 2011. The price impact of institutional herding. Review of Financial Studies 24, 892–925.

Ehsani, S., Linnainmaa, J., 2019. Factor momentum and the momentum factor, available at SSRN 3014521.

Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Frazzini, A., 2006. The disposition effect and underreaction to news. Journal of Finance 61, 2017–2046.

Georgopoulou, A., Wang, J. G., 2017. The trend is your friend: Time-series momentum strategies across equity and commodity markets. Review of Finance 31, 1557–1592.

Gertler, M., Karadi, P., 2015. Monetary policy surprises, credit costs, and economic activity. American Economic Journal: Macroeconomics 7, 44–76.

Gervais, S., Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14, 1–27.

Gómez-Cram, R., 2022. Late to recessions: Stocks and the business cycle. Journal of Finance 77, 809–1431.

Goyal, A., Jegadeesh, N., 2017. Cross-sectional and time-series tests of return predictability: What is the difference? Review of Financial Studies 31, 1748–1824.

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

Grinblatt, M., Han, B., 2002. The disposition effect and momentum. Tech. rep., National Bureau of Economic Research.

Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. Journal of Financial Economics 78, 311–339.

Gupta, T., Kelly, B. T., 2019. Factor momentum everywhere, available at SSRN 3300728.

Gutierrez, R. C., Kelley, E. K., 2008. The long-lasting momentum in weekly returns. Journal of Finance 63, 415–447.

Haldane, A. G., 2010. Patience and finance. Tech. rep., speech at Oxford China Business Forum, Beijing.

Harvey, C. R., Hoyle, E., Korgaonkar, R., Rattray, S., Sargaison, M., Van Hemert, O., 2018. The impact of volatility targeting. Journal of Portfolio Management 45, 14–33.

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

Hong, H., Stein, J., 2007. Disagreement and the stock market. Journal of Economic Perspectives 21, 109–128.

Hu, G. X., Pan, J., Wang, J., 2013. Noise as information for illiquidity. Journal of Finance 68, 2341–2382.

Huang, D., Li, J., Wang, L., Zhou, G., 2019. Time-series momentum: Is it there? Journal of Financial Economics Forthcoming.

Hurst, B., Ooi, Y. H., Pedersen, L. H., 2013. Demystifying managed futures. Journal of Investment Management 11, 42–58.

Israel, R., Moskowitz, J. T., 2012. The role of shorting, firm size, and time on market anomalies. Journal of Financial Economics 108, 275–301.

Jagannathan, R., Wang, Z., 1996. The conditional capm and the cross-section of expected returns. Journal of Finance 51, 3–53.

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

Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. Journal of Finance 56, 699–720.

Johnson, T. C., 2002. Rational momentum effects. Journal of Finance 13, 585–608.

Jordan, S. J., 2013. Is momentum a self-fulfilling prophecy? Quantitative Finance 4, 737–748.

Kim, A. Y., Tse, Y., Wald, J. K., 2016. Time series momentum and volatility scaling. Journal of Financial Markets 30, 103–124.

Lee, C., Swaminathan, B., 2000. Price momentum and trading volume. Journal of Finance 55, 2017–2069.

Levine, A., Pedersen, L. H., 2016. Which trend is your friend? Financial Analyst Journal 72, 51–66.

Lewellen, J., 2002. Momentum and autocorrelation in stock returns. Review of Financial Studies 15, 533–564.

Lewellen, J., Nagel, S., 2006. The conditional CAPM does not explain asset-pricing anomalies. Journal of Financial Economics 82, 289–314.

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

Liu, L. X., Zhang, L., 2013. A model of momentum. Tech. rep., working Paper.

Lo, A. W., MacKinlay, A. C., 1990. When are contrarian profits due to stock market overreaction? Review of Financial Studies 3, 175–205.

Maddala, G. S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press.

Makarov, I., Rytchkov, O., 2012. Forecasting the forecasts of others: Implications for asset pricing. Journal of Economic Theory 3, 941–966.

McLean, R. D., 2010. Idiosyncratic risk, long-term reversal, and momentum. Journal of Financial and Quantitative Analysis 45, 883–906.

Moreira, A., Muir, T., 2017. Volatility-managed portfolios. Journal of Finance 72, 1611–1644.

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

Novy-Marx, R., 2012. Is momentum really momentum? Journal of Financial Economics 103, 429–453.

Odean, T., 1998. Volume, volatility, price, and profit when all traders are above average. Journal of Finance 53, 1887–1934.

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

Pflueger, C., Siriwardane, E., Sunderam, A., 2018. A measure of risk appetite for the macroeconomy. Tech. rep., working Paper, National Bureau of Economic Research.

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

Shefrin, H., 2008. A Behavioral Approach to Asset Pricing. Academic Press.

Vayanos, D., Woolley, P., 2012. A theoretical analysis of momentum and value strategies. Tech. rep., working Paper, Paul Wooley Center for the Study of Capital Market Dysfunctionality, London School of Economics.

Vayanos, D., Woolley, P., 2013. An institutional theory of momentum and reversal. Review of Financial Studies 26, 1087–1145.

Verardo, M., 2009. Heterogeneous beliefs and momentum profits. Journal of Financial and Quantitative Analysis 44, 795–822.

Watanabe, M., 2008. Price volatility and investor behavior in an overlapping generations model with information asymmetry. Journal of Finance 63, 229–272.

Zhou, G., Zhu, Y., 2013. An equilibrium model of moving-average predictability and time-series momentum, available at SSRN 2326650.

Internet Appendix A. Proofs

We present an auxiliary lemma, which we use to prove some propositions of the main text.

Lemma 10. *Let Z_1 and Z_2 be bivariate standard normal random variables (mean zero, variance one) with correlation ρ. Then,*

$$\mathbf{E}[Z_1|Z_2 < z] = -\rho\frac{\phi(z)}{\Phi(z)}, \tag{A1}$$

$$\mathbf{E}[Z_1|Z_2 \geq z] = \rho\frac{\phi(z)}{1-\Phi(z)}, \tag{A2}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$ and $\Phi(z) = \int_{-\infty}^{z}\phi(\zeta)d\zeta$ are the probability density and cumulative distribution functions, respectively, of a standard normal random variable.

Proof of Lemma 10. These standard results follow from the fact that the conditional distribution of Z_1 given Z_2 is normal and from expected value calculations for truncated normal random variables. See, for example, Maddala (1983), p. 367. □

Proof of Proposition 1. Recall that $\mu_t = c + \varphi\mu_{t-1} + \varepsilon_t$, where $c = 0$, $\varphi \in (0,1)$ and $\{\varepsilon_t\}$ is normal white noise with variance $\sigma_\varepsilon^2 > 0$. Also, $r_t = \mu_t + z_t$, where $\{z_t\}$ is normal white noise with variance $\sigma_z^2 > 0$, and where each pair of innovations (ε_t, z_t) is jointly normal with correlation coefficient $\lambda \in (-1,1)$. Recall that the variance of r_t is $\sigma_r^2 = \frac{\sigma_\varepsilon^2}{1-\varphi^2} + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z$ and the noise level θ is defined by $\theta = \frac{\sigma_z^2+\lambda\sigma_\varepsilon\sigma_z}{\sigma_r^2}$.

Consider signal $s_{t,K} = \frac{1}{K}\sum_{i=1}^{K} r_{t-i}$, the arithmetic K-period average trailing excess return, for $K = 1, 2, \ldots$. The mean and variance of $s_{t,K}$ are as follows. The mean is: $\mathbf{E}[s_{t,K}] = \mathbf{E}\left[\frac{1}{K}\sum_{i=1}^{K} r_{t-i}\right] = \frac{1}{K}\sum_{i=1}^{K}\mathbf{E}\left[r_{t-i}\right] = 0$, where the first equality follows by definition of $s_{t,K}$; the second equality follows by linearity of the expectation operator; and the final equality follows from $\mathbf{E}[r_t] = 0$ for all t.

If $i = j$, then $\mathbf{E}[r_{t-i}r_{t-j}] = \mathbf{E}[r_{t-i}^2] = \mathbf{Var}[r_{t-i}] + (\mathbf{E}[r_{t-i}])^2 = \sigma_r^2$, where the second equality follows by the definition of variance; and the final equality follows from $\mathbf{E}[r_{t-i}] = 0$. If $i < j$, then $\mathbf{E}[r_{t-i}r_{t-j}] = \mathbf{Cov}[r_{t-i}, r_{t-j}] = \mathbf{Cov}[\mu_{t-i} + z_{t-i}, \mu_{t-j} + z_{t-j}] = \mathbf{Cov}[\mu_{t-i}, \mu_{t-j}] + \mathbf{Cov}[\mu_{t-i}, z_{t-j}] = \varphi^{j-i}\frac{\sigma_\varepsilon^2}{1-\varphi^2} + \varphi^{j-i}\lambda\sigma_\varepsilon\sigma_z = \varphi^{j-i}(1-\theta)\sigma_r^2$, where the first equality follows by definition of covariance and because $\mathbf{E}[r_{t-i}] = \mathbf{E}[r_{t-j}] = 0$; the second equality follows by definition of r_t; the third equality follows because z_{t-i} and z_{t-j} are independent for $i < j$, and z_{t-i} and μ_{t-j} are independent for $i < j$; the fourth equality follows by the standard autocovariance of the weak-sense stationary AR(1) process $\{\mu_t\}$ and because $\mathbf{Cov}[\mu_{t-i}, z_{t-j}] = \varphi^{j-i}\mathbf{Cov}[\varepsilon_{t-j}, z_{t-j}] = \varphi^{j-i}\lambda\sigma_\varepsilon\sigma_z$ for $i < j$; and the final equality follows by the definition of θ. Thus,

$$\mathbf{E}[r_{t-i}r_{t-j}] = \begin{cases} \sigma_r^2, & \text{if } i = j, \\ \varphi^{|i-j|}(1-\theta)\sigma_r^2, & \text{if } i \neq j. \end{cases} \tag{A3}$$

and

$$\begin{aligned}
\mathbf{Var}\left[s_{t,K}\right] &= \mathbf{E}\left[s_{t,K}^2\right] = \mathbf{E}\left[\left(\frac{1}{K}\sum_{i=1}^{K} r_{t-i}\right)^2\right] = \frac{1}{K^2}\sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{E}\left[r_{t-i}r_{t-j}\right] \\
&= \frac{1}{K^2}\left(K\sigma_r^2 + 2\sum_{i=1}^{K-1}\sum_{j=i+1}^{K}\varphi^{j-i}(1-\theta)\sigma_r^2\right) \\
&= \frac{\sigma_r^2}{K} + \frac{2(1-\theta)\sigma_r^2}{K^2}\sum_{i=1}^{K-1}\sum_{j=1}^{i}\varphi^j \\
&= \frac{\sigma_r^2}{K} + \frac{2(1-\theta)\sigma_r^2}{K^2}\sum_{i=1}^{K-1}\varphi\frac{1-\varphi^i}{1-\varphi} \\
&= \sigma_r^2\left[\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}\right] > 0,
\end{aligned} \tag{A4}$$

where the first equality follows by definition of variance and because $\mathbf{E}[s_{t,K}] = 0$; the second equality follows by definition of $s_{t,K}$; the third equality follows by expansion of the square and linearity of the expectation operator; the fourth equality follows by (A3); the fifth equality follows by the summation reordering identity $\sum_{i=1}^{K-1}\sum_{j=i+1}^{K}\varphi^{j-i} = \sum_{i=1}^{K-1}\sum_{j=1}^{i}\varphi^j$; the sixth equality follows by identity $\sum_{j=1}^{i}\varphi^j = \varphi\frac{1-\varphi^i}{1-\varphi}$; the seventh equality follows by identity $\sum_{i=1}^{K-1}(1-\varphi^i) = K-1-\varphi\frac{1-\varphi^{K-1}}{1-\varphi}$ and simplification; and the final inequality follows by standard properties of variance.

The covariance between r_t and $s_{t,K}$ satisfies:

$$\begin{aligned}
\mathbf{Cov}[r_t, s_{t,K}] &= \mathbf{E}\left[r_t \cdot \frac{1}{K}\sum_{i=1}^{K} r_{t-i}\right] = \frac{1}{K}\sum_{i=1}^{K}\mathbf{E}[r_t, r_{t-i}] = \frac{1}{K}\sum_{i=1}^{K}\varphi^i(1-\theta)\sigma_r^2 \\
&= \frac{(1-\theta)\sigma_r^2}{K}\varphi\frac{1-\varphi^K}{1-\varphi},
\end{aligned} \tag{A5}$$

where the first equality follows by definition of covariance and $s_{t,K}$ and because $\mathbf{E}[r_t] = \mathbf{E}[s_{t,K}] = 0$; the second equality follows by linearity of the expectation operator; the third equality follows by (A3); and the fourth equality follows by identity $\sum_{j=1}^{i}\varphi^j = \varphi\frac{1-\varphi^i}{1-\varphi}$. This covariance (A5) is strictly positive if $\theta < 1$, because it implies $1-\theta > 0$, and because $\varphi \in (0,1)$ implies $0 < \varphi^K < 1$ and $1-\varphi^K > 0$ for all $K = 1, 2, \ldots$.

By Section 2.1, excess returns of the two elementary time series (TS) momentum strategies, SLOW and FAST, can be expressed as:

$$r_{\text{SLOW},t} = r_t \cdot \text{sign}\,(s_{t,12}) \quad \text{and} \quad r_{\text{FAST},t} = r_t \cdot \text{sign}(s_{t,1}), \tag{A6}$$

where r_t is the excess return of the market. Because r_t and $s_{t,K}$ are linear functions of jointly normal sequences $\{\varepsilon_t\}$ and $\{z_t\}$, they are jointly normally distributed.

Let $Z_1 = \frac{r_t}{\sigma_r}$ and $Z_2 = \frac{s_{t,K}}{\mathbf{SD}[s_{t,K}]}$, where $\mathbf{SD}[s_{t,K}] = \sqrt{\mathbf{Var}[s_{t,K}]}$ is the standard deviation of $s_{t,K}$. Then, Z_1 and Z_2 are mean-zero unit-variance bivariate normal random variables with correlation ρ, which satisfies:

$$\begin{aligned}\rho = \mathbf{Corr}[Z_1, Z_2] = \mathbf{Cov}[Z_1, Z_2] = \mathbf{Cov}\left[\frac{r_t}{\sigma_r}, \frac{s_{t,K}}{\mathbf{SD}[s_{t,K}]}\right] = \frac{\mathbf{Cov}[r_t, s_{t,K}]}{\sigma_r \cdot \mathbf{SD}[s_{t,K}]} \\ = \frac{\frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}}{\sqrt{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}},\end{aligned} \tag{A7}$$

where the first equality follows by definition of ρ; the second equality follows because Z_1 and Z_2 have unit variance; the third equality follows by definition of Z_1 and Z_2; the fourth equality follows by the bilinearity of the covariance operator; and the fifth equality follows by substitution from (A4) and (A5) and cancellation of σ_r^2 in numerator and denominator. The correlation ρ of (A7) is strictly positive if $\theta < 1$, in which case the covariance in the fourth equality is strictly positive by (A4) and (A5), as noted earlier.

The expected strategy return satisfies:

$$\begin{aligned}\mathbf{E}[r_t \cdot \text{sign}\,(s_{t,K})] &= \mathbf{E}\left[\sigma_r \frac{r_t}{\sigma_r} \cdot \text{sign}\left(\frac{s_{t,K}}{\mathbf{SD}[s_{t,K}]}\right)\right] = \sigma_r \cdot \mathbf{E}[Z_1 \cdot \text{sign}\,(Z_2)] \\ &= \sigma_r\left(-\mathbf{E}\left[Z_1 | Z_2 < 0\right]\mathbf{P}\left[Z_2 < 0\right] + \mathbf{E}\left[Z_1 | Z_2 \geq 0\right]\mathbf{P}\left[Z_2 \geq 0\right]\right) \\ &= \sigma_r\left(\rho\frac{\phi\,(0)}{\Phi\,(0)}\Phi\,(0) + \rho\frac{\phi\,(0)}{1-\Phi\,(0)}\left(1-\Phi\,(0)\right)\right) \\ &= \sqrt{\frac{2}{\pi}}\sigma_r \frac{\frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}}{\sqrt{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}},\end{aligned} \tag{A8}$$

where the first equality follows because σ_r and $\mathbf{SD}[s_{t,K}]$ are positive; the second equality follows by definition of Z_1 and Z_2; the third equality follows by the law of total expectation; the fourth equality follows by (A1) and (A2), respectively, of Lemma 10; and the fifth equality follows because $\Phi(0) = \frac{1}{2}$, $\phi(0) = \frac{1}{\sqrt{2\pi}}$, and by (A7). The expected value of (A8) is strictly positive if $\theta < 1$ by the discussion following (A7).

The variance of the strategy return satisfies:

$$\begin{aligned}\mathbf{Var}[r_t \cdot \text{sign}\,(s_{t,K})] &= \mathbf{E}[(r_t \cdot \text{sign}\,(s_{t,K}))^2] - (\mathbf{E}[r_t \cdot \text{sign}\,(s_{t,K})])^2 \\ &= \sigma_r^2 - \left(\frac{2}{\pi}\sigma_r^2 \frac{\frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}}{\sqrt{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}}\right)^2 \\ &= \sigma_r^2 \left[1 - \frac{\frac{2}{\pi}\left(\frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}\right)^2}{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}\right] > 0,\end{aligned} \tag{A9}$$

where the first equality follows by definition of variance; the second equality follows because $\mathbf{E}[(r_t \cdot \text{sign}\,(s_{t,K}))^2] = \mathbf{E}[r_t^2] = \mathbf{Var}[r_t] = \sigma_r^2$ and from substitution of (A8); the third equality follows by factorization; and the final inequality follows by standard properties of variance.

The Sharpe ratio of the strategy return satisfies:

$$\begin{aligned}\mathbf{Sharpe}[r_t \cdot \text{sign}\,(s_{t,K})] &= \frac{\mathbf{E}[r_t \cdot \text{sign}\,(s_{t,K})]}{\mathbf{SD}[r_t \cdot \text{sign}\,(s_{t,K})]} \\ &= \frac{\frac{\sqrt{\frac{2}{\pi}}\sigma_r \frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}}{\sqrt{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}}}{\sqrt{\sigma_r^2\left[1 - \frac{\frac{2}{\pi}\left(\frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi}\right)^2}{\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}}\right]}} \\ &= \frac{\sqrt{\frac{2}{\pi}}(1-\theta)\varphi}{\sqrt{\frac{(1-\theta)[K(1-\varphi^2)-2\varphi(1-\varphi^K)]+\theta K(1-\varphi)^2}{(1-\varphi^K)^2} - \frac{2}{\pi}(1-\theta)^2\varphi^2}},\end{aligned} \tag{A10}$$

where the first equality follows by definition; the second equality follows by substitution from (A8) and (A9); and the third equality follows by simplification. If $\theta < 1$, then the Sharpe ratio of (A10) is strictly positive; otherwise, it is zero if $\theta = 1$ and negative if $\theta > 1$. By (A6), evaluation of (A10) at $K = 1$ and $K = 12$ with simplification and multiplying by $\sqrt{12}$ to annualize completes the proof. □

Proof of Proposition 2. We first show that $\theta < 1$. Recall that $\theta = \frac{\sigma_z^2+\lambda\sigma_\varepsilon\sigma_z}{\sigma_r^2}$ and $\sigma_r^2 = \frac{\sigma_\varepsilon^2}{1-\varphi^2}+\sigma_z^2+2\lambda\sigma_\varepsilon\sigma_z$. So, $\theta < 1$ if and only if $\sigma_z^2+\lambda\sigma_\varepsilon\sigma_z < \frac{\sigma_\varepsilon^2}{1-\varphi^2}+\sigma_z^2+2\lambda\sigma_\varepsilon\sigma_z$, which is equivalent to $\lambda > -\frac{1}{1-\varphi^2}\frac{\sigma_\varepsilon}{\sigma_z}$. In (D78) of Internet Appendix D, we show that $\mathbf{E}[r_t|r_{t-1}, r_{t-2}, \ldots] = \frac{c}{1-\delta} + (\varphi - \delta)\left(r_{t-1} + \delta r_{t-2} + \delta^2 r_{t-3} + \cdots\right)$, where $\delta = \varphi\frac{\sigma_z}{\sigma_w}$ and $\sigma_w^2 = \sigma_\varepsilon^2 + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z$. Because $\varphi > 0$ and $\delta > 0$, this conditional expectation is positively related to all past returns if and only if $\delta < \varphi$, which is equivalent to $\sigma_z^2 < \sigma_w^2$. This condition implies that $\lambda > -\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_z}$, which implies $\lambda > -\frac{1}{1-\varphi^2}\frac{\sigma_\varepsilon}{\sigma_z}$ because $\frac{1}{2} < \frac{1}{1-\varphi^2}$. Therefore, $\theta < 1$.

Next, we show that $\theta > 0$. By definition of θ, we have $\theta > 0$ if and only if $\sigma_z^2 + \lambda\sigma_\varepsilon\sigma_z > 0$, which is equivalent to $\lambda > -\frac{\sigma_z}{\sigma_\varepsilon}$. The condition $\sigma_z^2 \geq \frac{1}{2}\sigma_\varepsilon^2$ in the statement of this proposition implies that $\frac{\frac{1}{2}\sigma_\varepsilon^2}{\sigma_z^2} \leq 1$. Thus, $\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_z} = \frac{\frac{1}{2}\sigma_\varepsilon^2}{\sigma_z^2}\frac{\sigma_z}{\sigma_\varepsilon} < \frac{\sigma_z}{\sigma_\varepsilon}$. Therefore, $\lambda > -\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_z}$ implies that $\lambda > -\frac{\sigma_z}{\sigma_\varepsilon}$, which implies $\theta > 0$ and completes the proof.

$\square$

Proof of Proposition 3. Recall that $\mu_t = c + \varphi\mu_{t-1} + \varepsilon_t$, where $c = 0$, $\varphi \in (0,1)$, and $\{\varepsilon_t\}$ is normal white noise with variance $\sigma_\varepsilon^2 > 0$. Also, $r_t = \mu_t + z_t$, where $\{z_t\}$ is white noise with variance $\sigma_z^2 > 0$ and where each pair of innovations (ε_t, z_t) is jointly normal with correlation coefficient $\lambda \in (-1,1)$. For $\lambda = 0$, $\{\varepsilon_t\}$ and $\{z_t\}$ are independent and the variance of r_t is $\sigma_r^2 = \frac{\sigma_\varepsilon^2}{1-\varphi^2} + \sigma_z^2$ with the noise level $\theta = \frac{\sigma_z^2}{\sigma_r^2}$.

Let $y = [r_t \ s_{t,K}]'$ be a column vector of the return and K-month trailing return signal, respectively, where $s_{t,K} = \frac{1}{K}\sum_{i=1}^{K} r_{t-i}$ and $\mathbf{E}[s_{t,K}] = 0$ as in the proof of Proposition 1, and $'$ is the transpose operator. Let $x = [\mu_{t-1} \ \mu_{t-2}]'$ be a column vector of the two previous month's expected returns. Note that each component of y and x has zero mean. Because all components of y and x are linear combinations of independent normal random variables, y and x are multivariate normal random variables with covariance matrix

$$\begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix}, \tag{A11}$$

where $\Sigma_{yy} = \mathbf{Cov}[y,y]$, $\Sigma_{yx} = \mathbf{Cov}[y,x] = \Sigma_{xy}'$, and $\Sigma_{xx} = \mathbf{Cov}[x,x]$ are each 2×2 submatrices. By standard properties of the multivariate normal distribution, the conditional distribution of y given x is also multivariate normally distributed, $y|x \sim \mathcal{N}(\mu_{y|x}, \Sigma_{y|x})$, where

$$\mu_{y|x} = \Sigma_{yx}\Sigma_{xx}^{-1}x, \tag{A12}$$

$$\Sigma_{y|x} = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{yx}. \tag{A13}$$

The covariance between returns and trailing expected returns satisfies

$$\mathbf{Cov}[r_t, \mu_{t-i}] = \mathbf{Cov}[\mu_t + z_t, \mu_{t-i}] = \mathbf{Cov}[\mu_t, \mu_{t-i}] = (1-\theta)\sigma_r^2\varphi^i, \tag{A14}$$

for $i = 1, 2$, where the first equality follows by definition of r_t; the second equality follows by independence of z_t and μ_{t-i}; and the final equality follows by the autocovariance properties of the AR(1) expected return process. The covariances between trailing return signal $s_{t,K}$

and first and second previous expected returns, respectively, satisfy:

$$\mathbf{Cov}[s_{t,K}, \mu_{t-1}] = \frac{1}{K}\sum_{i=0}^{K-1}(1-\theta)\sigma_r^2\varphi^i = \frac{1-\theta}{K}\sigma_r^2\frac{1-\varphi^K}{1-\varphi}, \tag{A15}$$

$$\mathbf{Cov}[s_{t,K}, \mu_{t-2}] = \frac{1}{K}\left[(1-\theta)\sigma_r^2\varphi + \sum_{i=0}^{K-2}(1-\theta)\sigma_r^2\varphi^i\right] = \frac{1-\theta}{K}\sigma_r^2\left[\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right], \tag{A16}$$

where the first equality in each equality chain follows by the definition of $s_{t,K}$, the bilinearity of the covariance operator, because $\mathbf{Cov}[r_{t-i}, \mu_{t-1}] = \mathbf{Cov}[\mu_{t-i}, \mu_{t-1}]$ by definition of r_t and independence of z_t and μ_{t-1}, and by the autocovariance properties of the AR(1) expected return process evaluated at lags 0 up to $K-1$; and the final equalities follow by application of the identity $\sum_{i=0}^{N-1}\varphi^i = \frac{1-\varphi^N}{1-\varphi}$.

Thus, the covariance sub-matrices satisfy the following:

$$\Sigma_{yy} = \sigma_r^2 \cdot \begin{bmatrix} 1 & \frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi} \\ \frac{(1-\theta)}{K}\varphi\frac{1-\varphi^K}{1-\varphi} & \frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K} \end{bmatrix}, \tag{A17}$$

$$\Sigma_{yx} = \Sigma_{xy}' = \sigma_r^2(1-\theta) \cdot \begin{bmatrix} \varphi & \varphi^2 \\ \frac{1}{K}\frac{1-\varphi^K}{1-\varphi} & \frac{1}{K}\left(\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right) \end{bmatrix}, \tag{A18}$$

$$\Sigma_{xx} = \sigma_r^2(1-\theta) \cdot \begin{bmatrix} 1 & \varphi \\ \varphi & 1 \end{bmatrix}. \tag{A19}$$

Equation (A17) follows by definition, $\sigma_r^2 = \mathbf{Var}[r_t]$, and from(A5) and (A4). Equation (A18) follows by (A14), (A15), and (A16). Equation (A19) follows from the autocovariance properties of the AR(1) expected return process $\{\mu_t\}$.

Using (A18) and (A19), from standard matrix operations and simplification, we obtain:

$$\begin{aligned}
\Sigma_{yx}\Sigma_{xx}^{-1} &= \sigma_r^2(1-\theta) \cdot \begin{bmatrix} \varphi & \varphi^2 \\ \frac{1}{K}\frac{1-\varphi^K}{1-\varphi} & \frac{1}{K}\left(\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right) \end{bmatrix} \left(\sigma_r^2(1-\theta) \cdot \begin{bmatrix} 1 & \varphi \\ \varphi & 1 \end{bmatrix}\right)^{-1} \\
&= \begin{bmatrix} \varphi & \varphi^2 \\ \frac{1}{K}\frac{1-\varphi^K}{1-\varphi} & \frac{1}{K}\left(\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right) \end{bmatrix} \begin{bmatrix} \frac{1}{1-\varphi^2} & \frac{-\varphi}{1-\varphi^2} \\ \frac{-\varphi}{1-\varphi^2} & \frac{1}{1-\varphi^2} \end{bmatrix} \\
&= \begin{bmatrix} \varphi & 0 \\ \frac{1}{K} & \frac{1}{K}\frac{1-\varphi^{K-1}}{1-\varphi} \end{bmatrix}.
\end{aligned} \tag{A20}$$

From (A12) and (A20) we obtain:

$$\mathbf{E}[r_t|\mu_{t-1},\mu_{t-2}] = \varphi\mu_{t-1}, \tag{A21}$$

$$\mathbf{E}[s_{t,K}|\mu_{t-1},\mu_{t-2}] = \frac{1}{K}\left(\mu_{t-1} + \frac{1-\varphi^{K-1}}{1-\varphi}\mu_{t-2}\right). \tag{A22}$$

From (A20) and the transpose of (A18), and from standard matrix operations and simplification, we obtain:

$$\begin{aligned}\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy} &= \begin{bmatrix} \varphi & 0 \\ \frac{1}{K} & \frac{1}{K}\frac{1-\varphi^{K-1}}{1-\varphi} \end{bmatrix} \cdot \left(\sigma_r^2(1-\theta)\cdot\begin{bmatrix} \varphi & \frac{1}{K}\frac{1-\varphi^K}{1-\varphi} \\ \varphi^2 & \frac{1}{K}\left(\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right)\end{bmatrix}\right) \\ &= \sigma_r^2(1-\theta)\cdot\begin{bmatrix} \varphi^2 & \frac{1}{K}\varphi\frac{1-\varphi^K}{1-\varphi} \\ \frac{1}{K}\left(\varphi+\varphi^2\frac{1-\varphi^{K-1}}{1-\varphi}\right) & \frac{1}{K^2}\left(\frac{1-\varphi^K}{1-\varphi} + \frac{1-\varphi^{K-1}}{1-\varphi}\left(\varphi + \frac{1-\varphi^{K-1}}{1-\varphi}\right)\right)\end{bmatrix}.\end{aligned} \tag{A23}$$

By (A13), (A17), and (A23), with simplification we obtain:

$$\begin{aligned}\mathbf{Var}[s_{t,K}|\mu_{t-1},\mu_{t-2}] &= \sigma_r^2\left(\frac{(1-\theta)}{K}\frac{K(1-\varphi^2)-2\varphi(1-\varphi^K)}{K(1-\varphi)^2} + \frac{\theta}{K}\right) \\ &\quad - \sigma_r^2(1-\theta)\left(\frac{1}{K^2}\left(\frac{1-\varphi^K}{1-\varphi} + \frac{1-\varphi^{K-1}}{1-\varphi}\left(\varphi+\frac{1-\varphi^{K-1}}{1-\varphi}\right)\right)\right) \\ &= \frac{\sigma_r^2(1-\theta)}{K}\left[\frac{1+\varphi}{1-\varphi} + \frac{1}{K} - \frac{2(1-\varphi^K)+(1-\varphi^{K-1})^2}{K(1-\varphi)^2}\right] + \frac{\sigma_r^2\theta}{K}.\end{aligned} \tag{A24}$$

The conditional expected value of the sign of the momentum signal given μ_{t-1} and μ_{t-2} satisfies:

$$\begin{aligned}\mathbf{E}[\mathrm{sign}(s_{t,K})|\mu_{t-1},\mu_{t-2}] &= \mathbf{P}[s_{t,K}>0|\mu_{t-1},\mu_{t-2}] - \mathbf{P}[s_{t,K}\leq 0|\mu_{t-1},\mu_{t-2}] \\ &= 2\mathbf{P}[s_{t,K}>0|\mu_{t-1},\mu_{t-2}] - 1 \\ &= 2\mathbf{P}\left[\frac{s_{t,K}-\mathbf{E}[s_{t,K}|\mu_{t-1},\mu_{t-2}]}{\mathbf{SD}[s_{t,K}|\mu_{t-1},\mu_{t-2}]} > -\frac{\mathbf{E}[s_{t,K}|\mu_{t-1},\mu_{t-2}]}{\mathbf{SD}[s_{t,K}|\mu_{t-1},\mu_{t-2}]}\middle|\mu_{t-1},\mu_{t-2}\right] - 1 \\ &= 2\Phi\left(\frac{\mathbf{E}[s_{t,K}|\mu_{t-1},\mu_{t-2}]}{\mathbf{SD}[s_{t,K}|\mu_{t-1},\mu_{t-2}]}\right) - 1 \\ &= 2\Phi\left(\frac{\frac{1}{K}\left(\mu_{t-1}+\frac{1-\varphi^{K-1}}{1-\varphi}\mu_{t-2}\right)}{\sigma_r\sqrt{\frac{(1-\theta)}{K}\left[\frac{1+\varphi}{1-\varphi}+\frac{1}{K}-\frac{2(1-\varphi^K)+(1-\varphi^{K-1})^2}{K(1-\varphi)^2}\right]+\frac{\theta}{K}}}\right) - 1,\end{aligned} \tag{A25}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and where the first equality follows by the law of total expectation; the second equality follows by difference of probabilities of complementary events; the third equality follows by subtracting the conditional mean and dividing by conditional standard deviation ($\mathbf{SD}[s_{t,K}|\mu_{t-1},\mu_{t-2}]$) on both sides of the interior inequality; the fourth equality follows because the conditional distribution of demeaned and inverse-standard-deviation scaled $s_{t,K}$ is conditional standard normal and by $1-\Phi(-z)=\Phi(z)$ for all z; and the final equality follows by substitution from (A22) and the square root of (A24).

By (A13), (A17), and (A23) with simplification we obtain $\mathbf{Cov}[r_t, s_{t,K}|\mu_{t-1},\mu_{t-2}]=0$, which implies that r_t and $s_{t,K}$ are conditionally independent given μ_{t-1} and μ_{t-2}. Therefore, the conditional expected strategy return given μ_{t-1} and μ_{t-2} satisfies:

$$\begin{aligned}\mathbf{E}[r_t\cdot \operatorname{sign}(s_{t,K})|\mu_{t-1},\mu_{t-2}] &= \mathbf{E}[r_t|\mu_{t-1},\mu_{t-2}]\cdot \mathbf{E}[\operatorname{sign}(s_{t,K})|\mu_{t-1},\mu_{t-2}] \\ &= \varphi\mu_{t-1}\cdot\left[2\Phi\left(\frac{\frac{1}{K}\left(\mu_{t-1}+\frac{1-\varphi^{K-1}}{1-\varphi}\mu_{t-2}\right)}{\sigma_r\sqrt{\frac{(1-\theta)}{K}\left[\frac{1+\varphi}{1-\varphi}+\frac{1}{K}-\frac{2(1-\varphi^K)+(1-\varphi^{K-1})^2}{K(1-\varphi)^2}\right]+\frac{\theta}{K}}}\right)-1\right],\end{aligned} \tag{A26}$$

where the second equality follows by substitution from (A21) and (A25). Evaluation of (A26) at $K=1$ and $K=12$ with simplification we obtain the conditional expected returns of FAST and SLOW, respectively, given the most recent two (unobservable) consecutive return means, as stated in the proposition:

$$\mathbf{E}[r_{\text{FAST},t}|\mu_{t-1},\mu_{t-2}] = \varphi\mu_{t-1}\cdot\left[2\Phi\left(\frac{\mu_{t-1}}{\sigma_r\sqrt{\theta}}\right)-1\right], \tag{8}$$

$$\mathbf{E}[r_{\text{SLOW},t}|\mu_{t-1},\mu_{t-2}] = \varphi\mu_{t-1}\cdot\left[2\Phi\left(\frac{\frac{1}{12}\left(\mu_{t-1}+\frac{1-\varphi^{11}}{1-\varphi}\mu_{t-2}\right)}{\sigma_r\sqrt{\frac{(1-\theta)}{12}\left[\frac{1+\varphi}{1-\varphi}+\frac{1}{12}-\frac{2(1-\varphi^{12})+(1-\varphi^{11})^2}{12(1-\varphi)^2}\right]+\frac{\theta}{12}}}\right)-1\right], \tag{9}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Finally, we show (8) is strictly positive for $\mu_{t-1}\neq 0$ and obtain the inequalities (10) of the proposition. We use the identity: $\operatorname{sign}(2\Phi(z)-1)=\operatorname{sign}(z)$ for all z. Therefore, (8) is positive because, if $\mu_{t-1}>0$, then $2\Phi\left(\frac{\mu_{t-1}}{v\sqrt{\theta}}\right)-1>0$, so the product $\varphi\mu_{t-1}\cdot\left[2\Phi\left(\frac{\mu_{t-1}}{v\sqrt{\theta}}\right)-1\right]$ is positive; whereas, if $\mu_{t-1}<0$, then $2\Phi\left(\frac{\mu_{t-1}}{v\sqrt{\theta}}\right)-1<0$, so the product is positive. Note that this argument also proves the first inequality of (10).

For the second inequality of (10), we consider two cases. Case 1: $\mu_{t-1}<0<\mu_{t-2}$.

In this case, $|\mu_{t-1}| < \frac{1-\varphi^{12}}{1-\varphi}|\mu_{t-2}|$ implies $-\mu_{t-1} < \frac{1-\varphi^{12}}{1-\varphi}\mu_{t-2}$, which implies $\mu_{t-1} < 0 < \mu_{t-1} + \frac{1-\varphi^{12}}{1-\varphi}\mu_{t-2}$. Case 2: $\mu_{t-1} > 0 > \mu_{t-2}$. In this case, $|\mu_{t-1}| < \frac{1-\varphi^{12}}{1-\varphi}|\mu_{t-2}|$ implies $\mu_{t-1} < -\frac{1-\varphi^{12}}{1-\varphi}\mu_{t-2}$, which implies $\mu_{t-1} + \frac{1-\varphi^{12}}{1-\varphi}\mu_{t-2} < 0 < \mu_{t-1}$. In either case, μ_{t-1} and $\mu_{t-1} + \frac{1-\varphi^{12}}{1-\varphi}\mu_{t-2}$ are opposite sign. Therefore, by the identity $\text{sign}(2\Phi(z) - 1) = \text{sign}(z)$ for all z, the term in square brackets of (9) is always opposite sign of $\varphi\mu_{t-1}$ and, therefore, its product with $\varphi\mu_{t-1}$ is strictly negative, which completes the proof. $\square$

Proof of Proposition 4. We show first that the exact decomposition for the Sharpe ratio in (16) holds. First, note that

$$\begin{aligned}\mathbf{E}[r_{t+1}(a)] &= \mathbf{E}[((1-a)w_{\text{SLOW},t} + a w_{\text{FAST},t})r_{t+1}] \\ &= (1-a)\mathbf{E}[r_{\text{SLOW},t+1}] + a\mathbf{E}[r_{\text{FAST},t+1}],\end{aligned} \tag{A27}$$

where the first equality follows by (12), and the second equality follows by definition of SLOW and FAST in Section 2.1. Second,

$$\begin{aligned}\mathbf{E}[w_{\text{SLOW},t}\, w_{\text{FAST},t}\, r_{t+1}^2] &= \mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Bu}\\\text{Be}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Bu}\\\text{Be}\end{matrix}}] - \mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}] \\ &= (\mathbf{E}[r_{t+1}^2] - \mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]) - \mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}] \\ &= \mathbf{E}[r_{t+1}^2] - 2\mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}],\end{aligned} \tag{A28}$$

where the first equality follows by the law of total expectation, the fact that SLOW and FAST have the same signs after Bull or Bear and the opposite signs after Correction or Rebound, and the fact that the weights of SLOW and FAST each have magnitude one; and the second equality follows by the law of total expectation. Third,

$$\mathbf{Var}[w_{\text{SLOW},t}\, r_{t+1}] = \mathbf{E}[r_{t+1}^2] - (\mathbf{E}[w_{\text{SLOW},t}\, r_{t+1}])^2, \tag{A29}$$

$$\mathbf{Var}[w_{\text{FAST},t}\, r_{t+1}] = \mathbf{E}[r_{t+1}^2] - (\mathbf{E}[w_{\text{FAST},t}\, r_{t+1}])^2, \tag{A30}$$

$$\begin{aligned}\mathbf{Cov}[w_{\text{SLOW},t}\, r_{t+1}, w_{\text{FAST},t}\, r_{t+1}] &= \mathbf{E}[w_{\text{SLOW},t}\, w_{\text{FAST},t}\, r_{t+1}^2] - \mathbf{E}[w_{\text{SLOW},t}\, r_{t+1}]\mathbf{E}[w_{\text{FAST},t}\, r_{t+1}] \\ &= \mathbf{E}[r_{t+1}^2] - 2\mathbf{E}[r_{t+1}^2|{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}]\mathbf{P}[{\tiny\begin{matrix}\text{Co}\\\text{Re}\end{matrix}}] \\ &\quad - \mathbf{E}[w_{\text{SLOW},t}\, r_{t+1}]\mathbf{E}[w_{\text{FAST},t}\, r_{t+1}],\end{aligned} \tag{A31}$$

where the first two equalities follow by definition of variance and the fact that the weights of SLOW and FAST have magnitude one; the third equality follows by definition of covariance; and the fourth equality follows by (A28). Using these facts, the variance of a momentum strategy of any speed satisfies

$$
\begin{aligned}
\mathbf{Var}[r_{t+1}(a)] &= \mathbf{Var}[((1-a)w_{\mathrm{SLOW},t} + a\, w_{\mathrm{FAST},t})\, r_{t+1}] \\
&= (1-a)^2\mathbf{Var}[w_{\mathrm{SLOW},t}\, r_{t+1}] + a^2\mathbf{Var}[w_{\mathrm{FAST},t}\, r_{t+1}] \\
&\quad + 2(1-a)a\mathbf{Cov}[w_{\mathrm{SLOW},t}\, r_{t+1}, w_{\mathrm{FAST},t}\, r_{t+1}] \\
&= (1-a)^2[\mathbf{E}[r_{t+1}^2] - (\mathbf{E}[w_{\mathrm{SLOW},t} r_{t+1}])^2] + a^2[\mathbf{E}[r_{t+1}^2] - (\mathbf{E}[w_{\mathrm{FAST},t} r_{t+1}])^2] \\
&\quad + 2(1-a)a\,\{\mathbf{E}[r_{t+1}^2] - 2\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}] - \mathbf{E}[w_{\mathrm{SLOW},t} r_{t+1}]\mathbf{E}[w_{\mathrm{FAST},t} r_{t+1}]\} \\
&= \mathbf{E}[r_{t+1}^2] - 4(1-a)a\,\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}] \\
&\quad - \{(1-a)\mathbf{E}[w_{\mathrm{SLOW},t} r_{t+1}] + a\,\mathbf{E}[w_{\mathrm{FAST},t} r_{t+1}]\}^2 \\
&= (\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}] + \mathbf{E}[r_{t+1}^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}]) - 4(1-a)a\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}] \\
&\qquad - \{\mathbf{E}[((1-a)w_{\mathrm{SLOW},t} + a w_{\mathrm{FAST},t}) r_{t+1}]\}^2 \\
&= \mathbf{E}[r_{t+1}^2|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}] + (2a-1)^2\mathbf{E}[r_t^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}] - (\mathbf{E}[r_{t+1}(a)])^2, \qquad \text{(A32)}
\end{aligned}
$$

where the first equality follows by definition of $r_{t+1}(a)$ in (12); the second equality follows by definition of variance of a sum; the third equality follows by substitution of (A29), (A30), and (A31); the fourth equality follows by collecting terms; the fifth equality follows by the law of total expectation; and the final equality follows by definition of $r_{t+1}(a)$ in (12).

Next, note that $D(a, \mu)$ in (17), whose definition is repeated here for convenience, satisfies

$$
\begin{aligned}
D(a,\mu) &= \sqrt{\frac{\mathbf{E}[r_{t+1}^2] - \mu^2}{\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}] + (2a-1)^2\mathbf{E}[r_{t+1}^2|{}^{\mathrm{Co}}_{\mathrm{Re}}]\mathbf{P}[{}^{\mathrm{Co}}_{\mathrm{Re}}] - (\mathbf{E}[r_{t+1}(a)])^2}}, \qquad (17) \\
&= \sqrt{\frac{\mathbf{E}[r_{t+1}^2] - \mu^2}{\mathbf{Var}[r_{t+1}(a)]}}, \qquad \text{(A33)}
\end{aligned}
$$

where the last equality follows from (A32). In particular, applying (A29) and (A30), respectively, gives

$$
D(a, \mathbf{E}[r_{\mathrm{SLOW},t+1}]) = \sqrt{\frac{\mathbf{E}[r_{t+1}^2] - (\mathbf{E}[r_{\mathrm{SLOW},t+1}])^2}{\mathbf{Var}[r_{t+1}(a)]}} = \frac{\mathbf{SD}[r_{\mathrm{SLOW},t+1}]}{\mathbf{SD}[r_{t+1}(a)]}, \qquad \text{(A34)}
$$

$$
D(a, \mathbf{E}[r_{\mathrm{FAST},t+1}]) = \sqrt{\frac{\mathbf{E}[r_{t+1}^2] - (\mathbf{E}[r_{\mathrm{FAST},t+1}])^2}{\mathbf{Var}[r_{t+1}(a)]}} = \frac{\mathbf{SD}[r_{\mathrm{FAST},t+1}]}{\mathbf{SD}[r_{t+1}(a)]}. \qquad \text{(A35)}
$$

Hence, the Sharpe ratio of a momentum strategy of any speed satisfies the following:

$$\begin{aligned}
\mathbf{Sharpe}[r_{t+1}(a)] &= \frac{\mathbf{E}[r_{t+1}(a)]}{\mathbf{SD}[r_{t+1}(a)]} \\
&= \frac{(1-a)\mathbf{E}[r_{\text{SLOW},t+1}] + a\mathbf{E}[r_{\text{FAST},t+1}]}{\mathbf{SD}[r_{t+1}(a)]} \\
&= (1-a)\frac{\mathbf{E}[r_{\text{SLOW},t+1}]}{\mathbf{SD}[r_{\text{SLOW},t+1}]}\frac{\mathbf{SD}[r_{\text{SLOW},t+1}]}{\mathbf{SD}[r_{t+1}(a)]} + a\frac{\mathbf{E}[r_{\text{FAST},t+1}]}{\mathbf{SD}[r_{\text{FAST},t+1}]}\frac{\mathbf{SD}[r_{\text{FAST},t+1}]}{\mathbf{SD}[r_{t+1}(a)]} \\
&= (1-a)\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a, \mathbf{E}[r_{\text{SLOW},t+1}]) \\
&\quad + a\,\mathbf{Sharpe}[r_{\text{FAST},t+1}]D(a, \mathbf{E}[r_{\text{FAST},t+1}]),
\end{aligned} \tag{A36}$$

where the first equality follows by definition of the Sharpe ratio; the second equality follows by (A27); the third equality follows by factoring out each ratio of standard deviations in each summand; and the final equality follows by definition of the Sharpe ratio and equalities (A34) and (A35). The approximation in (18) of the proposition follows immediately by the assumption in the proposition.

Finally, we derive tight bounds for the multiplier $D(a, 0)$ in (19) on $a \in [0, 1]$. First, note that

$$(2a-1)^2 \leq 1, \quad a \in [0, 1], \tag{A37}$$

because $(2a-1)^2$ is a convex function of a, which therefore achieves its maximum value at the boundary of its domain, and is equal to 1 at both of its boundaries. Subtract 1 from both sides of (A37) and then multiply each side of the resultant inequality by the nonnegative quantity $\mathbf{E}[r_{t+1}^2|{}^{\text{Co}}_{\text{Re}}]\mathbf{P}[{}^{\text{Co}}_{\text{Re}}]$ to obtain:

$$[(2a-1)^2 - 1]\mathbf{E}[r_{t+1}^2|{}^{\text{Co}}_{\text{Re}}]\mathbf{P}[{}^{\text{Co}}_{\text{Re}}] \leq 0, \quad a \in [0, 1]. \tag{A38}$$

Add $\mathbf{E}[r_{t+1}^2]$ to both sides of (A38) and use the law of total expectation, $\mathbf{E}[r_{t+1}^2] = \mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + \mathbf{E}[r_{t+1}^2|{}^{\text{Co}}_{\text{Re}}]\mathbf{P}[{}^{\text{Co}}_{\text{Re}}]$, to obtain:

$$\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + (2a-1)^2\mathbf{E}[r_{t+1}^2|{}^{\text{Co}}_{\text{Re}}]\mathbf{P}[{}^{\text{Co}}_{\text{Re}}] \leq \mathbf{E}[r_{t+1}^2], \quad a \in [0, 1]. \tag{A39}$$

Therefore,

$$\begin{aligned}
D(a, 0) &= \sqrt{\frac{\mathbf{E}[r_{t+1}^2]}{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + (2a-1)^2\mathbf{E}[r_{t+1}^2|{}^{\text{Co}}_{\text{Re}}]\mathbf{P}[{}^{\text{Co}}_{\text{Re}}]}} \\
&\geq 1,
\end{aligned} \tag{19}$$

since the numerator inside the square root is always at least as large as the denominator, by (A39). The term $(2a-1)^2\mathbf{E}[r_{t+1}^2|_{\text{Re}}^{\text{Co}}]\mathbf{P}[_{\text{Re}}^{\text{Co}}]$ in the denominator is always nonnegative and equals zero at $a=\frac{1}{2}$, which therefore maximizes $D(a,0)$ on $[0,1]$. If $\mathbf{E}[r_{t+1}^2|_{\text{Re}}^{\text{Co}}]\mathbf{P}[_{\text{Re}}^{\text{Co}}]>0$ then the term is only zero at $a=\frac{1}{2}$, making it the unique maximizer of $D(a,0)$, which completes the proof. □

Proof of Proposition 5. Under the model of Section 2.2, $\mathbf{E}[r_{t+1}(a)]=0$ for all $a\in[0,1]$ because $c=0$. Then, the approximate equation (18) of Proposition 4 becomes exact:

$$\mathbf{Sharpe}[r_{t+1}(a)] = ((1-a)\,\mathbf{Sharpe}[r_{\text{SLOW},t+1}] + a\,\mathbf{Sharpe}[r_{\text{FAST},t+1}])\,D(a). \tag{A40}$$

For $\mathbf{Sharpe}[r_{\text{SLOW},t+1}]=\mathbf{Sharpe}[r_{\text{FAST},t+1}]$, equation (A40) reduces to $\mathbf{Sharpe}[r_{t+1}(a)]=\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a)$, which only depends on a via $D(a)$. By Proposition 4, $D(a)$ is maximized at $a=\frac{1}{2}$. Thus, $\mathbf{Sharpe}[r_{t+1}(a)]$ is maximized at $a=\frac{1}{2}$, which completes the proof. □

Proof of Proposition 6. First, note that

$$\begin{aligned}\mathbf{Cov}[w_t(a),(r_{t+1}-\mathbf{E}[r_{t+1}])^2] &= \mathbf{Cov}[w_t(a), r_{t+1}^2 - 2r_{t+1}\mathbf{E}[r_{t+1}] + (\mathbf{E}[r_{t+1}])^2]\\ &= \mathbf{Cov}[w_t(a), r_{t+1}^2] - 2\mathbf{Cov}[w_t(a), r_{t+1}]\mathbf{E}[r_{t+1}],\end{aligned} \tag{A41}$$

where the first equality follows by squaring the term inside the second covariance argument; and the second equality follows by bilinearity of the covariance operator and the fact that $(\mathbf{E}[r_{t+1}])^2$ is a constant and hence has zero covariance with $w_t(a)$.

Second,

$$\begin{aligned}\mathbf{Cov}[r_{t+1}(a), r_{t+1}] &= \mathbf{Cov}[w_t(a)r_{t+1}, r_{t+1}]\\ &= \mathbf{E}[w_t(a)r_{t+1}^2] - \mathbf{E}[w_t(a)r_{t+1}]\mathbf{E}[r_{t+1}]\\ &= (\mathbf{Cov}[w_t(a), r_{t+1}^2] + \mathbf{E}[w_t(a)]\mathbf{E}[r_{t+1}^2])\\ &\quad - (\mathbf{Cov}[w_t(a), r_{t+1}] + \mathbf{E}[w_t(a)]\mathbf{E}[r_{t+1}])\mathbf{E}[r_{t+1}]\\ &= \mathbf{E}[w_t(a)](\mathbf{E}[r_{t+1}^2] - (\mathbf{E}[r_{t+1}])^2)\\ &\quad - \mathbf{Cov}[w_t(a), r_{t+1}]\mathbf{E}[r_{t+1}] + \mathbf{Cov}[w_t(a), r_{t+1}^2]\\ &= \mathbf{E}[w_t(a)]\mathbf{Var}[r_{t+1}] - \mathbf{Cov}[w_t(a), r_{t+1}]\mathbf{E}[r_{t+1}] + \mathbf{Cov}[w_t(a), r_{t+1}^2]\\ &= \mathbf{E}[w_t(a)]\mathbf{Var}[r_{t+1}] + \mathbf{Cov}[w_t(a), r_{t+1}]\mathbf{E}[r_{t+1}]\\ &\quad + \mathbf{Cov}[w_t(a), (r_{t+1}-\mathbf{E}[r_{t+1}])^2],\end{aligned} \tag{A42}$$

where the first equality follows by definition of $r_{t+1}(a)=w_t(a)r_{t+1}$; the second equality follows by definition of covariance; the third equality follows by two applications of the definition of covariance (once with $\mathbf{E}[w_t(a)r_{t+1}^2] = \mathbf{Cov}[w_t(a), r_{t+1}^2] + \mathbf{E}[w_t(a)]\mathbf{E}[r_{t+1}^2]$ and again with $\mathbf{E}[w_t(a)r_{t+1}] = \mathbf{Cov}[w_t(a), r_{t+1}] - \mathbf{E}[w_t(a)]\mathbf{E}[r_{t+1}]$) similar to the decomposition in (24); the fourth equality follows by rearrangement of terms; the fifth equality follows by definition of variance; and the last equality follows by substitution for $\mathbf{Cov}[w_t(a), r_{t+1}^2]$ from its implicit relation in (A41).

Next, the expression for beta in (26) follows immediately from (A42) (equivalently, (25)) and the definition of beta as the contemporaneous covariance scaled by the underlying market return variance. Finally,

$$\begin{aligned}
&\mathbf{Alpha}[r_{t+1}(a)] \\
&\quad = \mathbf{E}[r_{t+1}(a)] - \mathbf{Beta}[r_{t+1}(a)]\mathbf{E}[r_{t+1}] \\
&\quad = \mathbf{Cov}[w_t(a), r_{t+1}] + \mathbf{E}[w_t(a)]\mathbf{E}[r_{t+1}] \\
&\qquad - \left(\mathbf{E}[w_t(a)] + \frac{\mathbf{Cov}[w_t(a), r_{t+1}]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}] + \frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}\right)\mathbf{E}[r_{t+1}] \\
&\quad = \mathbf{Cov}[w_t(a), r_{t+1}]\left(1 - \frac{(\mathbf{E}[r_{t+1}])^2}{\mathbf{Var}[r_{t+1}]}\right) - \frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}], \qquad \text{(A43)}
\end{aligned}$$

where the first equality follows by definition of alpha; the second equality follows by covariance decomposition of $\mathbf{E}[r_{t+1}(a)]$ as in (24) and substitution of the expression for beta in (26); and the last equality follows by rearrangement. The approximation in (28) follows immediately from the assumption in the proposition, and completes the proof. □

Proof of Lemma 7. The proof proceeds according to the following chain of equalities:

$$\begin{aligned}
\mathbf{Skew}[Y] &= \frac{\mathbf{E}[(Y - \mathbf{E}[Y])^3]}{(\mathbf{SD}[Y])^3} \\
&= \frac{\mathbf{E}[Y^3 - 3Y^2\mathbf{E}[Y] + 3Y\mathbf{E}[Y]^2 - \mathbf{E}[Y]^3]}{(\mathbf{SD}[Y])^3} \\
&= \frac{\mathbf{E}[Y^3]}{(\mathbf{SD}[Y])^3} - 3\frac{\mathbf{E}[Y](\mathbf{E}[Y^2] - \mathbf{E}[Y]^2)}{(\mathbf{SD}[Y])^3} - \frac{\mathbf{E}[Y]^3}{(\mathbf{SD}[Y])^3} \\
&= \frac{\mathbf{E}[Y^3]}{(\mathbf{SD}[Y])^3} - 3\frac{\mathbf{E}[Y]}{\mathbf{SD}[Y]} - \left(\frac{\mathbf{E}[Y]}{\mathbf{SD}[Y]}\right)^3 \\
&= \frac{\mathbf{E}[Y^3]}{(\mathbf{SD}[Y])^3} - \mathbf{Sharpe}[Y]\left(3 + (\mathbf{Sharpe}[Y])^2\right),
\end{aligned}$$

where the first equality is by definition of skewness; the second equality follows by expanding the cube inside the expectation of the numerator; the third equality follows by linearity of the expectation operator; the fourth equality follows by the fact that $(\mathbf{SD}[Y])^2 = \mathbf{Var}[Y] =$

$\mathbf{E}[Y^2] - \mathbf{E}[Y]^2$; and the final equality follows by substitution $\mathbf{Sharpe}[Y] = \frac{\mathbf{E}[Y]}{\mathbf{SD}[Y]}$ and gathering terms, which completes the proof. $\square$

Proof of Proposition 8. First, note that

$$\begin{aligned}
\mathbf{E}[r_{t+1}^3(a)] &= \mathbf{E}[(((1-a)\,w_{\text{SLOW},t} + a\,w_{\text{FAST},t})r_{t+1})^3] \\
&= \mathbf{E}\left[((1-a)^3 w_{\text{SLOW},t}^3 + 3(1-a)^2 a\, w_{\text{SLOW},t}^2\, w_{\text{FAST},t}\right. \\
&\qquad \left. + 3(1-a)a^2 w_{\text{SLOW},t}\, w_{\text{FAST},t}^2 + a^3 w_{\text{FAST},t}^3)\, r_{t+1}^3\right] \\
&= \mathbf{E}\left[((1-a)^3 w_{\text{SLOW},t} + 3(1-a)^2 a\, w_{\text{FAST},t}\right. \\
&\qquad \left. + 3(1-a)a^2 w_{\text{SLOW},t} + a^3 w_{\text{FAST},t})\, r_{t+1}^3\right] \\
&= (1-a)((1-a)^2 + 3a^2)\mathbf{E}[w_{\text{SLOW},t}\, r_{t+1}^3] + a(3(1-a)^2 + a^2)\mathbf{E}[w_{\text{FAST},t}\, r_{t+1}^3] \\
&= (1-a)((1-a)^2 + 3a^2)\frac{\mathbf{E}[(r_{\text{SLOW},t+1})^3]}{(\mathbf{SD}[r_{\text{SLOW},t+1}])^3}(\mathbf{SD}[r_{\text{SLOW},t+1}])^3 \\
&\qquad + a(3(1-a)^2 + a^2)\frac{\mathbf{E}[(r_{\text{FAST},t+1})^3]}{(\mathbf{SD}[r_{\text{FAST},t+1}])^3}(\mathbf{SD}[r_{\text{FAST},t+1}])^3,
\end{aligned} \tag{A44}$$

where the first equality follows by definition of $r_{t+1}(a)$ in (12); the second equality follows by expanding the cube inside the expectation; the third equality follows from the fact that $w_{\text{SLOW},t}^3 = w_{\text{SLOW},t}$, $w_{\text{FAST},t}^3 = w_{\text{FAST},t}$, and $w_{\text{SLOW},t}^2 = w_{\text{FAST},t}^2 = 1$; the fourth equality follows by linearity of the expectation operator and collecting terms; and the last equality follows from the fact that $w_{\text{SLOW},t}^3 = w_{\text{SLOW},t}$ and $w_{\text{FAST},t}^3 = w_{\text{FAST},t}$, by definitions $r_{\text{SLOW},t+1} = w_{\text{SLOW},t}\, r_{t+1}$ and $r_{\text{FAST},t+1} = w_{\text{FAST},t}\, r_{t+1}$, and by multiplying and dividing each summand by the corresponding cubed standard deviation term. Second,

$$\begin{aligned}
\mathbf{Skew}[r_{t+1}(a)] &= \frac{\mathbf{E}[r_{t+1}^3(a)]}{(\mathbf{SD}[r_{t+1}(a)])^3} - \mathbf{Sharpe}[r_{t+1}(a)](3 + (\mathbf{Sharpe}[r_{t+1}(a)])^2) \\
&= (1-a)((1-a)^2 + 3a^2)\frac{\mathbf{E}[(r_{\text{SLOW},t+1})^3]}{(\mathbf{SD}[r_{\text{SLOW},t+1}])^3}\frac{(\mathbf{SD}[r_{\text{SLOW},t+1}])^3}{(\mathbf{SD}[r_{t+1}(a)])^3} \\
&\quad + a(3(1-a)^2 + a^2)\frac{\mathbf{E}[(r_{\text{FAST},t+1})^3]}{(\mathbf{SD}[r_{\text{FAST},t+1}])^3}\frac{(\mathbf{SD}[r_{\text{FAST},t+1}])^3}{(\mathbf{SD}[r_{t+1}(a)])^3} \\
&\quad - \mathbf{Sharpe}[r_{t+1}(a)](3 + (\mathbf{Sharpe}[r_{t+1}(a)])^2) \\
&= (1-a)((1-a)^2 + 3a^2)D^3(a, \mathbf{E}[r_{\text{SLOW},t+1}]) \\
&\qquad \times \left(\mathbf{Skew}[r_{\text{SLOW},t+1}] + \mathbf{Sharpe}[r_{\text{SLOW},t+1}](3 + (\mathbf{Sharpe}[r_{\text{SLOW},t+1}])^2)\right) \\
&\quad + a(3(1-a)^2 + a^2)D^3(a, \mathbf{E}[r_{\text{FAST},t+1}]) \\
&\qquad \times \left(\mathbf{Skew}[r_{\text{FAST},t+1}] + \mathbf{Sharpe}[r_{\text{FAST},t+1}](3 + (\mathbf{Sharpe}[r_{\text{FAST},t+1}])^2)\right) \\
&\quad - 3\,\mathbf{Sharpe}[r_{t+1}(a)] - (\mathbf{Sharpe}[r_{t+1}(a)])^3,
\end{aligned} \tag{A45}$$

where the first equality follows by application of Lemma 7 at $Y = r_{t+1}(a)$; the second equality follows by substitution of (A44) for $\mathbf{E}[r^3_{t+1}(a)]$; and the last equality follows by application of Lemma 7 at $Y = r_{\text{SLOW},t+1}$ and $Y = r_{\text{FAST},t+1}$ and by definition of $D(a,\mu)$ in (17), for which $D(a, \mathbf{E}[r_{\text{SLOW},t+1}]) = \frac{\mathbf{SD}[r_{\text{SLOW},t+1}]}{\mathbf{SD}[r_{t+1}(a)]}$ and $D(a, \mathbf{E}[r_{\text{FAST},t+1}]) = \frac{\mathbf{SD}[r_{\text{FAST},t+1}]}{\mathbf{SD}[r_{t+1}(a)]}$. Continuing from (A45) above and substituting for $\mathbf{Sharpe}[r_{t+1}(a)]$ with (16) of Proposition 4, we have

$$\begin{aligned}
&\mathbf{Skew}[r_{t+1}(a)] \\
&= (1-a)((1-a)^2 + 3a^2)D^3(a, \mathbf{E}[r_{t+1}(0)]) \\
&\quad\quad \times \left(\mathbf{Skew}[r_{\text{SLOW},t+1}] + \mathbf{Sharpe}[r_{\text{SLOW},t+1}](3 + (\mathbf{Sharpe}[r_{\text{SLOW},t+1}])^2)\right) \\
&\quad + a(3(1-a)^2 + a^2)D^3(a, \mathbf{E}[r_{t+1}(1)]) \\
&\quad\quad \times \left(\mathbf{Skew}[r_{\text{FAST},t+1}] + \mathbf{Sharpe}[r_{\text{FAST},t+1}](3 + (\mathbf{Sharpe}[r_{\text{FAST},t+1}])^2)\right) \\
&\quad - 3[(1-a)\,\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)]) + a\,\mathbf{Sharpe}[r_{\text{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)])] \\
&\quad - [(1-a)\,\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)]) + a\,\mathbf{Sharpe}[r_{\text{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)])]^3.
\end{aligned} \tag{A46}$$

Continuing from (A46) above, expanding the cube and grouping terms, we have

$$\begin{aligned}
&\mathbf{Skew}[r_{t+1}(a)] \\
&= (1-a)((1-a)^2 + 3a^2)X^3(a, \mathbf{E}[r_{t+1}(0)])\mathbf{Skew}[r_{\text{SLOW},t+1}] \\
&\quad + a(3(1-a)^2 + a^2)X^3(a, \mathbf{E}[r_{t+1}(1)])\mathbf{Skew}[r_{\text{FAST},t+1}] \\
&\quad + 3[(1-a)\,\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)])\left(((1-a)^2 + 3a^2)D^2(a, \mathbf{E}[r_{t+1}(0)]) - 1\right) \\
&\quad\quad + a\,\mathbf{Sharpe}[r_{\text{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)])\left((3(1-a)^2 + a^2)D^2(a, \mathbf{E}[r_{t+1}(1)]) - 1\right)] \\
&\quad + 3(1-a)a[(1-a)\,\mathbf{Sharpe}[r_{\text{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)]) - a\,\mathbf{Sharpe}[r_{\text{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)])] \\
&\quad\quad \times \left((\mathbf{Sharpe}[r_{\text{FAST},t+1}])^2 D^2(a, \mathbf{E}[r_{t+1}(1)]) - (\mathbf{Sharpe}[r_{\text{SLOW},t+1}])^2 D^2(a, \mathbf{E}[r_{t+1}(0)])\right).
\end{aligned} \tag{A47}$$

Define

$$\gamma(a) := a(a^2 + 3(1-a)^2), \tag{A48}$$

and note that $\gamma(a) \in [0,1]$ with $\gamma(0) = 0$, $\gamma(\frac{1}{2}) = \frac{1}{2}$, and $\gamma(1) = 1$. Also, note that

$1 - \gamma(a) = (1-a)((1-a)^2 + 3a^2)$. Then,

$$\begin{aligned}
&\mathbf{Skew}[r_{t+1}(a)] \\
&= (1-\gamma(a))D^3(a, \mathbf{E}[r_{t+1}(0)])\mathbf{Skew}[r_{\mathrm{SLOW},t+1}] + \gamma(a)D^3(a, \mathbf{E}[r_{t+1}(1)])\mathbf{Skew}[r_{\mathrm{FAST},t+1}] \\
&\quad + 3[(1-a)\,\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)])\left(((1-a)^2 + 3a^2)D^2(a, \mathbf{E}[r_{t+1}(0)]) - 1\right) \\
&\qquad + a\,\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)])\left((3(1-a)^2 + a^2)D^2(a, \mathbf{E}[r_{t+1}(1)]) - 1\right)] \\
&\quad + 3(1-a)a[(1-a)\,\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}]D(a, \mathbf{E}[r_{t+1}(1)]) - a\,\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]D(a, \mathbf{E}[r_{t+1}(0)])] \\
&\qquad \times \left((\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}])^2 D^2(a, \mathbf{E}[r_{t+1}(1)]) - (\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}])^2 D^2(a, \mathbf{E}[r_{t+1}(0)])\right),
\end{aligned} \tag{A49}$$

where $\gamma(a)$ is as defined in (A48).

Approximate $D(a, \mathbf{E}[r_{t+1}(0)]) \approx D(a, \mathbf{E}[r_{t+1}(1)]) \approx D(a, 0)$. Then,

$$\begin{aligned}
\mathbf{Skew}[r_{t+1}(a)] \approx\ & [(1-\gamma(a))\mathbf{Skew}[r_{\mathrm{SLOW},t+1}] + \gamma(a)\mathbf{Skew}[r_{\mathrm{FAST},t+1}]]D^3(a,0) \\
& + 3[(1-a)\,\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]\left(((1-a)^2 + 3a^2)D^2(a,0) - 1\right) \\
& \quad + a\,\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}]\left((3(1-a)^2 + a^2)D^2(a,0) - 1\right)]D(a,0) \\
& + 3(1-a)a[(1-a)\,\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}] - a\,\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]] \\
& \quad \times \left((\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}])^2 - (\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}])^2\right) D^3(a,0).
\end{aligned} \tag{A50}$$

For the special case of $a = \frac{1}{2}$, this approximation becomes:

$$\begin{aligned}
&\mathbf{Skew}[r_{t+1}(\tfrac{1}{2})] \qquad\qquad (A51)\\
&\approx \tfrac{1}{2}(\mathbf{Skew}[r_{\mathrm{SLOW},t+1}] + \mathbf{Skew}[r_{\mathrm{FAST},t+1}])D^3(\tfrac{1}{2};0) \\
&\quad + 3[\tfrac{1}{2}\,\mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}] + \tfrac{1}{2}\,\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}]]D(\tfrac{1}{2};0)\left(D^2(\tfrac{1}{2};0) - 1\right) \\
&\quad + \tfrac{3}{4}\left(\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}] - \mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]\right)^2 \\
&\qquad \times \left(\tfrac{1}{2}(\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}] + \mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}])D(\tfrac{1}{2};0)\right) D^2(\tfrac{1}{2};0) \\
&= \tfrac{1}{2}(\mathbf{Skew}[r_{\mathrm{SLOW},t+1}] + \mathbf{Skew}[r_{\mathrm{FAST},t+1}])D^3(\tfrac{1}{2};0) \\
&\quad + 3\,\mathbf{Sharpe}[r_{t+1}(\tfrac{1}{2})]\left(D^2(\tfrac{1}{2};0) - 1\right) \\
&\quad + \tfrac{3}{4}\left(\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}] - \mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]\right)^2 \mathbf{Sharpe}[r_{t+1}(\tfrac{1}{2})]D^2(\tfrac{1}{2};0) \\
&= \tfrac{1}{2}(\mathbf{Skew}[r_{\mathrm{SLOW},t+1}] + \mathbf{Skew}[r_{\mathrm{FAST},t+1}])D^3(\tfrac{1}{2};0) \\
&\quad + 3\,\mathbf{Sharpe}[r_{t+1}(\tfrac{1}{2})]\left(D^2(\tfrac{1}{2};0)\left[1 + \left(\frac{\mathbf{Sharpe}[r_{\mathrm{FAST},t+1}] - \mathbf{Sharpe}[r_{\mathrm{SLOW},t+1}]}{2}\right)^2\right] - 1\right),
\end{aligned} \tag{A52}$$

where the middle equality follows from substitution of the approximation of $\mathbf{Sharpe}[r_{t+1}(\frac{1}{2})]$ as in (18) of Proposition 4, which holds under the same assumptions as made above. This completes the proof. $\square$

Proof of Proposition 9. First, the expected dynamic return satisfies:

$$\begin{aligned}\mathbf{E}[r_{t+1}(a_{s(t)})] &= \mathbf{E}[((1-a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t})r_{t+1}]\\ &= \mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] + (1-2a_{\text{Co}})\mathbf{E}[r_{t+1}|\text{Co}]\mathbf{P}[\text{Co}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]\\ &\quad + (2a_{\text{Re}}-1)\mathbf{E}[r_{t+1}|\text{Re}]\mathbf{P}[\text{Re}],\end{aligned} \tag{A53}$$

where the first equality follows by definition of $r_{t+1}(a_{s(t)}) = w_t(a_{s(t)})r_{t+1} = ((1-a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t})r_{t+1}$; and the second equality follows from the law of total expectation and the facts that $w_{\text{SLOW},t} = w_{\text{FAST},t} = 1$ after Bull, $w_{\text{SLOW},t} = 1$ and $w_{\text{FAST},t} = -1$ after Correction, $w_{\text{SLOW},t} = w_{\text{FAST},t} = -1$ after Bear, and $w_{\text{SLOW},t} = -1$ and $w_{\text{FAST},t} = 1$ after Rebound.

Second,

$$\begin{aligned}\mathbf{Var}[r_{t+1}(a_{s(t)})] &= \mathbf{Var}[((1-a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t})r_{t+1}]\\ &= \mathbf{E}[((1-a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t})^2 r_{t+1}^2] - (\mathbf{E}[r_{t+1}(a_{s(t)})])^2\\ &= \mathbf{E}[[(1-a_{s(t)})^2 + 2(1-a_{s(t)})a_{s(t)}w_{\text{SLOW},t}w_{\text{FAST},t} + a_{s(t)}^2]r_{t+1}^2]\\ &\quad - (\mathbf{E}[r_{t+1}(a_{s(t)})])^2\\ &= \mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + (2a_{\text{Co}}-1)^2\mathbf{E}[r_{t+1}^2|\text{Co}]\mathbf{P}[\text{Co}]\\ &\quad + (2a_{\text{Re}}-1)^2\mathbf{E}[r_{t+1}^2|\text{Re}]\mathbf{P}[\text{Re}] - (\mathbf{E}[r_{t+1}(a_{s(t)})])^2,\end{aligned} \tag{A54}$$

where the first equality follows by definition of $r_{t+1}(a_{s(t)}) = w_t(a_{s(t)})r_{t+1} = ((1-a_{s(t)})w_{\text{SLOW},t} + a_{s(t)}w_{\text{FAST},t})r_{t+1}$; the second equality follows by definition of variance; the third equality follows by expanding the square inside the expectation of the first summand; and the last equality follows by the law of total expectation and the facts that $w_{\text{SLOW},t}w_{\text{FAST},t} = 1$ after Bull or Bear, and $w_{\text{SLOW},t}w_{\text{FAST},t} = -1$ after Correction or Rebound, plus simplification.

Note that the expected return and variance of the dynamic strategy do not depend on the values of $a_{s(t)}$ in Bull or Bear cycles since both SLOW and FAST agree in these states.

Next, we compute the partial derivatives of the expected return, variance, and standard deviation with respect to each decision variable. Expected return:

$$\frac{\partial}{\partial a_{\text{Co}}}\mathbf{E}[r_{t+1}(a_{s(t)})] = -2\mathbf{E}[r_{t+1}|\text{Co}]\mathbf{P}[\text{Co}], \tag{A55}$$

$$\frac{\partial}{\partial a_{\text{Re}}}\mathbf{E}[r_{t+1}(a_{s(t)})] = 2\mathbf{E}[r_{t+1}|\text{Re}]\mathbf{P}[\text{Re}], \tag{A56}$$

Variance:

$$\begin{aligned}
\frac{\partial}{\partial a_{\text{Co}}}\mathbf{Var}[r_{t+1}(a_{s(t)})] &= 4(2a_{\text{Co}}-1)\mathbf{E}[r_{t+1}^2|\text{Co}]\mathbf{P}[\text{Co}] - 2\mathbf{E}[r_{t+1}(a_{s(t)})]\frac{\partial}{\partial a_{\text{Co}}}\mathbf{E}[r_{t+1}(a_{s(t)})] \\
&= 4(2a_{\text{Co}}-1)\mathbf{E}[r_{t+1}^2|\text{Co}]\mathbf{P}[\text{Co}] - 2\mathbf{E}[r_{t+1}(a_{s(t)})](-2\mathbf{E}[r_{t+1}|\text{Co}]\mathbf{P}[\text{Co}]) \\
&= 4\mathbf{P}[\text{Co}]\{(2a_{\text{Co}}-1)\mathbf{E}[r_{t+1}^2|\text{Co}] + \mathbf{E}[r_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\text{Co}]\}, \qquad \text{(A57)} \\
\frac{\partial}{\partial a_{\text{Re}}}\mathbf{Var}[r_{t+1}(a_{s(t)})] &= 4(2a_{\text{Re}}-1)\mathbf{E}[r_{t+1}^2|\text{Re}]\mathbf{P}[\text{Re}] - 2\mathbf{E}[r_{t+1}(a_{s(t)})]\frac{\partial}{\partial a_{\text{Re}}}\mathbf{E}[r_{t+1}(a_{s(t)})] \\
&= 4(2a_{\text{Re}}-1)\mathbf{E}[r_{t+1}^2|\text{Re}]\mathbf{P}[\text{Re}] - 2\mathbf{E}[r_{t+1}(a_{s(t)})](2\mathbf{E}[r_{t+1}|\text{Re}]\mathbf{P}[\text{Re}]) \\
&= 4\mathbf{P}[\text{Re}]\{(2a_{\text{Re}}-1)\mathbf{E}[r_{t+1}^2|\text{Re}] - \mathbf{E}[r_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\text{Re}]\}, \qquad \text{(A58)}
\end{aligned}$$

where the second equality in the equality chain of (A57) and (A58) follow from (A55) and (A56), respectively. Standard deviation:

$$\begin{aligned}
\frac{\partial}{\partial a_{\text{Co}}}\mathbf{SD}[r_{t+1}(a_{s(t)})] &= \frac{1}{2}\frac{1}{\mathbf{SD}[r_{t+1}(a_{s(t)})]}\frac{\partial}{\partial a_{\text{Co}}}\mathbf{Var}[r_{t+1}(a_{s(t)})] \\
&= \frac{2\mathbf{P}[\text{Co}]\{(2a_{\text{Co}}-1)\mathbf{E}[r_{t+1}^2|\text{Co}] + \mathbf{E}[r_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\text{Co}]\}}{\mathbf{SD}[r_{t+1}(a_{s(t)})]}, \qquad \text{(A59)} \\
\frac{\partial}{\partial a_{\text{Re}}}\mathbf{SD}[r_{t+1}(a_{s(t)})] &= \frac{1}{2}\frac{1}{\mathbf{SD}[r_{t+1}(a_{s(t)})]}\frac{\partial}{\partial a_{\text{Re}}}\mathbf{Var}[r_{t+1}(a_{s(t)})] \\
&= \frac{2\mathbf{P}[\text{Re}]\{(2a_{\text{Re}}-1)\mathbf{E}[r_{t+1}^2|\text{Re}] - \mathbf{E}[r_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\text{Re}]\}}{\mathbf{SD}[r_{t+1}(a_{s(t)})]}, \qquad \text{(A60)}
\end{aligned}$$

where (A59) and (A60) follow from (A57) and (A58), respectively.

Next, we compute the partial derivatives of the Sharpe ratio with respect to each decision variable. Sharpe ratio:

$$\begin{aligned}
\frac{\partial}{\partial a_{\text{Co}}}\mathbf{Sharpe}[r_{t+1}(a)] &= \frac{\partial}{\partial a_{\text{Co}}}\frac{\mathbf{E}[r_{t+1}(a_{s(t)})]}{\mathbf{SD}[r_{t+1}(a_{s(t)})]} \\
&= \frac{\mathbf{SD}[r_{t+1}(a_{s(t)})]\frac{\partial}{\partial a_{\text{Co}}}\mathbf{E}[r_{t+1}(a_{s(t)})] - \mathbf{E}[r_{t+1}(a_{s(t)})]\frac{\partial}{\partial a_{\text{Co}}}\mathbf{SD}[r_{t+1}(a_{s(t)})]}{\mathbf{Var}[r_{t+1}(a_{s(t)})]} \\
&= \frac{\mathbf{SD}[r_{t+1}(a_{s(t)})](-2\mathbf{E}[r_{t+1}|\text{Co}]\mathbf{P}[\text{Co}])}{\mathbf{Var}[r_{t+1}(a_{s(t)})]} \\
&\quad - \frac{\mathbf{E}[r_{t+1}(a_{s(t)})]\left(\frac{2\mathbf{P}[\text{Co}]\{(2a_{\text{Co}}-1)\mathbf{E}[r_{t+1}^2|\text{Co}]+\mathbf{E}[r_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\text{Co}]\}}{\mathbf{SD}[r_{t+1}(a_{s(t)})]}\right)}{\mathbf{Var}[r_{t+1}(a_{s(t)})]}
\end{aligned}$$

$$= -2\mathbf{P}[\mathrm{Co}]\left[\frac{(\mathbf{Var}[r_{t+1}(a_{s(t)})] + (\mathbf{E}[r_{t+1}(a_{s(t)})])^2)\mathbf{E}[r_{t+1}|\mathrm{Co}]}{\mathbf{SD}[r_{t+1}(a_{s(t)})]\mathbf{Var}[r_{t+1}(a_{s(t)})]} + \frac{(2a_{\mathrm{Co}} - 1)\mathbf{E}[r^2_{t+1}|\mathrm{Co}]\mathbf{E}[r_{t+1}(a_{s(t)})]}{\mathbf{SD}[r_{t+1}(a_{s(t)})]\mathbf{Var}[r_{t+1}(a_{s(t)})]}\right]$$
$$= -2\mathbf{P}[\mathrm{Co}]\frac{\mathbf{E}[r^2_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\mathrm{Co}] + (2a_{\mathrm{Co}} - 1)\mathbf{E}[r^2_{t+1}|\mathrm{Co}]\mathbf{E}[r_{t+1}(a_{s(t)})]}{\mathbf{SD}[r_{t+1}(a_{s(t)})]\mathbf{Var}[r_{t+1}(a_{s(t)})]}, \tag{A61}$$

where the second equality follows by the quotient rule for derivatives; the third equality follows from (A55) and (A59); and the last equality follows from the definition of variance. Similar steps for a_{Re} yield the following:

$$\frac{\partial}{\partial a_{\mathrm{Re}}}\mathbf{Sharpe}[r_{t+1}(a)] = 2\mathbf{P}[\mathrm{Re}]\frac{\mathbf{E}[r^2_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\mathrm{Re}] - (2a_{\mathrm{Re}} - 1)\mathbf{E}[r^2_{t+1}|\mathrm{Re}]\mathbf{E}[r_{t+1}(a_{s(t)})]}{\mathbf{SD}[r_{t+1}(a_{s(t)})]\mathbf{Var}[r_{t+1}(a_{s(t)})]}. \tag{A62}$$

Next, we apply the first order conditions to derive stationary points of the maximization problem. The first order conditions are: $\frac{\partial}{\partial a_{\mathrm{Co}}}\mathbf{Sharpe}[r_{t+1}(a)] = 0$ and $\frac{\partial}{\partial a_{\mathrm{Re}}}\mathbf{Sharpe}[r_{t+1}(a)] = 0$, or equivalently, by (A61) and (A62),

$$\mathbf{E}[r^2_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\mathrm{Co}] + (2a_{\mathrm{Co}} - 1)\mathbf{E}[r^2_{t+1}|\mathrm{Co}]\mathbf{E}[r_{t+1}(a_{s(t)})] = 0, \tag{A63}$$
$$\mathbf{E}[r^2_{t+1}(a_{s(t)})]\mathbf{E}[r_{t+1}|\mathrm{Re}] - (2a_{\mathrm{Re}} - 1)\mathbf{E}[r^2_{t+1}|\mathrm{Re}]\mathbf{E}[r_{t+1}(a_{s(t)})] = 0. \tag{A64}$$

Dividing each condition (A63) and (A64) through by $\mathbf{E}[r^2_{t+1}|\mathrm{Co}]$ and $\mathbf{E}[r^2_{t+1}|\mathrm{Re}]$, respectively, and rearranging yields the equivalent conditions:

$$(2a_{\mathrm{Co}} - 1)\mathbf{E}[r_{t+1}(a_{s(t)})] = -\frac{\mathbf{E}[r_{t+1}|\mathrm{Co}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Co}]}\mathbf{E}[r^2_{t+1}(a_{s(t)})], \tag{A65}$$
$$(2a_{\mathrm{Re}} - 1)\mathbf{E}[r_{t+1}(a_{s(t)})] = \frac{\mathbf{E}[r_{t+1}|\mathrm{Re}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Re}]}\mathbf{E}[r^2_{t+1}(a_{s(t)})]. \tag{A66}$$

By (A53) and (A54), and the fact that $\mathbf{E}[r^2_{t+1}(a_{s(t)})] = \mathbf{Var}[r^2_{t+1}(a_{s(t)})] + (\mathbf{E}[r_{t+1}(a_{s(t)})])^2$, these conditions are equivalent to $(2a_{\mathrm{Co}} - 1)(\mathbf{E}[r_{t+1}|\mathrm{Bu}]\mathbf{P}[\mathrm{Bu}] - \mathbf{E}[r_{t+1}|\mathrm{Be}]\mathbf{P}[\mathrm{Be}] + (2a_{\mathrm{Re}} - 1)\mathbf{E}[r_{t+1}|\mathrm{Re}]\mathbf{P}[\mathrm{Re}]) = -\frac{\mathbf{E}[r_{t+1}|\mathrm{Co}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Co}]}(\mathbf{E}[r^2_{t+1}|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}] + (2a_{\mathrm{Re}} - 1)^2\mathbf{E}[r^2_{t+1}|\mathrm{Re}]\mathbf{P}[\mathrm{Re}])$ and $(2a_{\mathrm{Re}} - 1)(\mathbf{E}[r_{t+1}|\mathrm{Bu}]\mathbf{P}[\mathrm{Bu}] + (1 - 2a_{\mathrm{Co}})\mathbf{E}[r_{t+1}|\mathrm{Co}]\mathbf{P}[\mathrm{Co}] - \mathbf{E}[r_{t+1}|\mathrm{Be}]\mathbf{P}[\mathrm{Be}]) = \frac{\mathbf{E}[r_{t+1}|\mathrm{Re}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Re}]}(\mathbf{E}[r^2_{t+1}|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}] + (2a_{\mathrm{Co}} - 1)^2\mathbf{E}[r^2_{t+1}|\mathrm{Co}]\mathbf{P}[\mathrm{Co}])$, so we can isolate the decision variables on either side of the condition equations: $(2a_{\mathrm{Co}} - 1) = -\frac{\mathbf{E}[r_{t+1}|\mathrm{Co}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Co}]}\frac{\mathbf{E}[r^2_{t+1}|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}]+(2a_{\mathrm{Re}}-1)^2\mathbf{E}[r^2_{t+1}|\mathrm{Re}]\mathbf{P}[\mathrm{Re}]}{\mathbf{E}[r_{t+1}|\mathrm{Bu}]\mathbf{P}[\mathrm{Bu}]-\mathbf{E}[r_{t+1}|\mathrm{Be}]\mathbf{P}[\mathrm{Be}]+(2a_{\mathrm{Re}}-1)\mathbf{E}[r_{t+1}|\mathrm{Re}]\mathbf{P}[\mathrm{Re}]}$ and $(2a_{\mathrm{Re}}-1) = \frac{\mathbf{E}[r_{t+1}|\mathrm{Re}]}{\mathbf{E}[r^2_{t+1}|\mathrm{Re}]}\frac{\mathbf{E}[r^2_{t+1}|{}^{\mathrm{Bu}}_{\mathrm{Be}}]\mathbf{P}[{}^{\mathrm{Bu}}_{\mathrm{Be}}]+(2a_{\mathrm{Co}}-1)^2\mathbf{E}[r^2_{t+1}|\mathrm{Co}]\mathbf{P}[\mathrm{Co}])}{\mathbf{E}[r_{t+1}|\mathrm{Bu}]\mathbf{P}[\mathrm{Bu}]+(1-2a_{\mathrm{Co}})\mathbf{E}[r_{t+1}|\mathrm{Co}]\mathbf{P}[\mathrm{Co}]-\mathbf{E}[r_{t+1}|\mathrm{Be}]\mathbf{P}[\mathrm{Be}]}$. Next, rewrite these conditions as

$$x(y) = -\frac{\mathbf{E}[r_{t+1}|\text{Co}]}{\mathbf{E}[r_{t+1}^2|\text{Co}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + y^2\mathbf{E}[r_{t+1}^2|\text{Re}]\mathbf{P}[\text{Re}]}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}] + y\mathbf{E}[r_{t+1}|\text{Re}]\mathbf{P}[\text{Re}]}, \tag{A67}$$

$$y(x) = \frac{\mathbf{E}[r_{t+1}|\text{Re}]}{\mathbf{E}[r_{t+1}^2|\text{Re}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}] + x^2\mathbf{E}[r_{t+1}^2|\text{Co}]\mathbf{P}[\text{Co}])}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] - x\mathbf{E}[r_{t+1}|\text{Co}]\mathbf{P}[\text{Co}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]}. \tag{A68}$$

To find a fixed point (x^*, y^*) of the above system such that $x^* = x(y^*)$ and $y^* = y(x^*)$, note that $x(0) = x(y(0))$ and $y(0) = y(x(0))$. Hence, $x^* = x(y(0))$ and $y^* = y(x(0))$ are fixed points. Specifically, we have $x(0) = -\frac{\mathbf{E}[r_{t+1}|\text{Co}]}{\mathbf{E}[r_{t+1}^2|\text{Co}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}]}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}]-\mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]}$ and $y(0) = \frac{\mathbf{E}[r_{t+1}|\text{Re}]}{\mathbf{E}[r_{t+1}^2|\text{Re}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}]}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}]-\mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]}$. Therefore,

$$a^*_{\text{Co}} = \frac{1}{2}\left(1 - \frac{\mathbf{E}[r_{t+1}|\text{Co}]}{\mathbf{E}[r_{t+1}^2|\text{Co}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}]}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]}\right), \tag{A69}$$

$$a^*_{\text{Re}} = \frac{1}{2}\left(1 + \frac{\mathbf{E}[r_{t+1}|\text{Re}]}{\mathbf{E}[r_{t+1}^2|\text{Re}]}\frac{\mathbf{E}[r_{t+1}^2|{}^{\text{Bu}}_{\text{Be}}]\mathbf{P}[{}^{\text{Bu}}_{\text{Be}}]}{\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}]}\right), \tag{A70}$$

comprise a stationary point of the Sharpe ratio maximization problem.

Moreover, the necessary stationary conditions (A65) and (A66) imply that $(2a_{\text{Co}} - 1)$ and $(2a_{\text{Re}} - 1)$ must be proportional to $-\frac{\mathbf{E}[r_{t+1}|\text{Co}]}{\mathbf{E}[r_{t+1}^2|\text{Co}]}$ and $\frac{\mathbf{E}[r_{t+1}|\text{Re}]}{\mathbf{E}[r_{t+1}^2|\text{Re}]}$, respectively, with the same proportionality constant, $\frac{\mathbf{E}[r_{t+1}^2(a_{s(t)})]}{\mathbf{E}[r_{t+1}(a_{s(t)})]}$. Since the Sharpe ratio is bounded above and below, and since this proportionality constant can be either positive or negative—depending on the sign of its denominator—the stationary point defined by (A69) and (A70), and the negative version of this point, are the only two stationary points. If $\mathbf{E}[r_{t+1}|\text{Bu}]\mathbf{P}[\text{Bu}] - \mathbf{E}[r_{t+1}|\text{Be}]\mathbf{P}[\text{Be}] > 0$, then expected return at stationary point (A69) and (A70) is positive, and this point is the unique maximizer, while the negative version of this point is the unique minimizer. This completes the proof. $\square$

Internet Appendix B. Data Sources

U.S. stock market returns of Sections 1-5 are the excess-weighted factor returns (Mkt-RF) from the Kenneth French Data Library. The data coverage is from 1926-07 to 2018-12. Section 6 makes use of local excess equity market returns for 20 countries. The data source is Datastream with coverage from 1980-01 to 2018-12. Some countries have later start dates in the sample: AUT, CHE, DNK, HKG, IRL, NLD, and SGP begin 1980-02; ESP begins 1987-04; FIN begins 1988-04; JPN begins 1986-08; NOR begins 1986-02; and SWE begins 1982-02.

Section 3 studies 16 macro time-series. From the FRED database of the Federal Reserve Bank of St. Louis, we obtained five indexes compiled by the Chicago Federal Reserve: CFNAI Personal Consumption and Housing (1967-03 to 2019-03); CFNAI Production and Income (1967-03 to 2019-03); CFNAI Sales, Orders and Inventories (1967-03 to 2019-03); CFNAI Employment, Unemployment and Hours (1967-03 to 2019-03); and the Chicago Fed National Financial Conditions Index (1967-03 to 2019-03). Also, from the FRED database we obtained the TED spread (1986-01 to 2019-03), the Consumer Sentiment Index of the University of Michigan (1978-01 to 2019-03), and NBER recessions-indicator time series. The series of monetary policy shocks are estimated as in Gertler and Karadi (2015) and the data coverage is from 1990-01 to 2012-06. The liquidity innovation measure is from Pastor and Stambaugh (2003) (from 1962-08 to 2018-12), the bond illiquidity metric (Noise) is from Hu, Pan, and Wang (2013) (from 1987-01 to 2016-11), and the high-volatility/low volatility spread is from Pflueger, Siriwardane, and Sunderam (2018) (quarterly data from 1970-06 to 2016-06). The U.S. Policy measure is from Baker, Bloom, and Davis (2016) and covers the period from 1985-01 to 2019-03. Lastly, the PMI ISM series is from Bloomberg (1948-01 to 2019-03) and from the Survey of Professional forecasts we obtained recession expectations (QTR1, median response) and corporate profit expectations (DCPROF2, % change of median response); both are from 1968-12 to 2019-03 at a quarterly frequency.

Internet Appendix C. Comparison with TS Momentum of Moskowitz et al. (2012)

Established time series (TS) momentum strategies such as those of Moskowitz et al. (2012) have been shown to generate significant alpha with respect to market returns. What do our intermediate-speed TS momentum strategies offer in addition? In this section, we compare the performance of the MED strategy from Section 4 (Table 2) with the performance of the TS momentum strategy established by Moskowitz et al. (2012) applied to U.S. market returns, which we label "MOP".

MOP explicitly embeds volatility targeting. Specifically, excess returns on MOP are defined by $r_{t+1}^{\mathrm{MOP}} = \mathrm{sign}(r_{t-12,t})\frac{15\%}{\sigma_t}r_{t+1}$, where $r_{t-12,t}$ is the trailing 12-month excess geometric return on U.S. value-weighted factor, the scalar 15% is the volatility target, r_{t+1} is the one-month U.S. excess value-weighted factor return, and σ_t is ex-ante annualized volatility defined as follows. The ex-ante annualized variance is $\sigma_t^2 = 261\sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - \overline{r}_t)^2$, where the scalar 261 scales the variance to be annual, the weights $(1-\delta)\delta^i$ add up to one, and $\overline{r}_t$ is the exponentially weighted average return computed similarly. The parameter δ is chosen so that the center of mass of the weights is $\sum_{i=0}^{\infty}(1-\delta)\delta^i i = \frac{\delta}{1-\delta} = 60$ days. The 15% annual volatility target is chosen because it is similar to the volatility of the market return.

Volatility targeting introduces a separate mechanism from TS momentum (Kim et al. (2016), Moreira and Muir (2017), Harvey et al. (2018)). Note that MED is not scaled by trailing volatility as in MOP. Rather, MED blends the signs of the trailing 12-month and 1-month arithmetic average returns, which embeds volatility timing effects intrinsically (as discussed in Section 4.2).

Table C.1 reports that MED generates significant alpha with respect MOP over the 50-year evaluation period from 1969 through 2018 used in our main analyses.[37] This evidence suggests that our intermediate-speed TS momentum strategies can offer a distinct source of alpha relative to strategies that explicitly embed volatility targeting. See Section 4.2.2 for additional discussion on relationships with volatility management.

[37]We obtain similar results (unreported) over the longer sample using data back to 1926.

Table C.1: **Performance Summary: MOP vs. MED**

	MOP	$a = \frac{1}{2}$ MED
Return and Risk		
Average (%) (anlzd.)	6.04	5.88
Volatility (%) (anlzd.)	16.83	11.60
Sharpe Ratio (anlzd.)	0.36	0.51
Alpha to MOP		
Alpha (%) (anlzd.)	0.00	3.14
Alpha t-statistic	—	2.53

Notes: This table reports the sample average, volatility, Sharpe ratio, as well as alpha and alpha t-statistic with respect to the TS momentum strategy of Moskowitz et al. (2012) (MOP). Excess returns on MOP are defined as $r_{t+1}^{\text{MOP}} = \text{sign}(r_{t-12,t}) \frac{15\%}{\sigma_t} r_{t+1}$, where $r_{t-12,t}$ is the trailing 12-month excess geometric return on U.S. value-weighted factor, the scalar 15% is the volatility target, r_{t+1} is the one-month U.S. excess value-weighted factor return, and σ_t is ex-ante annualized volatility defined as follows. The ex-ante annualized variance is $\sigma_t^2 = 261 \sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i} - \overline{r}_t)^2$, where the scalar 261 scales the variance to be annual, the weights $(1-\delta)\delta^i$ add up to one, and $\overline{r}_t$ is the exponentially weighted average return computed similarly. The parameter δ is chosen so that the center of mass of the weights is $\sum_{i=0}^{\infty}(1-\delta)\delta^i i = \frac{\delta}{1-\delta} = 60$ days. The 15% annual volatility target is chosen because it is similar to the volatility of the market return itself.

MED strategy weights are formed by mixing slow and fast strategies with mixing parameter $a = \frac{1}{2}$: $w_t(a) = (1-a)w_{\text{SLOW},t} + a w_{\text{FAST},t}$. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. MED excess returns are formed as $r_{t+1}(a) = w_t(a) r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return.

Data sources: Kenneth French Data Library. The evaluation period is 1969-01 to 2018-12.

Internet Appendix D. Exponentially Weighted Moving Averages

We analyze momentum strategies based on exponentially weighted moving averages (EWMAs) of past returns. We show that realized returns under our model can be recast equivalently as observations from an autoregressive moving average process with first-order autoregressive and first-order moving average components, ARMA(1,1). We use this equivalence to show analytically that the conditional expected return given all past realized returns is an exponentially weighted moving average of all past returns with the decay factor proportional to the persistence of the underlying expected return process.

We analyze two applications of this result for U.S. stock market returns over our 50-year sample. First, we consider a family of binary-position momentum strategies in which each strategy's decision to go long or short one unit in the underlying excess return process is based on the sign of truncated EWMAs of past returns. This family of strategies nests our SLOW and FAST: when the exponential decay factor goes to zero, we obtain FAST (full decay after one month), and when the decay factor goes to one (with truncation at the 12 past months of returns), we obtain SLOW. When the decay factor is close to the persistence of the underlying expected return process and truncation is expanded to include all past observations, we obtain a strategy that is long or short one unit based on the sign of conditional expected returns estimated under the model.

We find that the strategy with the decay factor just below the estimated persistence of expected returns ($\widehat{\varphi} = 0.85$) and truncation at the 12 past months performs best among all strategies in this family over a grid of possible parameter values. In particular, this strategy performs better than both SLOW and FAST. This result reinforces our conclusion that the underlying expected return process for the U.S. stock market is highly persistent and that longer-horizon averages (12 months or more) help to moderate the high noise content of realized returns. However, this best-performing binary EWMA-based strategy does not outperform MED, our static equal-blend of SLOW and FAST. MED exhibits statistically significant alpha with respect to this strategy and a higher Sharpe ratio. This result suggests there is additional information in the union of SLOW and FAST that the mixture strategy MED is able to capture compared to strategies that take binary positions based on the signs of EWMAs of trailing returns.

Second, we consider a family of continuous-position momentum strategies in which each strategy takes a position linear in the truncated EWMA of past returns. The motivation for this family of strategies is that strategies with positions of the binary variety, as in our first application, may constrain performance compared to strategies with weights continuously varying as a function of conditional expected returns. Because the EWMA proxies for conditional expected returns implied by our model, a strategy in this family varies its risk exposure in proportion to the implied risk compensation of the market. Rather than going fully long or short one unit, we invest in proportion to the magnitude of the truncated EWMA so that we take on more long-position investment risk in periods of steeper expected uptrend and more short-position investment risk in periods of steeper expected downtrend. We find that the best-performing strategy among this continuous-position family, over a grid of feasible decay factors, achieves a Sharpe ratio above that of the best binary-position strategy. Again, this result is consistent with the intuition that continuously varying weights can potentially improve performance over simpler binary strategies. This best-performing EWMA-based strategy, however, does not outperform MED. MED exhibits statistically significant alpha with respect to this EWMA strategy and a higher Sharpe ratio. This result suggests there is additional information in the union of SLOW and FAST that the mixture strategy MED is able to capture compared to this framework of continuous positions.

Other strategy types that take positions based on a nonlinear function of conditional expected returns could potentially perform better. We leave the topic of general nonlinear families of strategies to future work.

Internet Appendix D.A. Recasting as ARMA(1,1)

Recall the model of Section 2.2:

$$\mu_t = c + \varphi \mu_{t-1} + \varepsilon_t, \tag{3}$$

$$r_t = \mu_t + z_t, \tag{4}$$

$$\begin{pmatrix} \varepsilon_t \\ z_t \end{pmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon^2 & \lambda\sigma_\varepsilon\sigma_z \\ \lambda\sigma_\varepsilon\sigma_z & \sigma_z^2 \end{bmatrix} \right), \tag{D71}$$

where $\varphi \in (0,1)$, $\sigma_\varepsilon > 0$, $\sigma_z > 0$, $\lambda \in (-1,1)$, and where we define $\theta = \frac{\sigma_z^2 + \lambda\sigma_\varepsilon\sigma_z}{\sigma_r^2}$ and $\sigma_r^2 = \frac{\sigma_\varepsilon^2}{1-\varphi^2} + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z$. We can recast this model as the following equivalent ARMA(1,1):

$$r_t = c + \varphi r_{t-1} + w_t - \delta w_{t-1}, \tag{D72}$$

$$w_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_w^2), \tag{D73}$$

where

$$\delta = \varphi \frac{\sigma_z}{\sigma_w}, \tag{D74}$$

$$\sigma_w^2 = \sigma_\varepsilon^2 + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z. \tag{D75}$$

To see this, note the following:

$$\begin{aligned} r_t - \varphi r_{t-1} &= (\mu_t + z_t) - \varphi(\mu_{t-1} + z_{t-1}) = (c + \varphi\mu_{t-1} + \varepsilon_t + z_t) - \varphi\mu_{t-1} - \varphi z_{t-1} \\ &= c + \varepsilon_t + z_t - \varphi z_{t-1}, \end{aligned} \tag{D76}$$

where the first equality follows by substitution of (4) at dates t and $t-1$; the second equality follows by substitution of (3) for μ_t; and the final equality follows by simplification. Let $w_t = \varepsilon_t + z_t$. Then, $\mathbf{Var}[w_t] = \mathbf{Var}[\varepsilon_t + z_t] = \sigma_\varepsilon^2 + \sigma_z^2 + 2\lambda\sigma_\varepsilon\sigma_z = \sigma_w^2$. Now scale w_{t-1} by $\delta > 0$ so that δw_{t-1} has the same distribution as φz_{t-1}: $\mathbf{Var}[\delta w_{t-1}] = \delta^2\sigma_w^2 = \left(\varphi\frac{\sigma_z}{\sigma_w}\right)^2 \sigma_w^2 = \varphi^2\sigma_z^2 = \mathbf{Var}[\varphi z_{t-1}]$. Moreover, w_t is independent of w_{t-1}, because (ε_t, z_t) is independent of $(\varepsilon_{t-1}, z_{t-1})$. This fact plus substitution of w_t and w_{t-1} into (D76) and rearrangement of terms yields (D72) and (D73).

We can express (D72) equivalently as $(1-\varphi L)r_t = c + (1-\delta L)w_t$, where L is the lag operator ($Lr_t = r_{t-1}$ and $Lw_t = w_{t-1}$). For $\delta \in (0,1)$, we can invert the operator on the right-hand side to yield $w_t = (1-\delta L)^{-1}\left((1-\varphi L)r_t - c\right) = (1 + \delta L + \delta^2 L^2 + \cdots)(1-\varphi L)r_t - \frac{c}{1-\delta}$, which expands to

$$w_t = r_t - (\varphi - \delta)\sum_{i=0}^{\infty} \delta^i r_{t-1-i} - \frac{c}{1-\delta}. \tag{D77}$$

Likewise, because $\varphi \in (0,1)$, we can invert the operator on the left-hand side to yield $r_t = (1-\varphi L)^{-1}(c + (1-\delta L)w_t) = \frac{c}{1-\varphi} + (1 + \varphi L + \varphi^2 L^2 + \cdots)(1-\delta L)w_t = \frac{c}{1-\varphi} + w_t + (\varphi - \delta)\sum_{i=0}^{\infty} \varphi^i w_{t-1-i}$. Therefore, the information in the set of all shocks to time t, $\{w_s\}_{s\leq t}$, is equivalent to the information in the set of all returns to time t, $\{r_s\}_{s\leq t}$.

Thus, if $\delta \in (0,1)$, the conditional expected value of r_t given the history of all $\{r_s\}_{s<t}$ is

$$\begin{aligned} \mathbf{E}[r_t | r_{t-1}, r_{t-2}, \ldots] &= \mathbf{E}[c + \varphi r_{t-1} + w_t - \delta w_{t-1} | r_{t-1}, r_{t-2}, \ldots] \\ &= c + \varphi r_{t-1} - \delta w_{t-1} \\ &= c + \varphi r_{t-1} - \delta \left(r_{t-1} - (\varphi - \delta)\sum_{i=0}^{\infty} \delta^i r_{t-2-i} - \frac{c}{1-\delta} \right) \\ &= \frac{c}{1-\delta} + (\varphi - \delta)\left(r_{t-1} + \delta r_{t-2} + \delta^2 r_{t-3} + \cdots\right), \end{aligned} \tag{D78}$$

where the first equality follows by substitution of (D72) for r_t; the second equality follows because r_{t-1} and w_{t-1} are measurable with respect to the set of all returns to time $t-1$, because w_t is independent of $\{w_s\}_{s<t}$, which is equivalent to $\{r_s\}_{s<t}$, and $\mathbf{E}[w_t] = 0$; the third equality follows by substitution of (D77) at date $t-1$ for w_{t-1}; and the final equality follows by rearrangement and simplification.

In order for this conditional expectation to be classified as a momentum signal, the term $(\varphi - \delta)$ multiplying the EWMA of past returns needs to be positive. Therefore, we impose the condition $\delta < \varphi$. This condition implies that $\frac{\sigma_z}{\sigma_w} < 1$, or equivalently, that the ratio $\frac{\mathbf{Var}[z_t]}{\mathbf{Var}[z_t+\varepsilon_t]}$ is below one. Note that if the noise in realized returns (i.e., variance of z_t) is much larger than that in expected returns (i.e., variance of ε_t), then this ratio is close to one. For example, suppose that z_t and ε_t are uncorrelated. Then, $\frac{\sigma_z}{\sigma_w} = \sqrt{\frac{\theta}{(1-\theta)(1-\varphi^2)+\theta}}$. Substitution of our estimates $(\widehat{\varphi}, \widehat{\theta}) = (0.85, 0.86)$ from Section 4.1.3 into this expression gives a value of about 0.98 for this ratio.

Inspired by (D78), we define a signal family based on the truncated EWMA of past returns observed as of date t as

$$\mathrm{EWMA}_t(\delta, K) = \frac{c}{1-\delta} + (\varphi - \delta)(r_t + \delta r_{t-1} + \delta^2 r_{t-2} + \cdots + \delta^{K-1} r_{t-K-1}), \tag{D79}$$

where $0 \le \delta < \varphi < 1$ and $K = 1, 2, \ldots$. This signal can be used to take strategy positions in returns to be realized at date $t+1$.

Internet Appendix D.B. Binary EWMA Strategies

We consider the collection of binary strategies indexed by (δ, K) such that each strategy goes long one unit in the underlying excess return if $\mathrm{EWMA}_t(\delta, K) \ge 0$ and otherwise goes short one unit:

$$w_t(\delta, K) := \begin{cases} +1, & \text{if } \mathrm{EWMA}_t(\delta, K) \ge 0, \\ -1, & \text{if } \mathrm{EWMA}_t(\delta, K) < 0. \end{cases}$$

Intuitively, we can view $\mathrm{EWMA}_t(\delta, K)$ as an estimate of conditional expected returns at date t, assigning higher weight to more recent returns and lower weight to less recent returns. Then, the condition $\mathrm{EWMA}_t(\delta, K) \ge 0$ indicates periods of uptrend in which we desire to go long, while the condition $\mathrm{EWMA}_t(\delta, K) < 0$ indicates periods of downtrend in which we desire to go short. We set $c = 0$ so that this collection nests both SLOW and FAST, where $\mathrm{EWMA}_t(\delta = 1, K = 12)$ corresponds to SLOW of (1) and $\mathrm{EWMA}_t(\delta = 0, K = 1)$ corresponds to FAST of (2). Moreover, for $c = 0$, we can drop the multiplier $(\varphi - \delta)$ because it is positive and will not affect the sign of the weighted average of past returns.

Thus, this collection of strategies invests according to the sign of weighted past returns: $r_t + \delta r_{t-1} + \delta^2 r_{t-2} + \cdots + \delta^{K-1} r_{t-K-1}$. The strategy corresponding to $\text{EWMA}_t(\delta, \infty) = \lim_{K\to\infty} \text{EWMA}_t(\delta, K)$ exponentially weights all past returns.

We compute all strategies indexed by (δ, K) over a grid: $\delta = 0.00, 0.01, 0.02, \ldots, 0.85$ and $K = 1, 6, 12, 24, 36, 60, 120, 240, \infty$, where the upper bound on δ is based on $\widehat{\varphi} = 0.85$, our estimate from Section 4.1.3, and the condition $\delta < \varphi$. Among all strategies on this grid over the recent 50-year sample period, we find that the highest Sharpe ratio occurs at $K = 12$ and $\delta = 0.84$, just slightly below $\widehat{\varphi} = 0.85$. We label this best-performing strategy "Binary EWMA." Table D.1 reports its Sharpe ratio as well as the alpha t-statistic of MED with respect to Binary EWMA, where MED is our static equal-blend of FAST and SLOW. Binary EWMA generates a Sharpe ratio of 0.45, which is higher than both FAST (0.34) and SLOW (0.41), however, its Sharpe ratio remains lower than that of MED (0.51) and MED exhibits statistically significant alpha with respect to Binary EWMA—t-statistic of 1.96.[38]

Table D.1: **Performance Summary: Binary EWMA vs. MED**

	Binary EWMA	$a = \frac{1}{2}$ MED
Sharpe Ratio (anlzd.)	0.45	0.51
Alpha t-statistic to Binary EWMA	—	1.96

Notes: This table reports the Sharpe ratio and alpha t-statistic with respect to the Binary EWMA strategy for $\delta = 0.84$ and $K = 12$. Excess returns on Binary EWMA are defined as $r_{t+1}^{\text{Binary}} = \text{sign}(\text{EWMA}_t(\delta, K)) \cdot r_{t+1}$, where $\text{EWMA}_t(\delta, K) = r_t + \delta r_{t-1} + \delta^2 r_{t-2} + \cdots + \delta^{K-1} r_{t-K-1}$ is the exponentially weighted moving average of past excess returns and r_{t+1} is the 1-month U.S. excess value-weighted factor return.

MED strategy weights are formed by mixing slow and fast strategies with mixing parameter $a = \frac{1}{2}$: $w_t(a) = (1-a)w_{\text{SLOW},t} + a w_{\text{FAST},t}$. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. MED excess returns are formed as $r_{t+1}(a) = w_t(a) r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return.

Data source: Kenneth French Data Library. The evaluation period is 1969-01 to 2018-12.

Internet Appendix D.C. Linear EWMA Strategies

We consider the collection of linear strategies indexed by (δ, K) such that each strategy invests units proportional to $\text{EWMA}_t(\delta, K)$ in the underlying excess return:

$$w_t(\delta, K) = \kappa \cdot \text{EWMA}_t(\delta, K),$$

[38] We obtain similar results (unreported) over the longer sample using data back to 1926.

where $\kappa > 0$ is a positive scalar that governs the unconditional volatility of the strategy. For Sharpe ratios and alpha t-statistics we set $\kappa = 1$ without loss of generality. As in the binary application, we can view $\text{EWMA}_t(\delta, K)$ as an estimate of conditional expected returns at date t, assigning higher weight to more recent returns and lower weight to less recent returns. Rather than going fully long or short one unit, however, we invest in proportion to the magnitude of $\text{EWMA}_t(\delta, K)$. Because this family of strategies no longer nests SLOW and FAST for $c = 0$, we estimate c as $\widehat{c} = \widehat{\mathbf{E}[r_t]}(1 - \widehat{\varphi})$, where $\widehat{\mathbf{E}[r_t]} = 0.49\%$ is the in-sample monthly average excess return (over the recent 50-year period) and use $\varphi = 0.85$, our estimate from Section 4.1.3. For $K \to \infty$, this signal family nests the conditional expected return of (D78). As before, we compute all strategies indexed by (δ, K) over the following grid: $\delta = 0.00, 0.01, 0.02, \ldots, 0.85$ and $K = 1, 6, 12, 24, 36, 60, 120, 240, \infty$, where the upper bound on δ is based on $\widehat{\varphi} = 0.85$, our estimate from Section 4.1.3, and the condition $\delta < \varphi$.

We find that the highest Sharpe ratio among all linear strategies on this grid over the recent 50-year sample period occurs at $K = 12$ and $\delta = 0.76$. We label this strategy "Linear EWMA." Table D.1 reports its Sharpe ratio as well as the alpha t-statistic of MED with respect to Linear EWMA, where MED is our static equal-blend of FAST and SLOW. Linear EWMA generates a Sharpe ratio of 0.46, which is slightly higher than Binary EWMA (0.45), however, its Sharpe ratio remains lower than that of MED (0.51) and MED exhibits statistically significant alpha with respect to Linear EWMA—t-statistic of 2.48.[39]

Table D.2: **Performance Summary: Linear EWMA vs. MED**

	Linear EWMA	$a = \frac{1}{2}$ MED
Sharpe Ratio (anlzd.)	0.46	0.51
Alpha t-statistic to Linear EWMA	—	2.48

Notes: This table reports the Sharpe ratio and alpha t-statistic with respect to Linear EWMA for $c = 0.00073825$, $\varphi = 0.85$, $\delta = 0.76$ and $K = 12$. Excess returns on Linear EWMA are defined as $r_{t+1}^{\text{Linear}} = \text{EWMA}_t(\delta, K) \cdot r_{t+1}$, where $\text{EWMA}_t(\delta, K) = \frac{c}{1-\delta} + (\varphi - \delta)(r_t + \delta r_{t-1} + \delta^2 r_{t-2} + \cdots + \delta^{K-1} r_{t-K-1})$ is the truncated exponentially weighted moving average of past excess returns plus a constant and r_{t+1} is the 1-month U.S. excess value-weighted factor return.

MED strategy weights are formed by mixing slow and fast strategies with mixing parameter $a = \frac{1}{2}$: $w_t(a) = (1-a)w_{\text{SLOW},t} + a w_{\text{FAST},t}$. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. MED excess returns are formed as $r_{t+1}(a) = w_t(a) r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return.

Data source: Kenneth French Data Library. The evaluation period is 1969-01 to 2018-12.

[39]We obtain similar results (unreported) over the longer sample using data back to 1926.

Internet Appendix E. MOM6 vs. Various Speeds

Table E.1: **Momentum Performance by Speed**

	MOM6	$a = 0$ SLOW MOM12	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST MOM1
Return and Risk						
Average (%) (anlzd.)	4.75	6.46	6.17	5.88	5.59	5.30
Volatility (%) (anlzd.)	15.67	15.62	12.72	11.60	12.74	15.66
Sharpe Ratio (anlzd.)	0.30	0.41	0.48	0.51	0.44	0.34
Market Timing						
Average Position	0.33	0.46	0.39	0.32	0.25	0.18
Market Beta	−0.01	0.15	0.05	−0.04	−0.13	−0.23
Alpha (%) (anlzd.)	4.79	5.58	5.85	6.12	6.39	6.66
Alpha t-statistic	2.15	2.54	3.24	3.71	3.57	3.07
Tail Behavior						
Skewness	−0.61	−0.43	−0.13	0.02	0.03	0.15
Max. Drawdown (%)	−64.56	−43.43	−37.96	−34.43	−34.07	−44.53
Average (anlzd.)/|Max. DD|	0.07	0.14	0.16	0.17	0.17	0.12

Notes: This table reports the sample average, volatility, Sharpe ratio, skewness, maximum drawdown, ratio of average return to absolute maximum drawdown, average position, the market beta and alpha, and the alpha t-statistic for monthly returns of the MOM6 strategy and of TS momentum strategies of various speeds. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The FAST strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$, for $a \in [0,1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return (Mkt-RF) from the Kenneth French Data Library. A MOMk strategy is long one unit ($+1$) if the trailing k-month return is nonnegative, and otherwise it is short one unit (-1). The evaluation period is 1969-01 to 2018-12.

Internet Appendix F. Static Speed Performance—Since 1927

Table F.1: **Performance Summary by Speed**

	Market	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Return and Risk						
Average (%) (anlzd.)	7.72	7.60	7.00	6.40	5.81	5.21
Volatility (%) (anlzd.)	18.56	18.56	15.33	14.11	15.38	18.63
Sharpe Ratio (anlzd.)	0.42	0.41	0.46	0.45	0.38	0.28
Market Timing						
Average Position	1.00	0.45	0.39	0.33	0.27	0.21
Market Beta	1.00	0.07	0.05	0.03	0.00	−0.02
Alpha (%) (anlzd.)	0.00	7.03	6.62	6.21	5.80	5.38
Alpha t-statistic	—	3.61	4.11	4.18	3.58	2.74
Tail Behavior						
Skewness	0.19	−0.28	−0.02	0.14	0.12	0.15
Max. Drawdown (%)	−84.69	−68.73	−57.08	−64.00	−74.29	−82.76
Average (anlzd.)/|Max. DD|	0.09	0.11	0.12	0.10	0.08	0.06

Notes: This table reports the sample average, volatility, Sharpe ratio, average position, market beta and alpha, alpha t-statistic, skewness, maximum drawdown, and ratio of average return to absolute maximum drawdown for monthly returns of the long-only market strategy (Market) and of TS momentum strategies of various speeds. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals +1 if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals −1. The fast strategy weight, $w_{\text{FAST},t}$, equals +1 if the trailing 1-month return is nonnegative, and otherwise equals −1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$, for $a \in [0,1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return (Mkt-RF) from the Kenneth French Data Library. The evaluation period is 1927-07 to 2018-12.

Internet Appendix G. Static Speed Performance—Equity Factors

Table G.1: **Performance Summary: Equity Factors**

Panel A: Market	Mkt-Rf	$a=0$ SLOW	$a=\frac{1}{4}$	$a=\frac{1}{2}$ MED	$a=\frac{3}{4}$	$a=1$ FAST	Average of SLOW, FAST
Sharpe Ratio (anlzd.)	0.42	0.41	0.46	0.45	0.39	0.28	0.35
Alpha to Mkt-Rf t-stat.	—	3.61	4.11	4.18	3.58	2.74	—
Panel B: Size	**SMB**						
Sharpe Ratio (anlzd.)	0.22	0.30	0.35	0.37	0.35	0.31	0.30
Alpha to SMB t-stat.	—	2.29	2.79	3.08	3.02	2.73	—
Panel C: Value	**HML**						
Sharpe Ratio (anlzd.)	0.36	0.34	0.40	0.44	0.43	0.38	0.36
Alpha to HML t-stat.	—	2.78	3.35	3.69	3.57	3.15	—
Panel D: Profitability	**RMW**						
Sharpe Ratio (anlzd.)	0.41	0.38	0.50	0.59	0.59	0.52	0.45
Alpha to RMW t-stat.	—	2.85	3.71	4.33	4.30	3.84	—
Panel E: Investment	**CMA**						
Sharpe Ratio (anlzd.)	0.49	0.36	0.47	0.57	0.59	0.55	0.46
Alpha to CMA t-stat.	—	1.51	2.34	3.09	3.40	3.34	—

Notes: This table reports the sample Sharpe ratio and alpha t-statistic for monthly returns of the long-only strategy and of TS momentum strategies of various speeds for each of five equity factors (Mkt-Rf, SMB, HML, RMW, CMA) from Kenneth French Data Library. The evaluation period for Mkt-Rf, SMB, and HML is 1927-07 to 2018-12, and for RMW and CMA is 1964-07 to 2018-12 (first year of data before 1927-07 and 1964-07, respectively, is for computing the SLOW signal). The slow strategy weight in month $t+1$, $w_{\text{SLOW},t}$, equals +1 if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals −1. The fast strategy weight, $w_{\text{FAST},t}$, equals +1 if the trailing 1-month return is nonnegative, and otherwise equals −1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$, for $a \in [0,1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the factor return. The table also reports the average of the Sharpe ratios of SLOW and FAST.

Internet Appendix H. Timing vs. Static Component Performance

Table H.1: **Market Timing and Static Exposure, U.S. Stock Market**

Panel A: Performance of the Market-Timing Component					
$(w_t(a) - \mathbf{E}[w_t(a)])r_{t+1}$	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Average (%) (anlzd.)	3.76	3.88	4.01	4.13	4.26
Volatility (%) (anlzd.)	16.18	13.71	12.85	13.91	16.51
Sharpe Ratio (anlzd.)	0.23	0.28	0.31	0.30	0.26
Skewness	−0.02	0.17	0.23	0.22	0.29
Max. Drawdown (%)	−63.33	−49.85	−46.58	−43.76	−50.23
Avg. (%) (anlzd.)/|Max. DD|	0.06	0.08	0.09	0.09	0.08
Avg. Position	0.00	0.00	0.00	0.00	0.00
Market Beta	−0.31	−0.33	−0.36	−0.38	−0.41
Alpha (%) (anlzd.)	5.58	5.85	6.12	6.39	6.66
Alpha t-stat	2.54	3.24	3.71	3.57	3.07

Panel B: Performance of Static Component					
$\mathbf{E}[w_t(a)]r_{t+1}$	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Average (%) (anlzd.)	2.70	2.28	1.87	1.46	1.04
Volatility (%) (anlzd.)	7.14	6.05	4.95	3.86	2.76
Sharpe Ratio (anlzd.)	0.38	0.38	0.38	0.38	0.38
Skewness	−0.55	−0.55	−0.55	−0.55	−0.55
Max. Drawdown (%)	−28.54	−24.69	−20.66	−16.45	−12.04
Avg. (%) (anlzd.)/|Max. DD|	0.09	0.09	0.09	0.09	0.09
Avg. Position	0.46	0.39	0.32	0.25	0.18
Market Beta	0.46	0.39	0.32	0.25	0.18
Alpha (%) (anlzd.)	0.00	0.00	0.00	0.00	0.00

Notes: This table reports the sample average, volatility, Sharpe ratio, skewness, maximum drawdown, ratio of average return to absolute maximum drawdown, average position, market beta and alpha, and alpha t-statistic for monthly returns of the static (Panel B) and market-timing (Panel A) components of the long-only market strategy (Market) and of TS momentum strategies of various speeds. The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$, for $a \in [0, 1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted factor return (Mkt-RF) from the Kenneth French Data Library. The static component return is $\mathbf{E}[w_t(a)]r_{t+1}$. The market-timing component return is $(w_t(a) - \mathbf{E}[w_t(a)])r_{t+1}$. The evaluation period is 1969-01 to 2018-12.

Internet Appendix I. Beta and Alpha—Last 15 Years

Table I.1: **Beta and Alpha Decompositions by Speed**

Market Beta and Alpha	$a = 0$ SLOW	$a = \frac{1}{4}$	$a = \frac{1}{2}$ MED	$a = \frac{3}{4}$	$a = 1$ FAST
Beta	0.20	0.14	0.08	0.01	−0.05
Alpha (%) (anlzd.)	6.08	5.40	4.72	4.04	3.37
Alpha t-stat.	1.70	1.83	1.73	1.34	0.91
Beta Components					
Static	0.722	0.619	0.517	0.414	0.311
Market Timing	0.007	0.006	0.005	0.003	0.002
Volatility Timing	−0.528	−0.487	−0.445	−0.403	−0.362
Alpha Components (%) (anlzd.)					
Market Timing	2.09	1.72	1.36	0.99	0.63
Volatility Timing	4.00	3.69	3.37	3.06	2.74

Notes: This table reports the sample market beta, alpha, and alpha t-statistic of monthly returns for momentum strategies of various speeds repeated from Table 2. The table also reports the (additive) decomposition of beta into static, market-timing, and volatility-timing components according to estimates of the terms in

$$\mathbf{Beta}[r_{t+1}(a)] = \underbrace{\mathbf{E}[w_t(a)]}_{\substack{\text{static}\\ \text{component}}} + \underbrace{\frac{\mathbf{Cov}[w_t(a), r_{t+1}]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}]}_{\substack{\text{market timing}\\ \text{component}}} + \underbrace{\frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}}_{\substack{\text{volatility timing}\\ \text{component}}}; \tag{26}$$

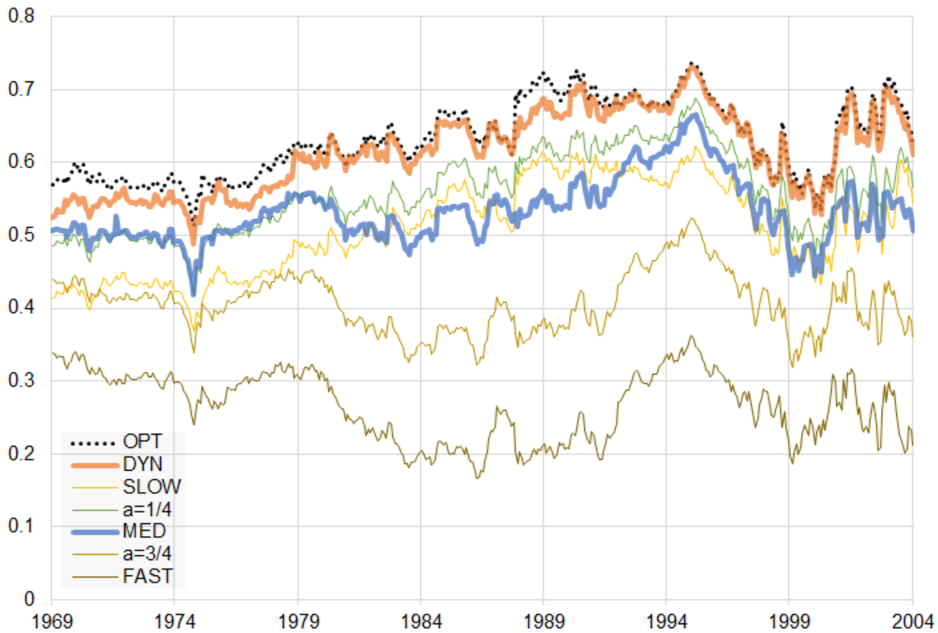
and the (additive) decomposition of alpha into market timing and volatility timing according to estimates of the terms in

$$\mathbf{Alpha}[r_{t+1}(a)] = \mathbf{Cov}[w_t(a), r_{t+1}]\left(1 - \frac{(\mathbf{E}[r_{t+1}])^2}{\mathbf{Var}[r_{t+1}]}\right) - \frac{\mathbf{Cov}[w_t(a), (r_{t+1} - \mathbf{E}[r_{t+1}])^2]}{\mathbf{Var}[r_{t+1}]}\mathbf{E}[r_{t+1}]. \tag{27}$$

The slow strategy weight applied to the market return in month $t+1$, $w_{\text{SLOW},t}$, equals $+1$ if the trailing 12-month return (arithmetic average monthly return) is nonnegative, and otherwise equals -1. The fast strategy weight, $w_{\text{FAST},t}$, equals $+1$ if the trailing 1-month return is nonnegative, and otherwise equals -1. Intermediate-speed strategy weights, $w_t(a)$ are formed by mixing slow and fast strategies with mixing parameter a: $w_t(a) = (1-a)w_{\text{SLOW},t} + aw_{\text{FAST},t}$, for $a \in [0, 1]$. Strategy returns are formed as $r_{t+1}(a) = w_t(a)r_{t+1}$, where r_{t+1} is the U.S. excess value-weighted market factor return (Mkt-RF) from the Kenneth French Data Library. The evaluation period is 2004-01 to 2018-12.

Internet Appendix J. DYN vs. Static Strategies on U.S. Equities

Figure J.1: **Sharpe Ratios (Month t to End of Sample): DYN vs. Static Strategies**



Beginning Month t to 2018-12

Notes: This figure plots, for each month t in 1969-01 to 2004-01, the Sharpe ratio of various strategies over the evaluation period beginning month t and ending December 2018. Strategies OPT and DYN are the same as those described in Table 6. Static-speed strategies are the same as those described in Table 2.