A Theoretical Analysis of Momentum and Value Strategies

Dimitri Vayanos*
LSE, CEPR and NBER

Paul Woolley†
LSE

July 5, 2012‡

Abstract

We explore implications of the rational theory of momentum and reversal in Vayanos and Woolley (2011) for empirical work and portfolio management. We compute closed-form Sharpe ratios of various implementations of momentum and value strategies, of combinations of these strategies, and for general investment horizons. For plausible parameter values, the correlation between momentum and value returns is negative, momentum exhibits positive serial correlation for short lags and zero for longer lags, and value exhibits positive serial correlation for short lags and negative for longer lags. While the Sharpe ratio for momentum exceeds that of value for short horizons, the comparison can reverse for long horizons.

*d.vayanos@lse.ac.uk
†p.k.woolley@lse.ac.uk
‡We thank seminar participants at the Bank of Canada, Boston University, Cass Business School, CEMFI, Chicago, Hitotsubashi University, LSE, Lugano, Oxford, Tilburg, USC, and participants at the Nippon Finance Association Conference for helpful comments.

1 Introduction

Momentum and value strategies underlie much of active management. Momentum strategies extrapolate recent trends, buying assets whose price has increased in the recent past and selling assets whose price has decreased. Value strategies exploit differences between price and measures of fundamental value, e.g, earnings or book equity, buying assets whose price is low relative to fundamental value and selling assets whose price is high. A large empirical literature documents that momentum and value strategies are profitable.[1]

The design of momentum and value strategies has mainly been driven by the empirical findings, while theory has provided little guidance. A theoretical perspective could be valuable, however, both to shed light on why the strategies are profitable and to suggest possible improvements in the way they are implemented. Theoretical guidance has been limited partly because of a lack of theories explaining the simultaneous presence of momentum and value effects. One could alternatively perform a partial-equilibrium analysis by specifying exogenous price processes that exhibit these effects. But since momentum and value strategies typically involve multiple assets, such a specification would involve many degrees of freedom, which an equilibrium model could help restrict.

In this paper we study the design and performance of momentum and value strategies within the theoretical framework of Vayanos and Woolley (VW 2011). In that theory, momentum and value effects arise because of flows between investment funds. Negative shocks to assets' fundamental values trigger outflows from funds holding those assets. Outflows cause asset sales, which amplify the shocks' negative effects. If the outflows are gradual because of, e.g., investor inertia or institutional constraints, then the amplification is also gradual and momentum effects arise. Moreover, because flows push prices away from fundamental value, value effects also arise. Both effects arise—and are sizeable in a calibration—despite investors and fund managers being rational.

We employ the model of VW as a laboratory to evaluate the performance of a host of active strategies. The model allows for multiple risky assets, and hence for the design of elaborate strategies. It also has a tractable linear structure, which allows for closed-form calculations of measures of strategy performance. We compute Sharpe ratios of various implementations of momentum and value strategies, of combinations of these strategies, and for general investment horizons.

[1]Short-run momentum was first documented by Jegadeesh and Titman (1993), and the value effect by Fama and French (1992). Closely related to the value effect is long-run reversal, first documented by DeBondt and Thaler (1985). Surveys of the literature include Fama (1991) and Schwert (2003).

We show that the Sharpe ratio of a strategy depends on how it loads on a time-varying premium of a risk factor associated with fund flows. We also compute the optimal strategy, and use it as a benchmark to evaluate momentum and value strategies and their combinations. We further decompose the Sharpe ratios of momentum and value strategies into intuitive components, whose relative importance we measure in our calibration.

We calibrate the model using evidence on fund flows and returns from a recent empirical literature. The calibration is as in VW, and yields static Sharpe ratios of 40% for momentum and 26% for value. We show additionally the following main results:

- Value strategies are less sensitive to implementation than momentum strategies. In particular, the quality of forecast of fundamental value has only a small effect on the Sharpe ratio of a value strategy.
- The correlation between momentum and value returns is slightly negative, equal to minus 3%. Thus, combining momentum and value strategies yields significant diversification benefits, as shown empirically by Asness, Moskowitz, and Pedersen (2009). The Sharpe ratio of the optimal combination is 48%.
- The Sharpe ratio of the optimal combination of momentum and value strategies is significantly smaller than of the overall optimal strategy, which is 61%. Thus, momentum and value strategies can be improved. This can be done by using information on fund flows.
- Returns of momentum strategies are positively autocorrelated over lags shorter than one year, and the autocorrelation over longer lags drops to zero. Returns of value strategies are also positively autocorrelated over lags shorter than one year, but the autocorrelation over longer lags is negative. Thus, over intervals longer than one year, momentum is a series of *i.i.d.* bets, but value exhibits mean reversion.
- The dynamic Sharpe ratio of momentum strategies decreases with the investment horizon when the horizon is short, and becomes essentially flat after a one-year horizon. The dynamic Sharpe ratio of value strategies, decreases with the investment horizon when the horizon is short. As the horizon increases, however, it increases and eventually overtakes the Sharpe ratio of momentum strategies.

Section 2 provides a brief overview of the VW model, and Section 3 of the main properties of the equilibrium in the case where information about fund managers' efficiency is asymmetric.

Section 4 computes the Sharpe ratio of a general trading strategy. Section 5 computes the Sharpe ratio of momentum strategies. We consider two implementations of momentum: one that uses raw past returns and one that uses risk-adjusted returns. Section 6 computes the Sharpe ratio of value strategies. We consider four implementations of value, obtained by using raw or risk-adjusted prices, and a perfect or a crude forecast for expected dividends. Section 7 computes the correlation between momentum and value strategies, as well as the Sharpe ratio that can be achieved by combining them. Sections 4-7 concern infinitesimal investment horizons and static Sharpe ratios. Section 8 extends the analysis to general non-infinitesimal horizons, and computes dynamic Sharpe ratios.

Behavioral theories of momentum and reversal include Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Hong and Stein (1999), and Barberis and Shleifer (2003). Rational theories include Albuquerque and Miao (2010), Cespa and Vives (2011), and Dasgupta, Prat, and Verardo (2011). Our theory has similarities to Barberis and Shleifer (2003) who emphasize flows across styles rather than across investment funds, and to Dasgupta, Prat, and Verardo (2011) who emphasize reputation concerns of fund managers. The only paper that assumes multiple assets, as we do, is Barberis and Shleifer (2003), and it explores implications for portfolio choice. It does not address, however, any of the issues that we are studying in this paper, e.g., the robustness of a strategy to implementation, the correlation between strategies, and the performance over long horizons. Koijen, Rodriguez, and Sbuelz (2006) perform portfolio optimization for an exogenous price process that exhibits momentum and reversal effects, in the case of one risky asset.

2 Model

In this section we describe briefly the model of Vayanos and Woolley (VW 2011), referring to VW for more detailed discussion and motivation of the assumptions. We consider only the case of asymmetric information because it is under that case that we evaluate the performance of trading strategies. Time t is continuous and goes from zero to infinity. There are N risky assets and a riskless asset. We refer to the risky assets as stocks. The riskless asset has an exogenous, continuously compounded return r. The stocks pay dividends over time, and their prices are determined endogenously in equilibrium. We denote by D_{nt} the cumulative dividend per share of stock $n = 1, .., N$, by S_{nt} the stock's price, and by π_n the stock's supply in terms of number of shares. We specify the stochastic process for dividends later in this section.

A competitive investor can invest in the riskless asset and in the stocks. The investor can access the stocks only through two investment funds: a passive fund that tracks mechanically a market index, and an active fund. We assume that the market index includes a fixed number η_n of shares of stock n. Thus, if the vectors $\pi \equiv (\pi_1, .., \pi_N)$ and $\eta \equiv (\eta_1, .., \eta_N)$ are collinear, the market index is capitalization-weighted and coincides with the market portfolio.

To ensure that the active fund can add value over the index fund, we assume that the market index differs from the true market portfolio characterizing equilibrium asset returns. This can be because the market index does not include some stocks. Alternatively, the market index can coincide with the market portfolio, but unmodelled buy-and-hold investors, such as firms' managers or founding families, can hold a portfolio different from the market portfolio. That is, buy-and-hold investors hold $\hat{\pi}_n$ shares of stock n, and the vectors π and $\hat{\pi} \equiv (\hat{\pi}_1, .., \hat{\pi}_N)$ are not collinear. To nest the two cases, we define a vector $\theta \equiv (\theta_1, .., \theta_N)$ to coincide with π in the first case and $\pi - \hat{\pi}$ in the second. The vector θ represents the residual supply left over from buy-and-hold investors, and is the true market portfolio characterizing equilibrium asset returns. We assume that θ is not collinear with the market index η, and set

$$\Delta \equiv \theta\Sigma\theta'\eta\Sigma\eta' - (\eta\Sigma\theta')^2 > 0.$$

The investor determines how to allocate her wealth between the riskless asset, the index fund, and the active fund. She maximizes expected utility of intertemporal consumption. Utility is exponential, i.e.,

$$-E\int_0^\infty \exp(-\alpha c_t - \beta t)dt, \tag{2.1}$$

where α is the coefficient of absolute risk aversion, c_t is consumption, and β is the discount rate. The investor's control variables are consumption c_t and the number of shares x_t and y_t of the index and active fund, respectively.

The active fund is run by a competitive manager, who can also invest his personal wealth in the fund. The manager determines the active portfolio and the allocation of his wealth between the riskless asset and the fund. He maximizes expected utility of intertemporal consumption. Utility is exponential, i.e.,

$$-E\int_0^\infty \exp(-\bar{\alpha}\bar{c}_t - \bar{\beta} t)dt, \tag{2.2}$$

where $\bar{\alpha}$ is the coefficient of absolute risk aversion, $\bar{c}_t$ is consumption, and $\bar{\beta}$ is the discount rate. The manager's control variables are consumption $\bar{c}_t$, the number of shares $\bar{y}_t$ of the active fund, and the active portfolio $z_t \equiv (z_{1t}, .., z_{Nt})$, where z_{nt} denotes the number of shares of stock n included in one share of the active fund. The assumption that the manager can invest his personal wealth in the active fund is for parsimony: in addition to choosing the active portfolio, the manager acts as trading counterparty to the investor's flows, and this eliminates the need to introduce additional agents into the model. The manager in our model can be viewed as the aggregate of all agents absorbing the investor's flows.

To ensure that the investor has a motive to move across funds and generate flows, we assume that she suffers a time-varying cost from investing in the active fund. This cost drives a wedge between the investor's net return from the fund, and the gross return made of the dividends and capital gains of the stocks held by the fund. The interpretation of the cost that best fits our model is as a managerial perk, although other interpretations such as managerial ability could fit more complicated versions of the model. The index fund entails no cost, so its gross and net returns coincide.

We model the cost as a flow (i.e., the cost between t and $t+dt$ is of order dt), and assume that the flow cost is proportional to the number of shares y_t that the investor holds in the active fund. We denote the coefficient of proportionality by C_t and assume that it follows the process

$$dC_t = \kappa(\bar{C} - C_t)dt + sdB_t^C, \tag{2.3}$$

where κ is a mean-reversion parameter, $\bar{C}$ is a long-run mean, s is a positive scalar, and B_t^C is a Brownian motion.

To remain consistent with the managerial-perk interpretation of the cost, we allow the manager to derive a benefit from the investor's participation in the active fund. We model the benefit in the same manner as the cost, i.e., a flow which is proportional to the number of shares y_t that the investor holds in the active fund. If the cost is a perk that the manager can extract efficiently, then the coefficient of proportionality for the benefit is C_t. We allow more generally the coefficient of proportionality to be λC_t, where $\lambda \geq 0$ is a constant that an be interpreted as the efficiency of perk extraction. Varying λ generates a rich specification of the manager's objective. When $\lambda = 0$, the manager cares about fund performance only through his personal investment in the fund, and his objective is similar to the fund investor's. When instead $\lambda > 0$, the manager is also concerned with commercial risk, i.e., the risk that the investor might reduce her participation in the fund.

We define one share of the fund by the requirement that its market value equals the equilibrium market value of the entire fund. Under this definition, the number of fund shares held by the investor and the manager in equilibrium sum to one, i.e.,

$$y_t + \bar{y}_t = 1. \tag{2.4}$$

We define one share of the index fund to coincide with the market index η.

We denote the vector of stocks' cumulative dividends by $D_t \equiv (D_{1t}, .., D_{Nt})'$ and the vector of stock prices by $S_t \equiv (S_{1t}, .., S_{Nt})'$, where v' denotes the transpose of the vector v. We assume that D_t follows the process

$$dD_t = F_t dt + \sigma dB_t^D, \tag{2.5}$$

where $F_t \equiv (F_{1t}, .., F_{Nt})'$ is a time-varying drift equal to the expected dividend rate, σ is a constant matrix of diffusion coefficients, and B_t^D is a d-dimensional Brownian motion independent of B_t^C. We assume a time-varying expected dividend F_t so that prices do not reveal the cost C_t to the investor perfectly. We model time-variation in F_t through the process

$$dF_t = \kappa(\bar{F} - F_t)dt + \phi\sigma dB_t^F \tag{2.6}$$

where the mean-reversion parameter κ is the same as for C_t for simplicity, $\bar{F}$ is a long-run mean, ϕ is a positive scalar, and B_t^F is a d-dimensional Brownian motion independent of B_t^C and B_t^D. The diffusion matrices for D_t and F_t are proportional for simplicity. We set $\Sigma \equiv \sigma\sigma'$.

We assume that the investor can adjust her active-fund holdings y_t to new information only gradually. Gradual adjustment can result from contractual restrictions or institutional decision lags. We model these frictions as a flow cost $\psi(dy_t/dt)^2/2$ that the investor must incur when changing y_t.

The manager observes all the variables in the model. The investor observes the returns and share prices of the index and active funds, but not the same variables for the individual stocks. She does not observe C_t and F_t.

3 Equilibrium

In this section we describe briefly the equilibrium of Vayanos and Woolley (VW 2011) in the case of asymmetric information. The equilibrium has with the following characteristics. The investor's conditional distribution of C_t is normal with mean $\hat{C}_t$. Stock prices are

$$S_t = \frac{\bar{F}}{r} + \frac{F_t - \bar{F}}{r+\kappa} - (a_0 + a_1\hat{C}_t + a_2 C_t + a_3 y_t), \tag{3.1}$$

where (a_0, a_1, a_2, a_3) are constant vectors. The first two terms in (3.1) are the present value of expected dividends, discounted at the riskless rate r, and the last term is a risk discount linear in the variables $(\hat{C}_t, C_t, y_t)$. The effects of $(\hat{C}_t, C_t, y_t)$ are described by the vectors (a_1, a_2, a_3). These vectors are

$$a_i = \gamma_i \Sigma p_f' \tag{3.2}$$

for $i = 1, 2, 3$, where $(\gamma_1, \gamma_2, \gamma_3)$ are constants (scalars) and

$$p_f \equiv \theta - \frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}\eta \tag{3.3}$$

is the "flow portfolio." The flow portfolio characterizes the flows that the investor generates when moving across funds. For example, when moving out of the active and into the index fund, she sells a slice of the flow portfolio. Indeed, the flows that she generates amount to selling stocks that the active fund overweights relative to the index fund, and buying stocks that the active fund underweights. Since market clearing requires that the active fund holds the true market portfolio θ net of the investor's holdings of the market index η, the stocks that the active fund overweights correspond to large components of θ relative to η, and hence to long positions in the flow portfolio. Conversely, the stocks that the active fund underweights correspond to short positions.

Eqs. (3.1) and (3.2) indicate that changes in $(\hat{C}_t, C_t, y_t)$ affect stock prices through the covariance with the flow portfolio. Consider, for example, an increase in the investor's holdings y_t of the active fund. This corresponds to a flow out of the index and into the active fund, and hence to a purchase of a slice of the flow portfolio by the investor. Market clearing requires that the manager takes the other side of this transaction (by changing his personal stake in the active fund). For the manager to be induced to do so, the expected returns of stocks that covary positively with the flow portfolio must decrease, and those of stocks that covary negatively must increase. Therefore, the price of the former stocks increases and of the latter decreases. This means that the constant γ_3 in (3.2) is negative, as confirmed in Proposition 3.1. The constants γ_1 and γ_2 are instead positive because increases in $\hat{C}_t$ and C_t correspond to outflows from the active fund. Indeed, an increase in the investor's estimate $\hat{C}_t$ of the cost C_t reduces her target holdings of the active fund, and corresponds to future outflows. And an increase in C_t forecasts a reduction of the investor's target holdings of the active fund in the future, as the investor learns that C_t has increased.

The dynamics of y_t in the equilibrium derived in VW are

$$v_t \equiv \frac{dy_t}{dt} = b_0 - b_1\hat{C}_t - b_2 y_t, \tag{3.4}$$

where (b_0, b_1, b_2) are constants. Intuitively, the investor's holdings y_t of the active fund should evolve towards a time-varying target, which is decreasing in $\hat{C}_t$. Thus, the constants b_1 and b_2 are positive, as confirmed in Proposition 3.1. The dynamics of $\hat{C}_t$ are

$$\begin{aligned} d\hat{C}_t =& \kappa(\bar{C} - \hat{C}_t)dt - \beta_1 \left\{ p_f \left[dD_t - E_t(dD_t)\right] - (C_t - \hat{C}_t)dt \right\} \\ & - \beta_2 p_f \left[dS_t + a_1 d\hat{C}_t + a_3 dy_t - E_t(dS_t + a_1 d\hat{C}_t + a_3 dy_t) \right], \end{aligned} \tag{3.5}$$

where

$$\beta_1 \equiv T \left[1 - (r+k)\frac{\gamma_2 \Delta}{\eta \Sigma \eta'} \right] \frac{\eta \Sigma \eta'}{\Delta}, \tag{3.6}$$

$$\beta_2 \equiv \frac{s^2 \gamma_2}{\frac{\phi^2}{(r+\kappa)^2} + \frac{s^2 \gamma_2^2 \Delta}{\eta \Sigma \eta'}}, \tag{3.7}$$

and T denotes the investor's conditional variance of C_t, which is constant in the steady state reached for $t \to \infty$. Eq. (3.5) characterizes the investor's dynamic learning about C_t, and is derived using Kalman filtering. The investor learns about C_t by observing the returns and prices of the index and active funds. Since returns are made of dividends and capital gains, observing fund returns and prices is equivalent to observing fund dividends and prices. The terms in β_1 and β_2 in (3.5) represent the learning from dividends and prices, respectively. In the case of dividends, the investor raises her estimate of C_t if the net dividends of the active fund are low relative to those of the index fund. This is because high C_t lowers the net dividends of the active fund. In the case of prices, the investor raises her estimate of C_t if the price of the index fund is low relative to that of the index fund. This is because high C_t forecasts future outflows by the active fund and this lowers the price of the active portfolio.

VW refer to an equilibrium satisfying (3.1), (3.2) and (3.4) as linear, and show that a unique linear equilibrium exists when the diffusion coefficient s of C_t is small. Moreover, numerical solutions for general values of s yield a linear equilibrium with the same properties as in Proposition 3.1.

Proposition 3.1 *For small s, there exists a unique linear equilibrium. The constants* $(b_1, b_2, \gamma_1, \gamma_2)$ *are positive and the constant* γ_3 *is negative.*

The properties of the equilibrium that are relevant for our analysis concern the comovement of stock returns, the cross section of expected returns, and the predictability of returns. We consider returns per share in excess of the riskless asset, denote them by $dR_t \equiv dD_t + dS_t - rS_t dt$, and refer to them simply as returns.

Proposition 3.2 shows that the covariance matrix of stock returns is the sum of a fundamental covariance, driven purely by cashflows, and a non-fundamental covariance, introduced by fund flows. The non-fundamental covariance between a pair of stocks is proportional to the product of the covariances between each stock in the pair and the flow portfolio. It is thus positive for stock pairs whose covariance with the flow portfolio has the same sign, and negative otherwise. Intuitively, two stocks move in the same direction in response to fund flows if they are both overweighted or both underweighted by the active fund, but move in opposite directions if one is overweighted and the other underweighted.

Proposition 3.2 *The covariance matrix of stock returns is*

$$Cov_t(dR_t, dR_t') = \left(f\Sigma + k\Sigma p_f' p_f \Sigma\right) dt, \tag{3.8}$$

where $f \equiv 1 + \phi^2/(r+\kappa)^2$ *and* k *are positive constants. The matrix* $f\Sigma dt$ *is the fundamental covariance, driven purely by cashflows, and the matrix* $k\Sigma p_f' p_f \Sigma dt$ *is the non-fundamental covariance, introduced by fund flows. The non-fundamental covariance is positive for stock pairs whose covariance with the flow portfolio has the same sign, and is negative otherwise.*

Proposition 3.3 shows that expected returns are given by a two-factor model, with the factors being the market index η and the flow portfolio p_f. Changes in $(\hat{C}_t, C_t, y_t)$ affect expected returns through the factor risk premium Λ_t associated to the flow portfolio. For example, an increase in y_t lowers Λ_t since $\gamma_3^R < 0$. It thus lowers the expected returns of stocks that covary positively with the flow portfolio and raises those of stocks that covary negatively.

Proposition 3.3 *Stocks' expected returns are given by the two-factor model*

$$E_t(dR_t) = \frac{r\alpha\bar{\alpha}}{\alpha+\bar{\alpha}} \frac{\eta\Sigma\theta'}{\eta\Sigma\eta'} Cov_t(dR_t, \eta dR_t) + \Lambda_t Cov_t(dR_t, p_f dR_t), \tag{3.9}$$

with the factors being the market index and the flow portfolio. The factor risk premium Λ_t associated to the flow portfolio is

$$\Lambda_t = r\bar{\alpha} + \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} \left(\gamma_1^R \hat{C}_t + \gamma_2^R C_t + \gamma_3^R y_t - k_1 \bar{q}_1 - k_2 \bar{q}_2 \right), \tag{3.10}$$

where $(\gamma_1^R, \gamma_2^R, \gamma_3^R, k_1, k_2, \bar{q}_1, \bar{q}_2)$ are constants. For small s, the constants (γ_1^R, γ_3^R) are negative and the constant γ_2^R is positive.

Propositions 3.4 and 3.5 study return predictability based on cashflows and past returns, respectively. A stock's cashflow shock or return predicts positively the stock's subsequent return in the short run, implying short-run momentum, and negatively the return in the long run, implying long-run reversal.

Proposition 3.4 *The covariance between cashflow shocks (dD_t, dF_t) at time t and returns at time $t' > t$ is given by*

$$Cov_t(dD_t, dR'_{t'}) = \frac{\beta_1(r+\kappa)Cov_t(dF_t, dR'_{t'})}{\beta_2\phi^2} = \left[\chi_1^D e^{-(\kappa+\rho)(t'-t)} + \chi_2^D e^{-b_2(t'-t)} \right] \Sigma p'_f p_f \Sigma dt dt', \tag{3.11}$$

where $(\chi_1^D, \chi_2^D, \rho)$ are constants. For small s, the term in the square bracket of (3.11) is positive if $t' - t < \hat{u}^D$ and negative if $t' - t > \hat{u}^D$, for a threshold $\hat{u}^D > 0$. A stock's cashflow shocks predict positively the stock's subsequent return for $t' - t < \hat{u}^D$ (short-run momentum) and negatively for $t' - t > \hat{u}^D$ (long-run reversal). They predict in the same manner the subsequent return of another stock when the covariance between each stock in the pair and the flow portfolio has the same sign, and in the opposite manner otherwise.

Proposition 3.5 *The covariance between stock returns at time t and those at time $t' > t$ is*

$$Cov_t(dR_t, dR'_{t'}) = \left[\chi_1 e^{-(\kappa+\rho)(t'-t)} + \chi_2 e^{-\kappa(t'-t)} + \chi_3 e^{-b_2(t'-t)} \right] \Sigma p'_f p_f \Sigma dt dt', \tag{3.12}$$

where (χ_1, χ_2, χ_3) are constants. For small s, the term in the square bracket of (3.12) is positive if $t' - t < \hat{u}$ and negative if $t' - t > \hat{u}$, for a threshold $\hat{u} > 0$. A stock's return predicts positively the stock's subsequent return for $t' - t < \hat{u}$ (short-run momentum) and negatively for $t' - t > \hat{u}$

(long-run reversal). It predicts in the same manner the subsequent return of another stock when the covariance between each stock in the pair and the flow portfolio has the same sign, and in the opposite manner otherwise.

The intuition for momentum and reversal is as follows. Suppose that a negative cashflow shock hits a stock that the active fund overweights. Since this lowers the active fund's performance relative to the index fund, it raises the investor's estimate of C_t, and lowers her target holdings of the active fund. The expectation of future outflows from the active fund lowers the prices of stocks that the active fund overweights, including the stock hit by the cashflow shock. Momentum arises if this decline is expected to continue in the short run. Reversal arises because the outflows cause the prices of the stocks that the active fund overweights to drop relative to fundamental value. Hence, these stocks' expected returns eventually rise.

The emergence of momentum is surprising. Indeed, why is the price decline expected to continue in the short run? And why do rational agents take the other side of the outflows that follow the cashflow shock, buying assets whose expected returns have decreased?[2] This is because of what we term the "bird in the hand" effect. The assets that experience a price drop and are expected to continue underperforming in the short run are those that the active fund overweights. The anticipation of outflows causes these assets to be underpriced and to guarantee agents an attractive return (bird in the hand) over a long horizon. Agents could earn an even more attractive return on average (two birds in the bush), by buying these assets after the outflows occur. This, however, exposes them to the risk that the outflows might not occur, in which case the assets would cease to be underpriced. In summary, short-run expected underperformance is possible because of long-run expected overperformance; and more generally, momentum is possible because of the subsequent reversal.

The bird-in-the-hand effect can be illustrated in the following simple example. An asset is expected to pay off 100 in Period 2. The asset price is 92 in Period 0, and 80 or 100 in Period 1 with equal probabilities. Buying the asset in Period 0 earns an investor a two-period expected capital gain of 8. Buying in Period 1 earns an expected capital gain of 20 if the price is 80 and 0 if the price is 100. A risk-averse agent might prefer earning 8 rather than 20 or 0 with equal probabilities, even though the expected capital gain between Periods 0 and 1 is negative.

[2]In our model, the only agent taking the other side of the investor's flows is the manager, but this assumption is only for simplicity.

4 Performance of Trading Strategies

In this section we define a performance measure for general trading strategies. Consider a trading strategy consisting of a vector of weights $w_t \equiv (w_{1t}, .., w_{Nt})$, where w_{nt} is the number of shares invested in stock n at time t. Part of the strategy's expected return is compensation for bearing risk that correlates with the market index. We focus on the remainder by index-adjusting the strategy, i.e., combining it with a position in the index such that the covariance between the overall position and the index is zero. The index-adjusted strategy is

$$\hat{w}_t \equiv w_t - \frac{Cov_t(w_t dR_t, \eta dR_t)}{Var_t(\eta dR_t)}\eta. \tag{4.1}$$

Note that the position in the index can be time-varying, reflecting possible time-variation in the covariance between the strategy and the index. We measure the performance of the strategy w_t by the Sharpe ratio of its index-adjusted version $\hat{w}_t$.[3] The Sharpe ratio is the ratio of expected return to standard deviation. We also divide by $\sqrt{dt}$ to express the Sharpe ratio in annualized terms, given that returns are evaluated over an infinitesimal period dt. The Sharpe ratio corresponding to the strategy w_t thus is

$$SR_w \equiv \frac{E(\hat{w}_t dR_t)}{\sqrt{Var(\hat{w}_t dR_t)dt}}. \tag{4.2}$$

Proposition 4.1 computes the Sharpe ratio under the prices in the equilibrium of Section 3 and in the steady state reached for $t \to \infty$. All subsequent calculations also concern the steady state.

Proposition 4.1 *The Sharpe ratio corresponding to the strategy w_t is*

$$SR_w = \frac{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right) E\left(\Lambda_t w_t \Sigma p_f'\right)}{\sqrt{f\left[E(w_t\Sigma w_t') - \frac{E[(w_t\Sigma\eta')^2]}{\eta\Sigma\eta'}\right] + kE[(w_t\Sigma p_f')^2]}}. \tag{4.3}$$

The use of the Sharpe ratio as a performance measure can be motivated based on portfolio optimization. Consider an investor with horizon dt and mean-variance preferences, who can invest

[3] Empirical studies that compute Sharpe ratios of momentum and value strategies typically consider long-short portfolios with zero initial investment, i.e., require the dollar weights to sum to zero. Our index-adjustment is in a similar spirit: the weights that sum to zero are the number of shares times the covariance between one share and the index rather than times the dollar value of one share. We define weights differently because this preserves linearity and simplifies the algebra.

in the riskless asset, the market index η and the strategy w_t. The investor chooses investments $\hat{x}$ in the index and $\hat{y}$ in the strategy to maximize

$$E(dW_t) - \frac{a}{2}Var(dW_t), \tag{4.4}$$

subject to the budget constraint

$$\begin{aligned} dW_t &= rW_t dt + \hat{x}\eta dR_t + \hat{y}w_t dR_t \\ &= rW_t dt + \hat{\hat{x}}\eta dR_t + \hat{y}\hat{w}_t dR_t, \end{aligned} \tag{4.5}$$

where a is the investor's risk-aversion coefficient,

$$\hat{\hat{x}} \equiv \hat{x} + \hat{y}\frac{Cov_t(w_t dR_t, \eta dR_t)}{Var_t(\eta dR_t)}$$

is the investor's overall exposure to the index, and the second step in (4.5) follows from (4.1). Substituting (4.5) into (4.4), and noting that $(\eta dR_t, \hat{w}_t dR_t)$ are orthogonal, we can write the investor's maximization problem as

$$\max_{\hat{\hat{x}},\hat{y}} \left\{ \hat{\hat{x}}E(\eta dR_t) + \hat{y}E(\hat{w}_t dR_t) - \frac{a}{2}\left[\hat{\hat{x}}^2 Var(\eta dR_t) + \hat{y}^2 Var(\hat{w}_t dR_t)\right]\right\}. \tag{4.6}$$

Lemma 4.1 *The solution to the maximization problem (4.6) is*

$$\hat{\hat{x}} = \frac{E(\eta dR_t)}{aVar(\eta dR_t)}, \tag{4.7}$$

$$\hat{y} = \frac{E(\hat{w}_t dR_t)}{aVar(\hat{w}_t dR_t)}. \tag{4.8}$$

The investor's maximum utility is

$$\frac{E(\eta dR_t)^2}{2aVar(\eta dR_t)} + \frac{SR_w^2 dt}{2a}. \tag{4.9}$$

The investor's maximum utility depends on the strategy's characteristics only through the Sharpe ratio, and is increasing in the ratio's square. Hence, if the investor must choose between two strategies with positive Sharpe ratios, i.e., where the optimal investment is a long rather than a short position, he prefers the one with the largest ratio. Proposition 4.2 determines a strategy that maximizes the Sharpe ratio.

Proposition 4.2 *The Sharpe ratio in (4.3) is maximized for the strategy $w_t = \Lambda_t p_f$. The maximum value of the Sharpe ratio is*

$$\max_{w_t} SR_w = \sqrt{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{\Delta}{\eta\Sigma\eta'}E(\Lambda_t^2)}. \tag{4.10}$$

The intuition for the optimal strategy can be derived from the two-factor model of Proposition 3.3. A strategy's expected return consists of a compensation for bearing risk that correlates with the index, and a compensation for bearing risk that correlates with the flow portfolio. Index adjustment isolates the latter component. Maximizing that component per unit of risk requires holding the flow portfolio since this eliminates uncompensated risk. Moreover, investment in the flow portfolio should be larger when the premium Λ_t associated to that risk factor is high.

Lemma 4.1 and Proposition 4.2 imply that the strategy $w_t = \Lambda p_f$ maximizes the utility of an investor with horizon dt and mean-variance preferences, who can invest in the riskless asset, the market index and the strategy. Note that since the set of strategies is unrestricted, the investor's maximization problem is equivalent to choosing freely investments in all stocks, knowing perfectly the structure of the equilibrium in Section 3.

The investor's optimal strategy can be contrasted to that of the manager in Section 3. Both investor and manager know perfectly the structure of the equilibrium and can choose freely investments in all assets to maximize their utility. The key difference between them is in their horizon: the manager has a long horizon since he maximizes the utility (2.2) over intertemporal consumption, while the investor has a short horizon since she maximizes the utility (4.4) over instantaneous changes in wealth. The difference in horizon implies a difference in optimal strategies: the manager always longs the flow portfolio, as is required from market clearing, while the investor can short it if $\Lambda_t < 0$.

The investment in the flow portfolio by the optimal strategy varies over time with Λ_t. Since time-variation in Λ_t is caused by fund flows, past and anticipated, the optimal strategy effectively exploits mispricing generated by flows. Momentum and value strategies exploit aspects of the flow-generated mispricing, and are imperfect approximations of the optimal strategy.

5 Momentum Strategies

We consider two implementations of a momentum strategy:

$$\left(w_t^M\right)' \equiv \int_{t-\tau}^{t} dR_u, \tag{5.1}$$

$$\left(w_t^{\hat{M}}\right)' \equiv \int_{t-\tau}^{t} d\hat{R}_u, \tag{5.2}$$

where

$$d\hat{R}_t \equiv dR_t - \frac{Cov_t(dR_t, \eta dR_t)}{Var_t(\eta dR_t)} \eta dR_t. \tag{5.3}$$

Under both implementations, a stock's momentum weight increases linearly in the stock's cumulative past return over the window $[t-\tau, t]$ for some $\tau > 0$. The two implementations differ in the measure of past returns used to construct momentum weights: raw returns in (5.1) and index-adjusted returns in (5.2). Momentum weights are typically constructed using raw returns in empirical work and investment practice. We also consider index-adjusted returns because the calculations are simpler and because the Sharpe ratio can be higher than with raw returns.

We introduce some notation that we also use in subsequent sections. For scalars $(\psi_1, \psi_2, \psi_3, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3)$ and a function $\nu(\omega, \mathcal{T})$, we define the function $G(\psi_1, \psi_2, \psi_3, \mathcal{T}, \nu)$ by

$$\begin{aligned} &G(\psi_1, \psi_2, \psi_3, \mathcal{T}, \nu) \equiv \\ &- \left[\psi_1 \nu(\kappa+\rho, \mathcal{T}) + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(\nu(\kappa+\rho, \mathcal{T}) - \nu(b_2, \mathcal{T})\right)\right] \beta_1 \left(1 + \frac{\beta_1 \gamma_1 \Delta}{\eta \Sigma \eta'}\right) \\ &- \left[(\psi_1+\psi_2)\nu(\kappa, \mathcal{T}) + \frac{\psi_3 b_1}{\kappa - b_2}\left(\nu(\kappa, \mathcal{T}) - \nu(b_2, \mathcal{T})\right)\right] s^2 \gamma_2 \left(1 + \frac{\beta_2 \gamma_1 \Delta}{\eta \Sigma \eta'}\right), \end{aligned}$$

the function $H(\psi_1,\psi_2,\psi_3,\hat\psi_1,\hat\psi_2,\hat\psi_3,\mathcal{T},\nu)$ by

$$
\begin{aligned}
&H(\psi_1,\psi_2,\psi_3,\hat\psi_1,\hat\psi_2,\hat\psi_3,\mathcal{T},\nu)\equiv\\
&\left[\frac{1}{2(\kappa+\rho)}\left(\hat\psi_1+\frac{\hat\psi_3 b_1}{\kappa+\rho-b_2}\right)\left(\psi_1-\frac{\psi_3 b_1}{\kappa+\rho+b_2}\right)\nu(\kappa+\rho,\mathcal{T})\right.\\
&\left.-\frac{\hat\psi_3 b_1}{(\kappa+\rho+b_2)(\kappa+\rho-b_2)}\left(\psi_1-\frac{\psi_3 b_1}{2b_2}\right)\nu(b_2,\mathcal{T})\right]\left[\beta_1^2+\frac{\phi^2\beta_2^2}{(r+\kappa)^2}+\frac{s^2\beta_2^2\gamma_2^2\Delta}{\eta\Sigma\eta'}\right]\frac{\Delta}{\eta\Sigma\eta'}\\
&+\left[\frac{1}{2\kappa+\rho}\left(\hat\psi_1+\hat\psi_2+\frac{\hat\psi_3 b_1}{\kappa-b_2}\right)\left(\psi_1-\frac{\psi_3 b_1}{\kappa+b_2}\right)\nu(\kappa,\mathcal{T})\right.\\
&-\frac{1}{2\kappa+\rho}\left(\hat\psi_1+\frac{\hat\psi_3 b_1}{\kappa+\rho-b_2}\right)\left(\frac{\psi_1\kappa}{\kappa+\rho}-\psi_2-\frac{\psi_3 b_1\kappa}{(\kappa+\rho)(\kappa+\rho+b_2)}\right)\nu(\kappa+\rho,\mathcal{T})\\
&\left.-\frac{\hat\psi_3 b_1}{(\kappa+b_2)(\kappa+\rho-b_2)}\left(\frac{2\psi_1\kappa\rho}{(\kappa-b_2)(\kappa+\rho+b_2)}+\psi_2-\frac{\psi_3\kappa b_1\rho}{b_2(\kappa-b_2)(\kappa+\rho+b_2)}\right)\nu(b_2,\mathcal{T})\right]\frac{s^2\beta_2\gamma_2\Delta}{\eta\Sigma\eta'}\\
&+\left[\frac{1}{2\kappa}\left(\hat\psi_1+\hat\psi_2+\frac{\hat\psi_3 b_1}{\kappa-b_2}\right)\left(\frac{\psi_1\rho}{2\kappa+\rho}+\psi_2-\frac{\psi_3 b_1\rho}{(2\kappa+\rho)(\kappa+b_2)}\right)\nu(\kappa,\mathcal{T})\right.\\
&-\frac{1}{2\kappa+\rho}\left(\hat\psi_1+\frac{\hat\psi_3 b_1}{\kappa+\rho-b_2}\right)\left(\frac{\psi_1\rho}{2(\kappa+\rho)}+\psi_2-\frac{\psi_3 b_1\rho}{2(\kappa+\rho)(\kappa+\rho+b_2)}\right)\nu(\kappa+\rho,\mathcal{T})\\
&\left.-\frac{\hat\psi_3 b_1\rho}{(\kappa+b_2)(\kappa-b_2)(\kappa+\rho-b_2)}\left(\frac{\psi_1\rho}{\kappa+\rho+b_2}+\psi_2-\frac{\psi_3 b_1\rho}{2b_2(\kappa+\rho+b_2)}\right)\nu(b_2,\mathcal{T})\right]s^2,
\end{aligned}
$$

and the functions $K_1(\psi_1,\psi_3,\mathcal{T},\nu)$ and $K_2(\psi_1,\psi_3,\mathcal{T},\nu)$ by

$$
\begin{aligned}
&K_1(\psi_1,\psi_3,\mathcal{T},\nu)\equiv-\frac{1}{2\kappa+\rho}\left(\psi_1-\frac{\psi_3 b_1}{\kappa+b_2}\right)\nu(\kappa,\mathcal{T})\frac{\phi^2\beta_2}{r+\kappa},\\
&K_2(\psi_1,\psi_3,\mathcal{T},\nu)\equiv-\left[\frac{1}{2\kappa+\rho}\left(\psi_1+\frac{\psi_3 b_1}{\kappa+\rho-b_2}\right)\nu(\kappa+\rho,\mathcal{T})-\frac{\psi_3 b_1}{(\kappa+b_2)(\kappa+\rho-b_2)}\nu(b_2,\mathcal{T})\right]\frac{\phi^2\beta_2}{r+\kappa}.
\end{aligned}
$$

We define the functions $\{\nu_i(\omega, \mathcal{T})\}_{i=0,..,6}$ by

$$\begin{aligned}
\nu_0(\omega, \mathcal{T}) \equiv& e^{-\omega \mathcal{T}}, \\
\nu_1(\omega, \mathcal{T}) \equiv& \frac{1 - e^{-\omega \mathcal{T}}}{\omega}, \\
\nu_2(\omega, \mathcal{T}) \equiv& \frac{1}{\omega}\left(\mathcal{T} - \frac{1 - e^{-\omega \mathcal{T}}}{\omega}\right), \\
\nu_3(\omega, \mathcal{T}) \equiv& e^{-\omega \mathcal{T}_1} \frac{1 - e^{-\omega \mathcal{T}_2}}{\omega}, \\
\nu_4(\omega, \mathcal{T}) \equiv& e^{-\omega \mathcal{T}_1} \frac{e^{\omega \min\{\mathcal{T}_1, \mathcal{T}_2\}} - 1}{\omega}, \\
\nu_5(\omega, \mathcal{T}) \equiv& e^{-\omega \mathcal{T}_1} \frac{e^{\omega (2\min\{\mathcal{T}_1, \mathcal{T}_2\} - \mathcal{T}_2)} + e^{-\omega \mathcal{T}_2} - 2}{\omega^2}, \\
\nu_6(\omega, \mathcal{T}) \equiv& \frac{e^{-\omega \mathcal{T}_1} - e^{-\omega \mathcal{T}_2}}{\omega},
\end{aligned}$$

where $\mathcal{T}$ is a scalar for $i = 0, 1, 2$, and a two-dimensional vector with components $(\mathcal{T}_1, \mathcal{T}_2)$ for $i = 3, 4, 5, 6$. We define the scalars $(L_1, L_2, \Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5, \Delta_6)$ by

$$\begin{aligned}
L_1 \equiv& \frac{r\alpha\bar{\alpha} f}{\alpha + \bar{\alpha}} \frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}, \\
L_2 \equiv& r\bar{\alpha}\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right) + (\gamma_1^R + \gamma_2^R)\bar{C} + \gamma_3^R \frac{b_0 - b_1\bar{C}}{b_2} - k_1\bar{q}_1 - k_2\bar{q}_2, \\
\Delta_1 \equiv& f\left[\eta\Sigma^3\eta' - \frac{(\eta\Sigma^2\eta')^2}{\eta\Sigma\eta'}\right] + k\left(\eta\Sigma^2 p_f'\right)^2, \\
\Delta_2 \equiv& f\left[\eta\Sigma^3 p_f' - \frac{\eta\Sigma^2\eta'\eta\Sigma^2 p_f'}{\eta\Sigma\eta'}\right] + k\eta\Sigma^2 p_f' p_f\Sigma^2 p_f', \\
\Delta_3 \equiv& f\left[p_f\Sigma^3 p_f' - \frac{(\eta\Sigma^2 p_f')^2}{\eta\Sigma\eta'}\right] + k\left(p_f\Sigma^2 p_f'\right)^2, \\
\Delta_4 \equiv& f\left[Tr(\Sigma^2) - \frac{\eta\Sigma^3\eta'}{\eta\Sigma\eta'}\right] + kp_f\Sigma^3 p_f', \\
\Delta_5 \equiv& \Delta_4 - \frac{\Delta_1}{\eta\Sigma\eta'}, \\
\Delta_6 \equiv& f\Delta_5 + k\Delta_3.
\end{aligned}$$

where $Tr(M)$ denotes the trace of the matrix M.

Proposition 5.1 *The Sharpe ratio of the momentum strategy (5.2), in which weights are constructed using index-adjusted past returns, is*

$$SR_{w^{\hat{M}}} = \frac{\left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \tau, \nu_1) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \tau, \nu_1) + L_2^2\tau\right] p_f \Sigma^2 p_f'}{\sqrt{\left[2G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \tau, \nu_2) + 2H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \tau, \nu_2) + L_2^2\tau^2\right] \Delta_3 + \tau\Delta_6}}. \quad (5.4)$$

Proposition 5.1 yields an intuitive decomposition of momentum profits. In the proposition's proof, we show that the numerator of the Sharpe ratio (5.4) can be written as

$$\frac{1}{dt}\int_{t-\tau}^{t} \left\{ E\left[Cov_u(d\hat{R}_u', d\hat{R}_t)\right] + Cov\left[E_u(d\hat{R}_u'), E_t(d\hat{R}_t)\right] + E(d\hat{R}_u')E(d\hat{R}_t) \right\}. \quad (5.5)$$

The three terms inside the curly bracket correspond to three distinct sources of momentum profits. The first term corresponds to profits earned because of the response of expected returns to shocks. As shown in Proposition 3.5, a shock to a stock's return moves the stock's short-run expected return in the same direction as the shock. Hence, a stock hit by a positive shock both receives higher weight in the momentum strategy and is expected to do better going forward, resulting in momentum profits.

The second term corresponds to profits earned because of time-variation in expected returns, shown in Proposition 3.3. Consider a stock whose expected return is temporarily high, i.e., its conditional expected return is higher than its unconditional average. Since such a stock is expected to perform well, its expected weight in the momentum strategy is higher than its unconditional average. Moreover, the stock is expected to do better than its unconditional average going forward, resulting in momentum profits.

The first two terms add up to the unconditional autocovariance of returns. Past returns can be informative about future expected returns because of the response of the latter to unexpected shocks: this is the first term, and is equal to the conditional autocovariance since shocks are the difference between returns and their conditional expectation. Past returns can also be informative because of their expected-return component, which is persistent: this is the second term, and is equal to the unconditional autocovariance of conditional expected returns.

The first two terms capture momentum profits earned from time-series variation in the expected return of each stock. Additional profits are earned because of cross-sectional variation in stocks' unconditional expected returns (index-adjusted). These correspond to the third term. Stocks with higher unconditional expected returns both receive higher weight in the momentum strategy on average and are expected to do better going forward, resulting in momentum profits.

The calibration in Vayanos and Woolley (2011) indicates that the dominant source of momentum profits is the first, i.e., the conditional autocovariance. This term is also key for the simultaneous emergence of short-run momentum and long-run reversal. Indeed, the second and third terms would imply that a momentum strategy is profitable at all lags.

Proposition 5.2 *The Sharpe ratio of the momentum strategy (5.1), in which weights are constructed using raw past returns can be derived from (5.4) by adding*

$$L_1 L_2 \tau \eta \Sigma^2 p_f' \tag{5.6}$$

to the numerator and

$$f\tau \frac{\Delta_1}{\eta \Sigma \eta'} + L_1^2 \tau^2 \Delta_1 + 2 L_1 L_2 \tau^2 \Delta_2 \tag{5.7}$$

to the term inside the square root in the denominator.

Using Propositions 5.1 and 5.2, we can examine whether a momentum strategy is best implemented using raw or index-adjusted past returns. Implementing the strategy using raw returns favors stocks whose cashflows covary highly with the market index. Indeed, the high covariance causes these stocks to have high expected (raw) return and hence high expected weight in the momentum strategy. Implementing instead the strategy using index-adjusted returns eliminates this effect because a stock's covariance with the index does not affect the stock's expected index-adjusted return.

Since trading strategies are evaluated based on their index-adjusted returns, the weight they give to a stock should depend on measures that predict the stock's own index-adjusted return. The latter return depends only on the stock's covariance with the flow portfolio. Hence, using raw returns to implement a momentum strategy can dominate using index-adjusted returns only if a stock's covariance with the index is informative about the stock's covariance with the flow portfolio. If instead the two covariances are unrelated, then raw returns only introduce noise relative to index-adjusted returns. Proposition 5.3 shows a result along these lines in the simple case where each stock has a "twin" that has the same weight in the market index η, and all pairs of twins have the same weight in the true market portfolio θ relative to η. Under these assumptions, differences in stocks' covariance with the flow portfolio arise within pairs of twins, but are unrelated to stocks' covariance with the market index.

Proposition 5.3 *Suppose that each stock n has a twin n' that has the same index weight ($\eta_n = \eta_{n'}$), cashflow covariance with each other stock ($\Sigma_{nm} = \Sigma_{n'm}$ for all $m \neq n, n'$), and cashflow variance ($\Sigma_{nn} = \Sigma_{n',n'}$). Suppose also that all pairs of twins have the same weight in θ relative to η ($\frac{\theta_n+\theta_{n'}}{\eta_n+\eta_{n'}}$ independent of n). If the momentum strategy (5.1) implemented using raw past returns yields a positive Sharpe ratio, then its counterpart (5.2) using index-adjusted past returns yields a higher Sharpe ratio, i.e.,*

$$SR_{w^{\hat{M}}} \geq SR_{w^M}. \tag{5.8}$$

Moreover, the inequality (5.8) is strict unless $\Sigma^{\frac{1}{2}}\eta'$ is an eigenvector of Σ.

6 Value Strategies

Value weights depend on the comparison between stock prices and expected dividends. To define and interpret these weights, we recall from (3.1) that prices are the present value of expected dividends, discounted at the riskless rate r, minus a risk discount. Lemma 6.1 shows that the risk discount is the present value of expected returns, discounted at r. Since from Proposition 3.3 expected returns consist of a compensation for bearing risk that correlates with the market index, and a compensation for bearing risk that correlates with the flow portfolio, a similar decomposition applies to the risk discount: it is the sum of discounts $(\Gamma_{1t}, \Gamma_{2t})$ arising because of stocks' correlation with the index and the flow portfolio, respectively.

Lemma 6.1 *Stock prices are*

$$S_t = \int_t^{\infty} \left[E_t(dD_{t'}) - E_t(dR_{t'})\right] e^{-r(t'-t)} = \frac{\bar{F}}{r} + \frac{F_t - \bar{F}}{r+\kappa} - \Gamma_{1t} - \Gamma_{2t}, \tag{6.1}$$

where

$$\Gamma_{1t} \equiv E_t\left[\int_t^{\infty} \frac{r\alpha\bar{\alpha}}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'} Cov_{t'}(dR_{t'}, \eta dR_{t'}) e^{-r(t'-t)}\right] = \frac{\alpha\bar{\alpha}f}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}\Sigma\eta', \tag{6.2}$$

$$\Gamma_{2t} \equiv E_t\left[\int_t^{\infty} \Lambda_{t'} Cov_{t'}(dR_{t'}, p_f dR_{t'}) e^{-r(t'-t)}\right] = \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\left[\int_t^{\infty} E\left(\Lambda_{t'}\right) e^{-r(t'-t)} dt'\right]\Sigma p_f', \tag{6.3}$$

are risk discounts arising because of stocks' correlation with the index and the flow portfolio, respectively.

We consider four implementations of a value strategy:

$$\left(w_t^V\right)' \equiv \frac{\bar{F}}{r} + \frac{\epsilon(F_t - \bar{F})}{r+\kappa} - S_t, \tag{6.4}$$

$$\left(w_t^{\hat{V}}\right)' \equiv \frac{\bar{F}}{r} + \frac{\epsilon(F_t - \bar{F})}{r+\kappa} - \hat{S}_t, \tag{6.5}$$

where $\epsilon \in \{0,1\}$ and

$$\hat{S}_t \equiv S_t + \Gamma_{1t}. \tag{6.6}$$

Under all implementations, a stock's value weight increases linearly in the difference between the present value of the stock's expected dividends, discounted at r, and the stock's price. The four implementations differ in the measures of expected dividends and price.

We consider two forecasts for expected dividends: an optimal forecast, which conditions on all information available at time t, and a crude forecast, which conditions only on the instantaneous dividend dD_t. The optimal forecast depends on the expected dividend rate F_t, which is a sufficient statistic for all other information. It is the forecast used by the manager in the model, enters in the equilibrium prices (3.1) and (6.1), and corresponds to $\epsilon = 1$. We also consider the crude forecast because F_t might not be observable to an investor outside the model. Because of the Brownian noise dB_t^D, observing dD_t yields no information on F_t. Hence, the crude forecast sets expected dividends equal to their unconditional mean $\bar{F}$, and corresponds to $\epsilon = 0$. The optimal and crude forecasts are polar cases; intermediate cases can be derived by setting ϵ to be strictly between zero and one in the propositions below.

We consider two price measures: the raw price in (6.4), and an index-adjusted price in (6.5). The latter is derived by adding back to the raw price the discount Γ_{1t} arising from a stock's correlation with the index. Besides simplifying the calculations, index-adjustment can raise the Sharpe ratio by providing a less noisy measure of flow-generated mispricing. Indeed, a stock's raw price can be low relative to the present value of expected dividends because the stock's cashflows covary highly with the index, or because of current or anticipated flows. Index-adjustment isolates the latter effect.

Using (6.1) and (6.6), we can express the value weights (6.4) and (6.5) in terms of the risk discounts $(\Gamma_{1t}, \Gamma_{2t})$:

$$\left(w_t^V\right)' \equiv \Gamma_{1t} + \Gamma_{2t} - \frac{(1-\epsilon)(F_t - \bar{F})}{r+\kappa}, \tag{6.7}$$

$$\left(w_t^{\hat{V}}\right)' \equiv \Gamma_{2t} - \frac{(1-\epsilon)(F_t - \bar{F})}{r+\kappa}. \tag{6.8}$$

When weights are computed using raw prices, they are equal to the sum of the discounts minus an error in forecasting expected dividends. Index-adjustment eliminates the risk discount Γ_{1t}. The forecast error is present only under the crude forecast ($\epsilon = 0$), and increases the weights of stocks corresponding to low components of the vector F_t. This is because the low expected dividends of those stocks are reflected in the price but not in the crude forecast, so the stocks appear cheap.

Proposition 6.1 *The Sharpe ratio of the value strategy (6.5), in which weights are constructed using index-adjusted prices, is*

$$SR_{w^{\hat{V}}} = \frac{\left[H(\gamma_1,\gamma_2,\gamma_3,\gamma_1^R,\gamma_2^R,\gamma_3^R,0,\nu_0) - \frac{1-\epsilon}{r+\kappa}K_1(\gamma_1^R,\gamma_3^R,0,\nu_0) + \frac{L_2^2}{r}\right] p_f \Sigma^2 p_f'}{\sqrt{\left[H(\gamma_1,\gamma_2,\gamma_3,\gamma_1,\gamma_2,\gamma_3,0,\nu_0) - \frac{2(1-\epsilon)}{r+\kappa}K_1(\gamma_1,\gamma_3,0,\nu_0) + \frac{L_2^2}{r^2}\right]\Delta_3 + \frac{(1-\epsilon)^2\phi^2}{2\kappa(r+\kappa)^2}\Delta_4}}. \tag{6.9}$$

Proposition 6.1 yields an intuitive decomposition of value profits. In the proposition's proof, we show that the numerator of the Sharpe ratio (6.9) can be written as

$$\frac{1}{dt}\left\{Cov\left[\Gamma_{2t}', E_t(d\hat{R}_t)\right] - Cov\left[\frac{(1-\epsilon)(F_t-\bar{F})'}{r+\kappa}, E_t(d\hat{R}_t)\right] + E(\Gamma_{2t}')E(d\hat{R}_t)\right\}. \tag{6.10}$$

The three terms inside the curly bracket correspond to three distinct sources of value profits. The first term corresponds to profits earned because of the covariance between a stock's risk discount Γ_{2t} and its expected return. Suppose that the risk discount is temporarily high because of flows, and so the stock receives high weight in the value strategy. The stock's expected return is temporarily high if the high discount is caused by current flows, and low if it is caused by anticipated future flows. Value profits are earned if the first scenario dominates. It does for small s, as we show in the proposition's proof, and our numerical solutions indicate the same for general values of s.

The second term is non-zero only under the crude forecast, and corresponds to profits earned because of the covariance between a stock's expected dividends F_t and its expected return. Suppose that expected dividends are temporarily low, and so the stock receives high weight in the value strategy because of the error under the crude forecast. The negative shock that caused the low expected dividends impacts expected return through the amplifying fund flows that it triggers. Current flows raise expected return, while anticipated future flows lower it. Value profits are earned if the first scenario dominates. We show that it does, as in the case of the first term.

The first two terms capture value profits earned from time-series variation in the expected return of each stock. Additional profits are earned because of cross-sectional variation in stocks' unconditional expected returns (index-adjusted). These correspond to the third term. Stocks with higher unconditional expected returns receive higher weight in the value strategy on average because of a high discount Γ_{2t}. They are also expected to do better going forward, resulting in value profits.

Proposition 6.2 *The Sharpe ratio of the value strategy (6.4), in which weights are constructed using raw prices can be derived from (6.9) by adding*

$$\frac{L_1 L_2}{r}\eta\Sigma^2 p_f' \tag{6.11}$$

to the numerator and

$$\frac{L_1^2}{r^2}\Delta_1 + 2\frac{L_1 L_2}{r^2}\Delta_2 \tag{6.12}$$

to the term inside the square root in the denominator.

Using Propositions 6.1 and 6.2, we can examine whether a value strategy is best implemented using raw or index-adjusted prices. Implementing the strategy using raw prices favors stocks whose cashflows covary highly with the market index. Indeed, the high covariance causes these stocks to have high discounts Γ_{1t} and hence high weight in the value strategy. Implementing instead the strategy using index-adjusted prices eliminates this effect. As in the case of momentum strategies, using index-adjusted prices is better if stocks' covariances with the index are unrelated to their covariances with the flow portfolio.

Proposition 6.3 *Suppose that each stock n has a twin n' that has the same index weight ($\eta_n = \eta_{n'}$), cashflow covariance with each other stock ($\Sigma_{nm} = \Sigma_{n'm}$ for all $m \neq n, n'$), and cashflow variance ($\Sigma_{nn} = \Sigma_{n',n'}$). Suppose also that all pairs of twins have the same weight in θ relative to η ($\frac{\theta_n+\theta_{n'}}{\eta_n+\eta_{n'}}$ independent of n). If the value strategy (6.4) implemented using raw prices yields a positive Sharpe* *ratio, then its counterpart (6.5) using index-adjusted prices yields a higher Sharpe ratio, i.e.,*

$$SR_{w^{\hat{V}}} \geq SR_{w^V}. \tag{6.13}$$

Moreover, the inequality (6.13) is strict unless $\Sigma^{\frac{1}{2}}\eta'$ *is an eigenvector of* Σ.

7 Combining Momentum and Value

In this section we compute the covariance between momentum and value strategies. This yields additional empirical predictions for the model. It also allows us to further explore the model's implications for portfolio management by determining how to best combine momentum and value strategies.

To determine the optimal portfolio of momentum and value, we extend the mean-variance optimization of Section 4. Instead of assuming that the investor has access to only one strategy w_t in addition to the market index, we assume that there are two strategies (w_t^A, w_t^B). The investor chooses investments $\hat{x}$ in the index and $(\hat{y}^A, \hat{y}^B)$ in the strategies to maximize (4.4) subject to the budget constraint

$$\begin{aligned} dW_t &= rW_t dt + \hat{x}\eta dR_t + \sum_{i=A,B} \hat{y}^i w_t^i dR_t \\ &= rW_t dt + \hat{\hat{x}}\eta dR_t + \sum_{i=A,B} \hat{y}^i \hat{w}_t^i dR_t, \end{aligned} \tag{7.1}$$

where

$$\hat{\hat{x}} \equiv \hat{x} + \sum_{i=A,B} \hat{y}^i \frac{Cov_t(w_t^i dR_t, \eta dR_t)}{Var_t(\eta dR_t)}$$

is the investor's overall exposure to the index. The maximization problem is solved in the Appendix (Lemma E.1). The Sharpe ratio of the optimal portfolio of (w_t^A, w_t^B) is

$$SR_{w^{AB}} \equiv \sqrt{\frac{SR^2_{w^A} + SR^2_{w^B} - 2SR_{w^A}SR_{w^B}Corr(w^A, w^B)}{1 - Corr(w^A, w^B)^2}}, \tag{7.2}$$

where

$$Corr(w^A, w^B) \equiv \frac{Cov\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right)}{\sqrt{Var(\hat{w}_t^A dR_t) Var(\hat{w}_t^B dR_t)}}$$

denotes the correlation between the index-adjusted versions $(\hat{w}_t^A, \hat{w}_t^B)$ of the strategies (w_t^A, w_t^B). For simplicity, by correlation and covariance between (w_t^A, w_t^B), we mean from now on the correlation and covariance between their index-adjusted versions.

Proposition 7.1 *The covariance between the momentum strategy (5.2) and the value strategy (6.5), which use index-adjusted returns and prices, respectively, is*

$$Cov(\hat{w}_t^{\hat{M}} dR_t, \hat{w}_t^{\hat{V}} dR_t) = \Bigg\{ \Bigg[G(\gamma_1, \gamma_2, \gamma_3, \tau, \nu_1) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1, \gamma_2, \gamma_3, \tau, \nu_1) - \frac{1-\epsilon}{r+\kappa} K_1(\gamma_1^R, \gamma_3^R, \tau, \nu_1)$$

$$+ \frac{L_2^2}{r}\tau \Bigg] \Delta_3 - \frac{(1-\epsilon)\phi^2}{(r+\kappa)^2} \nu_1(\kappa, \tau)(\Delta_5 + \beta_2 \gamma_1 \Delta_3) \Bigg\} dt. \tag{7.3}$$

The covariance between two strategies (w_t^A, w_t^B) is high if they both give high weight to the same stocks or to stocks that covary positively. In the proposition's proof we show that this covariance is

$$Cov\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right) = \left\{ Cov\left[\hat{w}_t^A, (f\Sigma + k\Sigma p_f' p_f \Sigma) \hat{w}_t^{B'}\right] + E\left(\hat{w}_t^A\right)(f\Sigma + k\Sigma p_f' p_f \Sigma) E\left(\hat{w}_t^{B'}\right) \right\} dt. \tag{7.4}$$

The first term in the curly bracket is the covariance generated by temporal variation in the weights that the strategies give to stocks. This term involves the covariance between weights, adjusted by the covariance matrix $(f\Sigma + k\Sigma p_f' p_f \Sigma)dt$ of stock returns. Temporal variation in weights generates negative covariance between momentum and value strategies—both for small s as we show in the proposition's proof, and for general values of s as our numerical solutions indicate. Intuitively, a stock hit by a positive shock receives temporarily high weight in the momentum strategy. It also receives temporarily low weight in the value strategy because its risk discount is reduced by the amplifying fund flows that the shock triggers. Moreover, if the value strategy uses the crude forecast for dividends, there is an additional effect lowering the stock's weight: a positive shock to expected dividends is reflected in the stock's price but not in the strategy's forecast for dividends. The negative covariance between stocks' momentum and value weights generates negative covariance between momentum and value strategies, i.e., between these strategies' returns.

The second term is the covariance generated by the average weights that the strategies give to stocks. These weights are identical for value and momentum strategies. Intuitively, stocks with high unconditional expected returns receive high average weight in the momentum strategy because their past returns are high on average. They also receive high average weight in the value strategy because their risk discount (which is the present value of expected returns) is high. This effect generates positive covariance between momentum and value strategies. It corresponds to the last term in the square bracket in (7.3), while the effect of temporal variation corresponds to the other terms.

The covariance between momentum and value strategies is negative if the effect of temporal variation in weights dominates that of average weights. This is the case when the time-variation in the cost C_t (measured by the diffusion coefficient s), which determines fund flows, is large relative to the long-run mean $\bar{C}$ of C_t, which influences unconditional expected returns.

A negative covariance between momentum and value strategies implies large diversification benefits from combining these strategies. As (7.2) confirms, negative covariance raises the Sharpe ratio of the optimal portfolio, holding the Sharpe ratios of the two strategies constant (and assuming that they are positive).

When momentum and value strategies are constructed using raw returns and prices, the effect of average weights is larger because there is more cross-sectional variation in unconditional expected returns. This can raise the covariance.

Proposition 7.2 *The covariance between the momentum strategy (5.1) and the value strategy (6.4), which use raw returns and prices, respectively, can be derived from (7.3) by adding*

$$\left[-\frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}\nu_1(\kappa,\tau)\frac{\Delta_1}{\eta\Sigma\eta'}+\frac{L_1^2}{r}\tau\Delta_1+\frac{2L_1L_2}{r}\tau\Delta_2\right]dt. \tag{7.5}$$

8 Momentum and Value over Long Horizons

Sections 4-7 assume that a strategy's performance is evaluated over an infinitesimal investment horizon. This section considers instead a general non-infinitesimal horizon. Denoting the horizon by T, the Sharpe ratio corresponding to a strategy w_t is

$$SR_{w,T}\equiv\frac{E\left(\int_t^{t+T}\hat{w}_u dR_u\right)}{\sqrt{Var\left(\int_t^{t+T}\hat{w}_u dR_u\right)T}}, \tag{8.1}$$

expressed in annualized terms. We motivate the use of the Sharpe ratio based on portfolio optimization in the Appendix, but under stronger assumptions than in the case of an infinitesimal horizon.

The expected return in (8.1) is

$$E\left(\int_t^{t+T} \hat{w}_u dR_u\right) = \int_t^{t+T} E(\hat{w}_u dR_u) = \frac{TE(\hat{w}_t dR_t)}{dt} \tag{8.2}$$

and the variance is

$$\begin{aligned} Var\left(\int_t^{t+T} \hat{w}_u dR_u\right) =& \int_t^{t+T} Var(\hat{w}_u dR_u) + 2\int_{t\le u'<u\le t+T} Cov(\hat{w}_{u'} dR_{u'}, \hat{w}_u dR_u) \\ =& \frac{TVar(\hat{w}_t dR_t)}{dt} + 2\int_{t<u\le t+T} (t+T-u)\frac{Cov(\hat{w}_t dR_t, \hat{w}_u dR_u)}{dt}, \end{aligned} \tag{8.3}$$

where the second step in each case follows because unconditional moments are time-invariant in the steady state. The expected return is the sum of instantaneous expected returns. The variance differs from the sum of instantaneous variances because it includes the autocovariance of returns. If the autocovariance is positive, then the variance exceeds the sum of instantaneous variances, and vice-versa. Because of the autocovariance, the Sharpe ratio can depend on the investment horizon. Eqs. (4.2) and (8.1)-(8.3) imply that the Sharpe ratio SR_w^T for investment horizon T is linked to its infinitesimal-horizon counterpart SR_w through

$$SR_w^T = \frac{SR_w}{\sqrt{1 + 2\int_{t<u\le t+T}\left(1 - \frac{u-t}{T}\right)\frac{Cov(\hat{w}_t dR_t, \hat{w}_u dR_u)}{Var(\hat{w}_t dR_t)}}}. \tag{8.4}$$

Therefore, SR_w^T exceeds SR_w if the autocovariance is positive, and vice-versa.

Lemma 8.1 computes the autocovariance of returns for linear strategies. Linear strategies have weights that are integrals of the Brownian shocks with constant coefficients. The momentum strategies (5.1) and (5.2), and the value strategies (6.4) and (6.5) are linear.

Lemma 8.1 *The covariance between the return of a linear strategy w_t at time t and that at time $t' > t$ is*

$$Cov(\hat{w}_t dR_t, \hat{w}_{t'} dR_{t'}) = C_1 + C_2, \tag{8.5}$$

where

$$\begin{aligned} C_1 \equiv& E\left[Cov_t\left(\hat{w}_t dR_t, \hat{w}_{t'} dR_{t'}\right)\right] \\ =& \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(\hat{w}_t\Lambda_{t'}\right) Cov_t(dR_t, w_{t'}\Sigma p_f') + E\left(\hat{w}_t w_{t'}\Sigma p_f'\right) Cov_t(dR_t, \Lambda_{t'})\right] dt', \end{aligned} \tag{8.6}$$

and

$$
\begin{aligned}
C_2 \equiv & Cov\left[E_t(\hat{w}_t dR_t), E_u(\hat{w}_{t'} dR_{t'})\right] \\
= & \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)^2 \Big[E(\Lambda_t)^2 Cov(w_t\Sigma p_f', w_{t'}\Sigma p_f') + E(w_t\Sigma p_f')^2 Cov(\Lambda_t, \Lambda_{t'}) \\
& + E(w_t\Sigma p_f')E(\Lambda_t)\left[Cov(\Lambda_t, w_{t'}\Sigma p_f') + Cov(w_t\Sigma p_f', \Lambda_{t'})\right] \\
& + Cov(\Lambda_t, \Lambda_{t'})Cov(w_t\Sigma p_f', w_{t'}\Sigma p_f') + Cov(\Lambda_t, w_{t'}\Sigma p_f')Cov(w_t\Sigma p_f', \Lambda_{t'})\Big] dtdt'. \qquad (8.7)
\end{aligned}
$$

The autocovariance of returns is the sum of two terms, given by (8.6) and (8.7). The first term characterizes how expected returns respond to shocks, and is positive if they increase following positive shocks. The second term is the autocovariance of expected returns, and is positive because expected returns are persistent. The calculation of each term is complicated by the fact that returns concern trading strategies rather than individual stocks. This is because the autocovariance of strategy returns is affected not only by the time-variation of stock expected returns but also by that of strategy weights. For example, if weights change rapidly over time, as is the case for momentum strategies that employ a short window of past returns, then the autocovariance of strategy returns is close to zero.

According to the two-factor model of Proposition 3.3, the expected index-adjusted return of a trading strategy w_t at time t is the product of the strategy's loading $(f + k\Delta/(\eta\Sigma\eta'))w_t\Sigma p_f'$ on the flow portfolio, times the premium Λ_t associated to that factor. Time-variation in $w_t\Sigma p_f'$ is driven by that in strategy weights, and time-variation in Λ_t reflects that in stock expected returns. Both sources of variation are relevant for the effects captured in (8.6) and (8.7). Eq. (8.6) shows that the response of the expected return to shocks is a sum of a term characterizing the response of $w_t\Sigma p_f'$ and a term characterizing the response of Λ_t. Each term is the product of a covariance, which characterizes the response, and an expectation. For example, a positive response of $w_t\Sigma p_f'$ translates to a positive response of the expected return if on average Λ_t is positive (first term inside the square bracket in (8.6)). Eq. (8.7) shows that the autocovariance of expected returns is a sum of terms that involve the autocovariance of Λ_t, that of $w_t\Sigma p_f'$, and their cross-autocovariances. For example, a positive autocovariance of $w_t\Sigma p_f'$ translates to a positive autocovariance of the expected return (first term inside the square bracket in (8.7)), and especially so if the autocovariance of Λ_t is positive (fifth term).

The linearity of trading strategies ensures that factor loadings are normally distributed. Since, in addition, Λ_t is normally distributed, expected returns of trading strategies are products of normals. Autocovariances involving those products can be written as sums of terms involving expectations and autocovariances of the products' components, which are normal. Lemma 8.2 determines expectations and autocovariances that involve Λ_t and are common to all linear strategies.

Lemma 8.2 *The risk premium Λ_t has the following properties:*

$$E(\Lambda_t) = \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} L_2, \tag{8.8}$$

$$Cov(\Lambda_t, \Lambda_{t'}) = \frac{1}{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)^2} H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, t' - t, \nu_0), \tag{8.9}$$

$$E\left(\hat{w}_t w_{t'} \Sigma p'_f\right) Cov_t(dR_t, \Lambda_{t'}) = \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} \left[E(w_t \Sigma p'_f)^2 + Cov(w_t \Sigma p'_f, w_{t'} \Sigma p'_f)\right] \Lambda_{R,t'-t} dt, \tag{8.10}$$

where

$$\Lambda_{R,t'-t} \equiv G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t' - t, \nu_0).$$

Propositions 8.1 and 8.2 compute the autocovariance of returns of the momentum strategy (5.2) and the value strategy (6.5), which use index-adjusted returns and prices. Combining Propositions 8.1 and 8.2 with (8.4), we can compute the Sharpe ratios of these strategies for a general investment horizon T.

Proposition 8.1 *The covariance between the return of the momentum strategy (5.2) at time t and* *that at time $t' > t$ can be derived from (8.6) and (8.7) by substituting (8.8)-(8.10),*

$$E(w_t\Sigma^i v') = L_2\tau p_f\Sigma^{i+1}v', \tag{8.11}$$

$$Cov(w_t\Sigma^i v', \Lambda_{t'}) = \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}}\left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_3) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_3)\right] p_f\Sigma^{i+1}v', \tag{8.12}$$

$$\begin{aligned} Cov(\Lambda_t, w_{t'}\Sigma p_f') = {} & \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}}\Big[H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_4) \\ & + \left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \tau + t - t', \nu_1) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \tau + t - t', \nu_1)\right] 1_{\{t'-t<\tau\}}\Big] p_f\Sigma^2 p_f', \end{aligned} \tag{8.13}$$

$$\begin{aligned} Cov(w_t\Sigma p_f', w_{t'}\Sigma p_f') = {} & \left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_5) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_5)\right] (p_f\Sigma^2 p_f')^2 \\ & + (\tau + t - t')\Delta_3 1_{\{t'-t<\tau\}} \end{aligned} \tag{8.14}$$

$$\begin{aligned} E\left(\hat{w}_t\Lambda_{t'}\right) Cov_t(dR_t, w_{t'}\Sigma p_f') = {} & \left[E(w_t\Sigma p_f')E(\Lambda_t) + Cov(w_t\Sigma p_f', \Lambda_{t'})\right] w_{MR1,\mathcal{T}}dt \\ & + \left[\left(E(w_t\Sigma^2 p_f') - \frac{E(w_t\Sigma\eta')}{\eta\Sigma\eta'}\eta\Sigma^2 p_f'\right) E(\Lambda_t) + Cov\left(w_t\Sigma^2 p_f' - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\Sigma^2 p_f', \Lambda_{t'}\right)\right] w_{MR2,\mathcal{T}}dt, \end{aligned} \tag{8.15}$$

where $i = 1, 2$, $v \in \{p_f, \eta\}$, $\mathcal{T} \equiv (t' - t, \tau)$, $1_{\mathcal{S}}$ is equal to one if condition $\mathcal{S}$ is satisfied and zero otherwise,

$$w_{MR1,\mathcal{T}} \equiv \left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, \mathcal{T}, \nu_4) + k1_{\{t'-t<\tau\}}\right] p_f\Sigma^2 p_f',$$

$$w_{MR2,\mathcal{T}} \equiv f1_{\{t'-t<\tau\}}.$$

Proposition 8.2 *The covariance between the return of the value strategy (6.5) at time t and that* *at time $t' > t$ can be derived from (8.6) and (8.7) by substituting (8.8)-(8.10),*

$$E(w_t\Sigma^i v') = \frac{L_2}{r} p_f \Sigma^{i+1} v', \tag{8.16}$$

$$Cov(w_t\Sigma^i v', \Lambda_{t'}) = \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} \left[H(\gamma_1,\gamma_2,\gamma_3,\gamma_1^R,\gamma_2^R,\gamma_3^R,t'-t,\nu_0) - \frac{1-\epsilon}{r+\kappa}K_2(\gamma_1^R,\gamma_3^R,t'-t,\nu_0)\right] p_f\Sigma^{i+1}v', \tag{8.17}$$

$$Cov(\Lambda_t, w_{t'}\Sigma p_f') = \frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} \left[H(\gamma_1^R,\gamma_2^R,\gamma_3^R,\gamma_1,\gamma_2,\gamma_3,t'-t,\nu_0) - \frac{1-\epsilon}{r+\kappa}K_1(\gamma_1^R,\gamma_3^R,t'-t,\nu_0)\right] p_f\Sigma^2 p_f', \tag{8.18}$$

$$\begin{aligned} Cov(w_t\Sigma p_f', w_{t'}\Sigma p_f') = &\Big[H(\gamma_1,\gamma_2,\gamma_3,\gamma_1,\gamma_2,\gamma_3,t'-t,\nu_0) - \frac{1-\epsilon}{r+\kappa}K_1(\gamma_1,\gamma_3,t'-t,\nu_0) \\ &-\frac{1-\epsilon}{r+\kappa}K_2(\gamma_1,\gamma_3,t'-t,\nu_0)\Big]\left(p_f\Sigma^2p_f'\right)^2 + \frac{(1-\epsilon)^2\phi^2}{2\kappa(r+\kappa)^2}\nu_0(\kappa,t'-t)p_f\Sigma^3p_f', \end{aligned} \tag{8.19}$$

$$\begin{aligned} E\left(\hat{w}_t\Lambda_{t'}\right)Cov_t(dR_t, w_{t'}\Sigma p_f') = &\left[E(w_t\Sigma p_f')E(\Lambda_t) + Cov(w_t\Sigma p_f', \Lambda_{t'})\right] w_{VR1,t'-t}dt \\ + &\left[\left(E(w_t\Sigma^2p_f') - \frac{E(w_t\Sigma\eta')}{\eta\Sigma\eta'}\eta\Sigma^2p_f'\right)E(\Lambda_t) + Cov\left(w_t\Sigma^2p_f' - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\Sigma^2p_f', \Lambda_{t'}\right)\right] w_{VR2,t'-t}dt, \end{aligned} \tag{8.20}$$

where $i = 1, 2$, $v \in \{p_f, \eta\}$,

$$w_{VR1,t'-t} \equiv \left[G(\gamma_1,\gamma_2,\gamma_3,t'-t,\nu_0) - \frac{(1-\epsilon)\phi^2\beta_2\gamma_1}{(r+\kappa)^2}\nu_0(\kappa,t'-t)\right] p_f\Sigma^2p_f',$$

$$w_{VR2,t'-t} \equiv -\frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}\nu_0(\kappa,t'-t).$$

9 Calibration

9.1 Model Parameters

We set the riskless rate r to 4%. We assume that there are $N = 10$ stocks, which we interpret as industry sectors. We assume that the market index η includes one share of each stock, i.e., $\eta = \mathbf{1}$, where $\mathbf{1} \equiv (1, .., 1)$, and that the true market portfolio includes one share of each stock on average, i.e., $\bar{\theta} \equiv \sum_{n=1}^{N} \theta_n/N = 1$. These are normalizations because we can redefine one share of each stock and of the index, leaving Sharpe ratios unchanged. We assume that stocks are symmetric in the sense that they all have the same standard deviation of dividends and the same pairwise correlations. (Our closed-form solutions for Sharpe ratios, however, do not require any symmetry.) Hence, the covariance matrix of dividends is $\Sigma = \hat{\sigma}^2(I + \omega \mathbf{1}'\mathbf{1})$, where I is the identity matrix and $(\hat{\sigma}, \omega)$ are scalars. We calibrate $\hat{\sigma}$ using the Sharpe ratio SR_η of the market index η. Closed-form solutions for SR_η and for all other quantities used in the calibration are in the Appendix. We express SR_η in annualized terms, and set it to 30%. This is equal to the Sharpe ratio of the S&P500 index, assuming an annual expected excess return of 4.5% and a standard deviation of 15%. The implied value of $\hat{\sigma}$ is 0.22.[4] We calibrate ω using the correlation between industry sectors and the market. Ang and Chen (2002) find that the average correlation between the returns of an industry sector and of a broad market index is 87% across the 13 sectors that they consider. The implied value of ω is seven. We set ϕ to 0.3. This parameter determines the size of shocks to the expected dividend rate F_t relative to dividends D_t, and has small effects on our calibration results.

VW show that the only characteristic of the true market portfolio θ that affects Sharpe ratios when stocks are symmetric is $\sigma(\theta) \equiv \sqrt{\sum_{n=1}^{N}(\theta_n - \bar{\theta})^2}$. This is the standard deviation across stocks of the number of shares included in θ, and must be strictly positive so that θ differs from the market index η. We calibrate $\sigma(\theta)$ using the average deviation between the weight that an active fund gives to an industry sector and the sector's weight in a broad market index. Kacperczyk, Sialm, and Zheng (KSZ1 2005) find that the sum of squared deviations across the ten sectors that they consider is 4.36% for the median fund, implying an average deviation of 6.6% ($10 \times 6.6\%^2 = 4.36\%$). To map this into a value for $\sigma(\theta)$, we adjust for the fact that θ is the sum of active- and index-fund holdings. The holdings of active funds are about ten times those of index funds in KSZ1's sample period, so the average deviation for a combined active and index fund (which is what θ represents) is 6%. The implied value of $\sigma(\theta)$ is 0.6.

To calibrate the diffusion coefficient s of the cost C_t, we recall the cost's interpretation as minus the return gap. Kacperczyk, Sialm, and Zheng (KSZ 2008) find that the top decile of mutual funds in terms of lagged one-year return gap earn a monthly CAPM alpha of 0.273%, while the bottom decile earn -0.431%.[5] Since in our model there is only one active fund, we interpret the differential between deciles in a time-series rather than a cross-sectional sense. The implied value of s is 1.6. We set the persistence parameter κ of the cost to 0.3. This is consistent with KSZ's finding that shocks

[4]Note that $\hat{\sigma}$ is a volatility per share rather than per dollar because this is how returns are expressed in our model. We calibrate using Sharpe ratios because these are comparable for per-share and per-dollar returns.

[5]KSZ derive two different sets of estimates; we focus on those derived using a back-testing procedure that reduces estimation noise.

to the return gap shrink to about one-third of their size within four years ($\log(3)/0.3 = 3.7$). We set the long-run mean $\bar{C}$ of the cost to zero, consistent with KSZ's finding that the average return gap in the cross-section is zero. With $\bar{C} = 0$, negative values of the cost are equally likely as positive values, which means that the cost cannot be interpreted solely as a managerial perk or operational cost. Hence, we emphasize again the managerial-ability interpretation, and for consistency set the parameter λ to zero.

We calibrate the adjustment-cost parameter ψ using the empirical response of fund flows to performance. Coval and Stafford (2007) find that a positive shock to a fund's return generates flows into the fund in each of the next four quarters, with the effect dying off in the fifth. We set $\psi = 1.2$, which ensures that following a positive shock to the active fund's return, the investor's holdings y_t in the fund increase in the next four quarters and start decreasing afterwards.

We set the investor's coefficient of absolute risk aversion α to one. This is a normalization because we can redefine the units of the consumption good, leaving Sharpe ratios unchanged. To calibrate the risk aversion of the manager, we recall that he can be interpreted as an aggregate of all "smart-money" agents with the expertise to exploit mispricings. We are interested in the capital that these experts own, rather than in the capital they might manage on behalf of outsiders, since only the former can be used to exploit mispricings generated by outsiders' flows. Since most of the financial expertise lies within the financial industry, the capital of experts can be linked to that industry's GDP share. Philippon (2008) reports that the GDP share of the Finance and Insurance industry was 5.5% on average during 1960-2007 in the US. We view this as an upper bound since only part of that industry concerns asset markets, and set the manager's coefficient of absolute risk aversion $\bar{\alpha}$ to 30. This means that the manager accounts for 3.2% (=1/(30+1)) of aggregate risk tolerance.[6]

As an independent check for our choices of s and $\bar{\alpha}$, we compute two additional quantities: the turnover and the return variance generated by fund flows. Lou (2011) finds that the standard deviation of a stock's quarterly turnover generated by fund flows is 0.7%. Since the funds in Lou's sample account for about 10% of market capitalization, the standard deviation of flow-generated volume is 7% of assets managed by these funds; we find 7.8%. Greenwood and Thesmar (GT 2001) find that fund flows explain 8% of stock return variance; we find 16%. GT's sample, however, includes less than half of all professionally-managed wealth. Accounting for that, and for possible measurement noise, is likely to produce a number even larger than 16%. Raising s and $\bar{\alpha}$ to match

[6]Risk tolerance in our model is independent of capital because of exponential utility. Our choice of $\bar{\alpha}$ is based on the notion that risk tolerance is proportional to capital, which is true under power utility.

such a number would raise the Sharpe ratios of momentum and value strategies that we find in the next section.

9.2 Results

We first compute Sharpe ratios over infinitesimal investment horizons. The maximum Sharpe ratio across all strategies (Proposition 4.2) is 61%. Since stocks are assumed symmetric, the momentum and value strategies (5.2) and (6.5), which use index-adjusted returns and prices, are equivalent to (5.1) and (6.4), which use raw returns and prices. We refer to them as the momentum and the value strategy, respectively.

Figure 1 plots the Sharpe ratio of the momentum strategy as a function of the length τ of the window over which past returns are calculated. This Sharpe ratio is positive for windows of less than three years, and then turns negative. Thus, a strategy based on short-run momentum is profitable, and so is one based on long-run reversal. The highest Sharpe ratio of momentum is achieved using a window of four months, and is 40%. Moreover, windows from one to 11 months yield Sharpe ratios larger than 30%, the ratio of the market index. The Sharpe ratio of momentum converges to zero as the window length goes to zero because very recent performance is a very noisy signal of future fund flows.

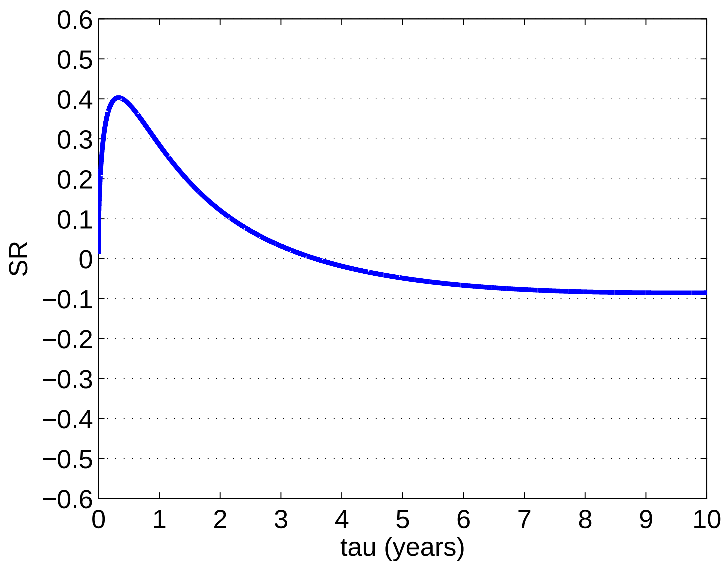

Figure 1: Sharpe ratio of the momentum strategy as a function of the length τ of the window over which past returns are calculated. The window length is measured in years, and the Sharpe ratio is expressed in annualized terms. The parameters for which the figure is drawn are described in Section 9.1.

Recall from Section 5 (Eq. (5.5)) that momentum profits can be decomposed into three sources. The first is the positive short-run response of expected returns to shocks: shocks hit by positive shocks receive high weight in the momentum strategy, and are expected to do well in the short run. The second is the time-series variation of expected returns (regardless of how these respond to shocks): stocks whose conditional expected returns are higher than their unconditional averages receive high expected weight in the momentum strategy, and are expected do well in the short run because conditional expected returns are persistent. The third is the cross-sectional variation of unconditional expected returns: stocks with high unconditional expected returns receive high expected weight in the momentum strategy, and are expected to do well going forward. The first source of profits is dominant in our calibration: for example, 62% of the maximum Sharpe ratio in Figure 1 is generated by the first source, 36% by the second, and 2% by the third.

The Sharpe ratio of the value strategy that uses the optimal forecast for expected dividends ($\epsilon = 1$) is 25.5%. Surprisingly, the value strategy that uses the crude forecast ($\epsilon = 0$) yields a slightly higher Sharpe ratio of 26%. Thus, using a crude forecast for expected dividends does not impair the Sharpe ratio of a value strategy, and can even enhance it. This is because the forecast error helps predict expected returns. Indeed, stocks whose expected dividends are temporarily low appear to be cheap under the crude forecast, and hence receive high weight in the value strategy. The negative shock that caused the low expected dividends also raises the expected return of these stocks through the amplifying fund flows that it triggers. This raises the value strategy's expected return, as pointed out in Section 6 (second term in (6.10)), and can raise the Sharpe ratio.

The correlation between the momentum and the value strategy is negative. For example, it is minus 3% between the momentum strategy that achieves the maximum Sharpe ratio in Figure 1 and the value strategy that uses the crude forecast ($\epsilon = 0$). Thus, combining momentum and value has significant diversification benefits. The Sharpe ratio of the optimal combination is 48%, significantly larger than that of the optimal momentum (40%) and the optimal value (26%). At the same time, it is significantly smaller than the maximum Sharpe ratio across all strategies, which is 61%. Thus, momentum and value strategies can be improved, possibly by using information on fund flows. Our model can yield predictions as to what type of fund-flow information can raise the Sharpe ratio.

We next turn to non-infinitesimal investment horizons. Figure 2 plots the autocorrelation of returns of the momentum strategy that achieves the maximum Sharpe ratio in Figure 1 and of the value strategy that uses the crude forecast ($\epsilon = 0$). The autocorrelation concerns instantaneous returns, which we assume are over intervals with the same infinitesimal length dt. We plot the autocorrelation as a function of the time lag between the intervals. We also divide by dt since the autocorrelation is of order dt.

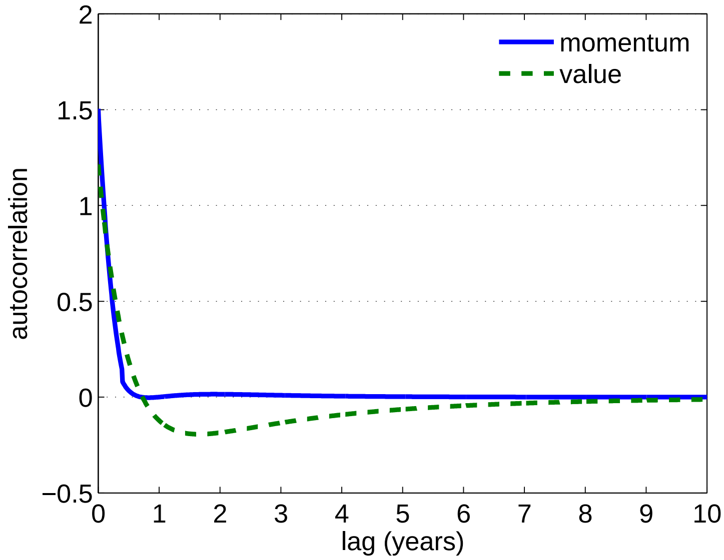

Figure 2: Return autocorrelations of the momentum and the value strategy. The momentum strategy uses a four-month window for past returns ($\tau = 4/12$), and the value strategy uses the crude forecast for expected dividends ($\epsilon = 0$). The autocorrelation concerns returns over intervals with the same infinitesimal length dt. It is plotted as a function of the time lag between the intervals, and is divided by dt. The parameters for which the figure is drawn are described in Section 9.1.

Figure 2 shows that the return autocorrelation is positive for both the momentum and the value strategy over short lags. Intuitively, since strategy weights exhibit some persistence, the short-run momentum in stock returns translates to short-run momentum in the returns of trading strategies. The momentum and the value strategy differ, however, in their return autocorrelation over longer lags. This autocorrelation is close to zero for the momentum strategy because momentum weights change rapidly. It is negative, however, for the value strategy. Indeed, since value weights change slowly, the the long-run reversal in stock returns translates to long-run reversal in the returns of the value strategy.

Figure 2 suggests that the long-horizon Sharpe ratios of the momentum and the value strategy can differ significantly from their short-horizon counterparts. Figure 3 plots the Sharpe ratios of the two strategies as a function of the investment horizon. We consider the momentum strategy that achieves the maximum Sharpe ratio in Figure 1 and the value strategy that uses the crude forecast ($\epsilon = 0$). The Sharpe ratios of both strategies decrease with the investment horizon when the horizon is small. This is because the returns of the strategies are positively autocorrelated over short lags, and hence the strategies' risk increases with the horizon. Because, however, the return autocorrelation of the momentum strategy dies off to zero quickly, the same is true for the decrease in the strategy's Sharpe ratio, with the ratio becoming essentially flat for horizons longer than one year. The Sharpe ratio of the value strategy, on the other hand, increases for horizons longer than one year, and eventually overtakes that of momentum. The increase is because the returns of the value strategy are negatively autocorrelated over long lags, and hence the strategy's risk decreases with the horizon when the horizon becomes sufficiently long.

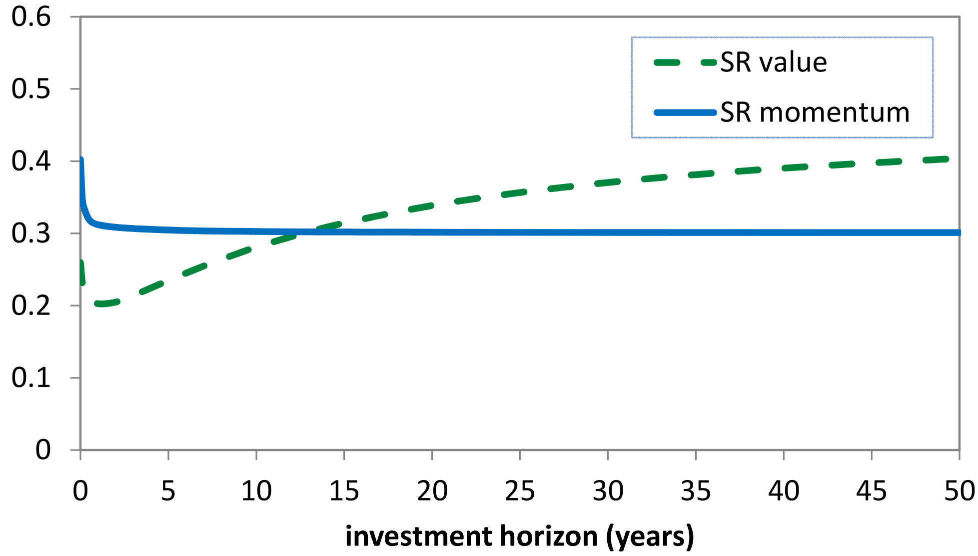

Figure 3: Sharpe ratios of the momentum and the value strategy as a function of the investment horizon. The momentum strategy uses a four-month window for past returns ($\tau = 4/12$), and the value strategy uses the crude forecast for expected dividends ($\epsilon = 0$). The investment horizon is measured in years, and the Sharpe ratios are expressed in annualized terms. The parameters for which the figure is drawn are described in Section 9.1.

Appendix

A Equilibrium

The proofs of Propositions 3.1-3.5 are in Vayanos and Woolley (VW 2011). Additional results of VW that we use in subsequent proofs are the properties of the flow portfolio

$$\eta\Sigma p_f' = 0,$$

$$\theta\Sigma p_f' = p_f\Sigma p_f' = \frac{\Delta}{\eta\Sigma\eta'},$$

and the characterizations of the investor's stock holdings

$$x_t\eta + y_t z_t = y_t p_f + \frac{\bar{\alpha}}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}, \tag{A.1}$$

of stock returns

$$\begin{aligned} dR_t = &\left\{ra_0 + \left[\gamma_1^R\hat{C}_t + \gamma_2^R C_t + \gamma_3^R y_t - \kappa(\gamma_1+\gamma_2)\bar{C} - b_0\gamma_3\right]\Sigma p_f'\right\}dt + \left(\sigma + \beta_1\gamma_1\Sigma p_f' p_f\sigma\right)dB_t^D \\ &+ \frac{\phi}{r+\kappa}\left(\sigma + \beta_2\gamma_1\Sigma p_f' p_f\sigma\right)dB_t^F - s\gamma_2\left(1 + \frac{\beta_2\gamma_1\Delta}{\eta\Sigma\eta'}\right)\Sigma p_f' dB_t^C, \end{aligned} \tag{A.2}$$

and of the dynamics of $\hat{C}_t$

$$d\hat{C}_t = \kappa(\bar{C}-\hat{C}_t)dt - \rho(C_t-\hat{C}_t)dt - \beta_1 p_f\sigma dB_t^D - \beta_2\left(\frac{\phi p_f\sigma dB_t^F}{r+\kappa} - \frac{s\gamma_2\Delta dB_t^C}{\eta\Sigma\eta'}\right). \tag{A.3}$$

B Evaluation of Trading Strategies

Proof of Proposition 4.1: Using (3.8), we can write (4.1) as

$$\hat{w}_t \equiv w_t - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta. \tag{B.1}$$

The expected return of the index-adjusted strategy is

$$
\begin{aligned}
E(\hat{w}_t dR_t) =& E\left[E_t(\hat{w}_t dR_t)\right] \\
=& E\left[\hat{w}_t E_t(dR_t)\right] \\
=& E\left\{\left(w_t - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\right)\left[\frac{r\alpha\bar{\alpha}}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}\left(f\Sigma + k\Sigma p_f' p_f\Sigma\right)\eta' + \Lambda_t\left(f\Sigma + k\Sigma p_f' p_f\Sigma\right)p_f'\right] dt\right\} \\
=& E\left\{\left(w_t - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\right)\left[\frac{r\alpha\bar{\alpha}}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'} f\Sigma\eta' + \Lambda_t\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\Sigma p_f'\right] dt\right\} \\
=& \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right) E\left(\Lambda_t w_t \Sigma p_f'\right) dt, \qquad \text{(B.2)}
\end{aligned}
$$

where the third step follows from (3.8), (3.9) and (B.1). Note that the second, third, fourth and fifth steps in the derivation of (B.2) imply that

$$
E_t(\hat{w}_t dR_t) = \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\Lambda_t w_t \Sigma p_f' dt. \qquad \text{(B.3)}
$$

The variance of the index-adjusted strategy is

$$
\begin{aligned}
Var(\hat{w}_t dR_t) =& E[Var_t(\hat{w}_t dR_t)] + Var[E_t(\hat{w}_t dR_t)] \\
=& E[Var_t(\hat{w}_t dR_t)] \\
=& E\left[\left(w_t - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\right)\left(f\Sigma + k\Sigma p_f' p_f \Sigma\right)\left(w_t - \frac{w_t\Sigma\eta'}{\eta\Sigma\eta'}\eta\right)'\right] dt \\
=& \left\{ f\left[E(w_t\Sigma w_t') - \frac{E\left[(w_t\Sigma\eta')^2\right]}{\eta\Sigma\eta'}\right] + kE\left[(w_t\Sigma p_f')^2\right]\right\} dt, \qquad \text{(B.4)}
\end{aligned}
$$

where the second step follows because $E[Var_t(w_t dR_t)]$ is of order dt and $Var[E_t(w_t dR_t)]$ of order $(dt)^2$, and the third step follows from (3.8) and (B.1). Eq. (4.3) follows from (4.2), (B.2) and (B.4). ∎

Proof of Lemma 4.1: Eqs. (4.7) and (4.8) follow from (4.6). Substituting (4.7) and (4.8) into the investor's utility, which is the term in curly brackets in (4.6), we find (4.9). ∎

Proof of Proposition 4.1: To show that the Sharpe ratio is maximized for $w_t = \Lambda_t p_f$, we write, for any given t, the strategy w_t as a linear combination of the market index, the flow portfolio, and an orthogonal component, i.e.,

$$
w_t = \lambda_{1t}\eta + \lambda_{2t} p_f + \check{w}_t, \qquad \text{(B.5)}
$$

where $\eta\Sigma\breve{w}_t = p_f\Sigma\breve{w}_t = 0$. Substituting w_t from (B.5), we can write (4.3) as

$$SR_w = \frac{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{\Delta}{\eta\Sigma\eta'}E(\Lambda_t\lambda_{2t})}{\sqrt{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{\Delta}{\eta\Sigma\eta'}E(\lambda_{2t}^2) + fE(\breve{w}_t\Sigma\breve{w}_t')}}. \tag{B.6}$$

The Sharpe ratio is maximized for $\breve{w}_t = 0$. Substituting into (B.6), we find

$$SR_w = \frac{E(\Lambda_t\lambda_{2t})}{\sqrt{E(\lambda_{2t}^2)}}\sqrt{\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{\Delta}{\eta\Sigma\eta'}}. \tag{B.7}$$

The Cauchy-Schwarz inequality implies that the term

$$\frac{E(\Lambda_t\lambda_{2t})}{\sqrt{E(\lambda_{2t}^2)}}$$

is maximized when λ_{2t} is proportional to Λ_t. Therefore, the Sharpe ratio is maximized by the strategy $\Lambda_t p_f$. Setting $w_t = \Lambda_t p_f$ in (4.3), we find the right-hand side of (4.10). ∎

C Momentum

We first prove four lemmas.

Lemma C.1 *The values of $(\hat{C}_t, C_t, y_t, F_t)$ in the steady state reached for $t \to \infty$ are*

$$\hat{C}_t = \bar{C} + \int_{-\infty}^t e^{-\kappa(t-u)}sdB_u^C - \int_{-\infty}^t e^{-(\kappa+\rho)(t-u)}\left[\beta_1 p_f\sigma dB_u^D + \frac{\phi\beta_2 p_f\sigma dB_u^F}{r+\kappa} + s\left(1 - \frac{\beta_2\gamma_2\Delta}{\eta\Sigma\eta'}\right)dB_u^C\right], \tag{C.1}$$

$$C_t = \bar{C} + \int_{-\infty}^t e^{-\kappa(t-u)}sdB_u^C, \tag{C.2}$$

$$y_t = \frac{b_0 - b_1\bar{C}}{b_2} + \int_{-\infty}^t \frac{b_1}{\kappa - b_2}\left[e^{-\kappa(t-u)} - e^{-b_2(t-u)}\right]sdB_u^C$$
$$- \int_{-\infty}^t \frac{b_1}{\kappa+\rho-b_2}\left[e^{-(\kappa+\rho)(t-u)} - e^{-b_2(t-u)}\right]\left[\beta_1 p_f\sigma dB_u^D + \frac{\phi\beta_2 p_f\sigma dB_u^F}{r+\kappa} + s\left(1 - \frac{\beta_2\gamma_2\Delta}{\eta\Sigma\eta'}\right)dB_u^C\right], \tag{C.3}$$

$$F_t = \bar{F} + \int_{-\infty}^t e^{-\kappa(t-u)}\phi\sigma dB_u^F, \tag{C.4}$$

Proof: The dynamics of $(\hat{C}_t, C_t, y_t)$ are (A.3), (2.3) and (3.4). Integrating this system with initial conditions $(\hat{C}_0, C_0, y_0)$, and letting $t \to \infty$, we find (C.1)-(C.3). Integrating (2.6) with initial condition F_0, and letting $t \to \infty$, we find (C.4). ∎

Lemma C.2 *The functions* (G, H, K_1, K_2) *satisfy*

$$Cov_t(dR_t, \psi_1 \hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}) = G(\psi_1, \psi_2, \psi_3, t' - t, \nu_0)\Sigma p'_f dt, \tag{C.5}$$

$$Cov\left(\psi_1 \hat{C}_t + \psi_2 C_t + \psi_3 y_t, \hat{\psi}_1 \hat{C}_{t'} + \hat{\psi}_2 C_{t'} + \hat{\psi}_3 y_{t'}\right) = H(\psi_1, \psi_2, \psi_3, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3, t' - t, \nu_0), \tag{C.6}$$

$$Cov\left(\psi_1 \hat{C}_t + \psi_2 C_t + \psi_3 y_t, F_{t'}\right) = K_1(\psi_1, \psi_3, t' - t, \nu_0)\Sigma p'_f, \tag{C.7}$$

$$Cov\left(F_t, \psi_1 \hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}\right) = K_2(\psi_1, \psi_3, t' - t, \nu_0)\Sigma p'_f, \tag{C.8}$$

where $t' > t$ *and*

$$\nu_0(\omega, t) \equiv e^{-\omega t}.$$

Proof: Since the covariance in (C.5) is conditional as of time t, it involves only the Brownian terms in dR_t, and not the drift terms (which are known at time t). Using (A.2) and (C.1)-(C.3), and noting that the only non-zero covariances are between Brownian increments of the same process and as of time t, we find

$$
\begin{aligned}
& Cov_t(dR_t, \psi_1 \hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}) \\
= & -\left(\sigma + \beta_1 \gamma_1 \Sigma p_f' p_f \sigma\right) \left[\psi_1 e^{-(\kappa+\rho)(t'-t)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-t)} - e^{-b_2(t'-t)}\right)\right] \beta_1 \sigma' p_f' dt \\
& - \frac{\phi}{r+\kappa}\left(\sigma + \beta_2 \gamma_1 \Sigma p_f' p_f \sigma\right) \left[\psi_1 e^{-(\kappa+\rho)(t'-t)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-t)} - e^{-b_2(t'-t)}\right)\right] \frac{\phi \beta_2}{r+\kappa} \sigma' p_f' dt \\
& - s\gamma_2 \left(1 + \frac{\beta_2 \gamma_1 \Delta}{\eta \Sigma \eta'}\right) \Sigma p_f' \left[(\psi_1+\psi_2) e^{-\kappa(t'-t)} + \frac{\psi_3 b_1}{\kappa - b_2}\left(e^{-\kappa(t'-t)} - e^{-b_2(t'-t)}\right)\right. \\
& - \left.\left[\psi_1 e^{-(\kappa+\rho)(t'-t)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-t)} - e^{-b_2(t'-t)}\right)\right] \left(1 - \frac{\beta_2 \gamma_2 \Delta}{\eta \Sigma \eta'}\right)\right] s dt \\
= & \left\{ - \left[\psi_1 e^{-(\kappa+\rho)(t'-t)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-t)} - e^{-b_2(t'-t)}\right)\right] \right. \\
& \times \left[\beta_1 \left(1 + \frac{\beta_1 \gamma_1 \Delta}{\eta \Sigma \eta'}\right) + \left(\frac{\phi^2 \beta_2}{(r+\kappa)^2} - s^2 \gamma_2 \left(1 - \frac{\beta_2 \gamma_2 \Delta}{\eta \Sigma \eta'}\right)\right)\left(1 + \frac{\beta_2 \gamma_1 \Delta}{\eta \Sigma \eta'}\right)\right] \\
& - \left.\left[(\psi_1+\psi_2) e^{-\kappa(t'-t)} + \frac{\psi_3 b_1}{\kappa - b_2}\left(e^{-\kappa(t'-t)} - e^{-b_2(t'-t)}\right)\right] s^2 \gamma_2 \left(1 + \frac{\beta_2 \gamma_1 \Delta}{\eta \Sigma \eta'}\right)\right\} \Sigma p_f' dt \\
= & \left\{ - \left[\psi_1 e^{-(\kappa+\rho)(t'-t)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-t)} - e^{-b_2(t'-t)}\right)\right] \beta_1 \left(1 + \frac{\beta_1 \gamma_1 \Delta}{\eta \Sigma \eta'}\right) \right. \\
& - \left.\left[(\psi_1+\psi_2) e^{-\kappa(t'-t)} + \frac{\psi_3 b_1}{\kappa - b_2}\left(e^{-\kappa(t'-t)} - e^{-b_2(t'-t)}\right)\right] s^2 \gamma_2 \left(1 + \frac{\beta_2 \gamma_1 \Delta}{\eta \Sigma \eta'}\right)\right\} \Sigma p_f' dt,
\end{aligned}
$$

where the third step follows from (3.7). This yields (C.5). Using (C.1)-(C.3) and noting that the only non-zero covariances are between Brownian increments of the same process and as of the same time $u \in (-\infty, t]$, we find

$$Cov\left(\psi_1\hat{C}_t + \psi_2 C_t + \psi_3 y_t, \hat{\psi}_1\hat{C}_{t'} + \hat{\psi}_2 C_{t'} + \hat{\psi}_3 y_{t'}\right)$$

$$= \int_{-\infty}^{t} \left[\psi_1 e^{-(\kappa+\rho)(t-u)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t-u)} - e^{-b_2(t-u)}\right)\right]$$

$$\times \left[\hat{\psi}_1 e^{-(\kappa+\rho)(t'-u)} + \frac{\hat{\psi}_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-u)} - e^{-b_2(t'-u)}\right)\right]\left[\beta_1^2 + \frac{\phi^2\beta_2^2}{(r+\kappa)^2}\right]\frac{\Delta}{\eta\Sigma\eta'}du$$

$$+ \int_{-\infty}^{u} \left[(\psi_1+\psi_2)e^{-\kappa(t-u)} + \frac{\psi_3 b_1}{\kappa-b_2}\left(e^{-\kappa(t-u)} - e^{-b_2(t-u)}\right)\right.$$

$$- \left[\psi_1 e^{-(\kappa+\rho)(t-u)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t-u)} - e^{-b_2(t-u)}\right)\right]\left(1 - \frac{\beta_2\gamma_2\Delta}{\eta\Sigma\eta'}\right)\Bigg]$$

$$\times \left[(\hat{\psi}_1+\hat{\psi}_2)e^{-\kappa(t'-u)} + \frac{\hat{\psi}_3 b_1}{\kappa-b_2}\left(e^{-\kappa(t'-u)} - e^{-b_2(t'-u)}\right)\right.$$

$$- \left[\hat{\psi}_1 e^{-(\kappa+\rho)(t'-u)} + \frac{\hat{\psi}_3 b_1}{\kappa+\rho-b_2}\left(e^{-(\kappa+\rho)(t'-u)} - e^{-b_2(t'-u)}\right)\right]\left(1 - \frac{\beta_2\gamma_2\Delta}{\eta\Sigma\eta'}\right)\Bigg] s^2 du.$$

Using (C.1)-(C.4), we similarly find

$$Cov\left(\psi_1\hat{C}_t + \psi_2 C_t + \psi_3 y_t, F_{t'}\right)$$

$$= -\int_{-\infty}^{t} \left[\psi_1 e^{-(\kappa+\rho)(t-u)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left[e^{-(\kappa+\rho)(t-u)} - e^{-b_2(t-u)}\right]\right] e^{-\kappa(t'-u)}\frac{\phi^2\beta_2\Sigma p_f'}{r+\kappa}du$$

and

$$Cov\left(F_t, \psi_1\hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}\right)$$

$$= -\int_{-\infty}^{t} e^{-\kappa(t-u)}\left[\psi_1 e^{-(\kappa+\rho)(t'-u)} + \frac{\psi_3 b_1}{\kappa+\rho-b_2}\left[e^{-(\kappa+\rho)(t'-u)} - e^{-b_2(t'-u)}\right]\right]\frac{\phi^2\beta_2\Sigma p_f'}{r+\kappa}du.$$

Integrating all products of exponentials, and summing, we find (C.6), (C.7) and (C.8).

The covariance $Cov_t(dR_t, dR_t')$ between stock returns at times t and t', derived in Proposition 3.5, can be expressed in terms of the function G. Since it involves only the drift terms in $dR_{t'}$ and not the Brownian terms (which have zero covariance with information up to time t'),

$$\begin{aligned} Cov_t(dR_t, dR_{t'}') =& Cov_t(dR_t, E_{t'}(dR_{t'}')) \\ =& Cov_t(dR_t, \gamma_1^R\hat{C}_{t'} + \gamma_2^R C_{t'} + \gamma_3^R y_{t'})p_f\Sigma dt' \\ =& G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t'-t, \nu_0)\Sigma p_f' p_f\Sigma dt dt', \end{aligned} \quad (C.9)$$

where the second step follows from (A.2) and the third from (C.5). ∎

Lemma C.3 *The expected dividend rate F_t has the following properties:*

$$Cov_t(dR_t, F'_{t'}) = \frac{\phi^2}{r+\kappa}\left(\Sigma + \beta_2\gamma_1\Sigma p'_f p_f \Sigma\right)\nu_0(\kappa, t'-t)dt, \tag{C.10}$$

$$Cov(F_t, F'_{t'}) = \frac{\phi^2\Sigma}{2\kappa}\nu_0(\kappa, t'-t). \tag{C.11}$$

Proof: Since the covariance in (C.10) is conditional as of time t, it involves only the Brownian terms in dR_t, and not the drift terms. Using (A.2) and (C.4), and noting that the only non-zero covariances are between the Brownian increments of the process F_t and as of time t, we find

$$Cov_t(dR_t, F'_{t'}) = \frac{\phi}{r+\kappa}\left(\sigma + \beta_2\gamma_1\Sigma p'_f p_f \sigma\right)\phi\sigma' e^{-\kappa(t'-t)}dt,$$

which yields (C.10). Using (C.4), and noting that the only non-zero covariances are between the Brownian increments of the process F_t and as of the same time $u \in (-\infty, t]$, we find

$$Cov(F_t, F'_{t'}) = \int_{-\infty}^{t} \phi^2\Sigma e^{-\kappa(t-u)}e^{-\kappa(t'-u)}du$$

Integrating, we find (C.11). ∎

Lemma C.4 *Index-adjusted returns have the following properties:*

$$E_t(d\hat{R}_t) = \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\Lambda_t\Sigma p'_f dt = \left[r\bar{\alpha}\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right) + \left(\gamma_1^R\hat{C}_t + \gamma_2^R C_t + \gamma_3^R y_t - k_1\bar{q}_1 - k_2\bar{q}_2\right)\right]\Sigma p'_f dt, \tag{C.12}$$

$$Cov_t(dR_t, d\hat{R}'_t) = Cov_t(d\hat{R}_t, d\hat{R}'_t) = \left[f\left(\Sigma - \frac{\Sigma\eta'\eta\Sigma}{\eta\Sigma\eta'}\right) + k\Sigma p'_f p_f\Sigma\right]dt, \tag{C.13}$$

$$Cov_t(d\hat{R}_t, \psi_1\hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}) = G(\psi_1, \psi_2, \psi_3, t'-t, \nu_0)\Sigma p'_f dt, \tag{C.14}$$

$$Cov_t(dR_t, d\hat{R}'_{t'}) = Cov_t(d\hat{R}_t, d\hat{R}'_{t'}) = G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t'-t, \nu_0)\Sigma p'_f p_f \Sigma dt dt', \tag{C.15}$$

$$Cov_t(d\hat{R}_t, F'_{t'}) = \frac{\phi^2}{r+\kappa}\left(\Sigma - \frac{\Sigma\eta'\eta\Sigma}{\eta\Sigma\eta'} + \beta_2\gamma_1\Sigma p'_f p_f\Sigma\right)\nu_0(\kappa, t'-t)dt, \tag{C.16}$$

where $t' > t$.

Proof: Eqs. (3.8) and (5.3) imply that

$$
\begin{aligned}
d\hat{R}_t =& dR_t - \frac{(f\Sigma + \Sigma p_f' p_f \Sigma)\eta'}{\eta(f\Sigma + \Sigma p_f' p_f \Sigma)\eta'} \eta dR_t \\
=& \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) dR_t,
\end{aligned} \tag{C.17}
$$

where I denotes the identity matrix. The first equality in (C.12) holds because

$$
\begin{aligned}
E_t(d\hat{R}_t) &= \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) E_t(dR_t) \\
&= \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) \left[\frac{r\alpha\bar{\alpha}}{\alpha + \bar{\alpha}} \frac{\eta\Sigma\theta'}{\eta\Sigma\eta'} (f\Sigma + \Sigma p_f' p_f \Sigma)\eta' + \Lambda_t (f\Sigma + \Sigma p_f' p_f \Sigma) p_f' \right] dt \\
&= \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) \left[L_1 \Sigma\eta' + \Lambda_t \left(f + \frac{k\Delta}{\eta\Sigma\eta'} \right) \Sigma p_f' \right] dt \\
&= \left(f + \frac{k\Delta}{\eta\Sigma\eta'} \right) \Lambda_t \Sigma p_f' dt,
\end{aligned} \tag{C.18}
$$

where the second step follows from (3.8) and (3.9). The second equality in (C.12) holds because of (3.10). Note that the second and third steps in the derivation of (C.18) imply that

$$
\begin{aligned}
E_t(dR_t) &= \left[L_1 \Sigma\eta' + \left(f + \frac{k\Delta}{\eta\Sigma\eta'} \right) \Lambda_t \Sigma p_f' \right] dt \\
&= \left[L_1 \Sigma\eta' + \left[r\bar{\alpha} \left(f + \frac{k\Delta}{\eta\Sigma\eta'} \right) + \left(\gamma_1^R \hat{C}_t + \gamma_2^R C_t + \gamma_3^R y_t - k_1 \bar{q}_1 - k_2 \bar{q}_2 \right) \right] \Sigma p_f' \right] dt.
\end{aligned} \tag{C.19}
$$

Eq. (C.13) holds because of (3.8),

$$
\begin{aligned}
Cov_t(dR_t, d\hat{R}_t') =& Cov_t(dR_t, dR_t') \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right)', \\
Cov_t(d\hat{R}_t, d\hat{R}_t') =& \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) Cov_t(dR_t, dR_t') \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right)'.
\end{aligned}
$$

Eq. (C.14) holds because of (C.5) and

$$
Cov_t(d\hat{R}_t, \psi_1 \hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}) = \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right) Cov_t(dR_t, \psi_1 \hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}).
$$

Eq. (C.15) holds because of (C.9),

$$
Cov_t(dR_t, d\hat{R}_{t'}') = Cov_t(dR_t, dR_{t'}') \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'} \right)',
$$

and

$$Cov_t(d\hat{R}_t, d\hat{R}'_{t'}) = \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'}\right) Cov_t(dR_t, dR'_{t'}) \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'}\right)'.$$

Eq. (C.16) holds because of (C.10) and

$$Cov_t(d\hat{R}_t, F'_{t'}) = \left(I - \frac{\Sigma\eta'\eta}{\eta\Sigma\eta'}\right) Cov_t(dR_t, F'_{t'}).$$

∎

Proof of Proposition 5.1: The numerator in (4.3) is equal to

$$\begin{aligned} &\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(\Lambda_t\right)E\left(w_t^{\hat{M}}\Sigma p'_f\right) + Cov\left(\Lambda_t, w_t^{\hat{M}}\Sigma p'_f\right)\right] \\ =&\left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\int_{t-\tau}^{t}\left[E\left(\Lambda_t\right)E\left(d\hat{R}'_u\Sigma p'_f\right) + Cov\left(\Lambda_t, d\hat{R}'_u\Sigma p'_f\right)\right], \end{aligned} \tag{C.20}$$

where the second step follows from (5.2). The term in square brackets inside the integral in (C.20) can be written as

$$\begin{aligned} &E\left(\Lambda_t\right)E\left(d\hat{R}'_u\Sigma p'_f\right) + E\left[Cov_u(\Lambda_t, d\hat{R}'_u\Sigma p'_f)\right] + Cov\left[E_u(\Lambda_t), E_u(d\hat{R}'_u)\Sigma p'_f\right] \\ =&E\left(\Lambda_t\right)E\left(d\hat{R}'_u\Sigma p'_f\right) + E\left[Cov_u(\Lambda_t, d\hat{R}'_u\Sigma p'_f)\right] + Cov\left[\Lambda_t, E_u(d\hat{R}'_u)\Sigma p'_f\right] \\ =&E\left(d\hat{R}'_u\right)E\left(\Lambda_t\Sigma p'_f\right) + E\left[Cov_u(d\hat{R}'_u, \Lambda_t\Sigma p'_f)\right] + Cov\left[E_u(d\hat{R}'_u), \Lambda_t\Sigma p'_f\right], \end{aligned} \tag{C.21}$$

where the second step follows because

$$\begin{aligned} Cov\left[E_u(X_t), Y_u\right] =&E\left[E_u(X_t)Y_u\right] - E\left[E_u(X_t)\right]E(Y_u) \\ =&E\left[E_u(X_tY_u)\right] - E(X_t)E(Y_u) \\ =&E(X_tY_u) - E(X_t)E(Y_u) \\ =&Cov(X_t, Y_u), \end{aligned} \tag{C.22}$$

for random variables (X_t, Y_u) that depend on information at time t and $u < t$, respectively. Using (C.12) and (C.21), we can write (C.20) as

$$\frac{1}{dt}\int_{t-\tau}^{t}\left\{E(d\hat{R}'_u)E\left[E_t(d\hat{R}_t)\right] + E\left[Cov_u\left(d\hat{R}'_u, E_t(d\hat{R}_t)\right)\right] + Cov\left[E_u(d\hat{R}'_u), E_t(d\hat{R}_t)\right]\right\}. \tag{C.23}$$

Eq. (C.23) coincides with (5.5) because the covariance $Cov_u(d\hat{R}'_u, d\hat{R}_t)$ involves only the drift terms in $d\hat{R}_t$ and not the Brownian terms. Eqs. (C.12) and (C.1)-(C.3) imply that

$$E(d\hat{R}_t) = E\left[E_t(d\hat{R}_t)\right] = L_2 \Sigma p'_f dt, \tag{C.24}$$

and hence

$$E(d\hat{R}'_u)E(d\hat{R}_t) = L_2^2 p_f \Sigma^2 p'_f dudt. \tag{C.25}$$

Eq. (C.15) implies that

$$\begin{aligned} E\left[Cov_u\left(d\hat{R}'_u, d\hat{R}_t\right)\right] =& E\left\{Tr\left[Cov_u\left(d\hat{R}_u, d\hat{R}'_t\right)\right]\right\} \\ =& G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t'-t, \nu_0) Tr\left(\Sigma p'_f p_f \Sigma\right) dudt \\ =& G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t'-t, \nu_0) p_f \Sigma^2 p'_f dudt, \end{aligned} \tag{C.26}$$

where the second step follows from (C.9) and the third because matrices inside a trace commute. Eq. (C.12) implies that

$$\begin{aligned} Cov\left[E_u(d\hat{R}'_u), E_t(d\hat{R}_t)\right] =& Cov\left(\gamma_1^R \hat{C}_u + \gamma_2^R C_u + \gamma_3^R y_u, \gamma_1^R \hat{C}_t + \gamma_2^R C_t + \gamma_3^R y_t\right) p_f \Sigma^2 p'_f dudt \\ =& H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, t-u, \nu_0) p_f \Sigma^2 p'_f dudt, \end{aligned} \tag{C.27}$$

where the second step follows from (C.6). Substituting (C.25)-(C.27) into (5.5), and integrating, we find the numerator in (5.4).

The term inside the square root in the denominator in (4.3) is equal to

$$\begin{aligned} & f\left[E\left(w_t^{\hat{M}}\right)\Sigma E\left(w_t^{\hat{M}'}\right) - \frac{\left[E\left(w_t^{\hat{M}}\right)\Sigma\eta'\right]^2}{\eta\Sigma\eta'}\right] + k\left[E\left(w_t^{\hat{M}}\right)\Sigma p'_f\right]^2 \\ & + f\left[Cov\left(w_t^{\hat{M}}, \Sigma w_t^{\hat{M}'}\right) - \frac{Var\left(w_t^{\hat{M}}\Sigma\eta'\right)}{\eta\Sigma\eta'}\right] + kVar\left(w_t^{\hat{M}}\Sigma p'_f\right) \\ =& \int_{t-\tau}^{t}\int_{t-\tau}^{t}(T_1 + T_2) \\ =& \int_{t-\tau}^{t}\int_{t-\tau}^{t} T_1 + 2\int_{t-\tau \le u' < u \le t} T_2 + \int_{t-\tau}^{t} T_3, \end{aligned} \tag{C.28}$$

where

$$T_1 \equiv f\left[E(d\hat{R}'_{u'})\Sigma E(d\hat{R}_u) - \frac{E(d\hat{R}'_{u'})\Sigma\eta' E(d\hat{R}'_u)\Sigma\eta'}{\eta\Sigma\eta'}\right] + kE(d\hat{R}'_{u'})\Sigma p'_f E(d\hat{R}'_u)\Sigma p'_f,$$

$$T_2 \equiv f\left[Cov(d\hat{R}'_{u'}, \Sigma d\hat{R}_u) - \frac{Cov(d\hat{R}'_{u'}\Sigma\eta', d\hat{R}'_u\Sigma\eta')}{\eta\Sigma\eta'}\right] + kCov(d\hat{R}'_{u'}\Sigma p'_f, d\hat{R}'_u\Sigma p'_f),$$

$$T_3 \equiv f\left[Cov(d\hat{R}'_u, \Sigma d\hat{R}_u) - \frac{Var(d\hat{R}'_u\Sigma\eta')}{\eta\Sigma\eta'}\right] + kVar(d\hat{R}'_u\Sigma p'_f),$$

and the second step in (C.28) follows from (5.2), and the third from separating non-diagonal from diagonal terms. Eq. (C.24) implies that

$$T_1 = L_2^2\Delta_3 du' du. \tag{C.29}$$

We can write (T_2, T_3) as

$$T_2 = T_{2a} + T_{2b}, \tag{C.30}$$

$$T_3 = T_{3a} + T_{3b} = T_{3a}, \tag{C.31}$$

where

$$T_{2a} \equiv f\left[E\left[Cov_{u'}(d\hat{R}'_{u'}, \Sigma d\hat{R}_u)\right] - \frac{E\left[Cov_{u'}(d\hat{R}'_{u'}\Sigma\eta', d\hat{R}'_u\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Cov_{u'}(d\hat{R}'_{u'}\Sigma p'_f, d\hat{R}'_u\Sigma p'_f)\right],$$

$$T_{2b} \equiv f\left[Cov\left[E_{u'}(d\hat{R}'_{u'}), \Sigma E_u(d\hat{R}_u)\right] - \frac{Cov\left[E_{u'}(d\hat{R}'_{u'})\Sigma\eta', E_u(d\hat{R}'_u)\Sigma\eta'\right]}{\eta\Sigma\eta'}\right] + kCov\left[E_{u'}(d\hat{R}'_{u'})\Sigma p'_f, E_u(d\hat{R}'_u)\Sigma p'_f\right],$$

$$T_{3a} \equiv f\left[E\left[Cov_u(d\hat{R}'_u, \Sigma d\hat{R}_u)\right] - \frac{E\left[Var_u(d\hat{R}'_u\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Var_u(d\hat{R}'_u\Sigma p'_f)\right],$$

$$T_{3b} \equiv f\left[Cov\left[E_u(d\hat{R}'_u), \Sigma E_u(d\hat{R}_u)\right] - \frac{Var\left[E_u(d\hat{R}'_u)\Sigma\eta'\right]}{\eta\Sigma\eta'}\right] + kVar\left[E_u(d\hat{R}'_u)\Sigma p'_f\right].$$

(The term $E_{u'}(d\hat{R}_u)$ can be replaced by $E_u(d\hat{R}_u)$ in T_{2b} because of (C.22), and the second step in (C.31) follows because T_{3a} is of order du and T_{3b} of order $(du)^2$.) Eq. (C.15) implies that

$$\begin{aligned}
T_{2a} =& f \left[E\left\{ Tr\left[Cov_{u'}(d\hat{R}_{u'}, d\hat{R}'_u \Sigma) \right] \right\} - \frac{E\left[Cov_{u'}(\eta \Sigma d\hat{R}_{u'}, d\hat{R}'_u \Sigma \eta') \right]}{\eta \Sigma \eta'} \right] \\
& + kE\left[Cov_{u'}(p_f \Sigma d\hat{R}'_{u'}, d\hat{R}'_u \Sigma p'_f) \right] \\
=& f \left[E\left\{ Tr\left[Cov_{u'}(d\hat{R}_{u'}, d\hat{R}'_u)\Sigma \right] \right\} - \frac{\eta \Sigma E\left[Cov_{u'}(d\hat{R}_{u'}, d\hat{R}'_u) \right] \Sigma \eta'}{\eta \Sigma \eta'} \right] \\
& + kp_f \Sigma E\left[Cov_{u'}(d\hat{R}'_{u'}, d\hat{R}'_u) \right] \Sigma p'_f \\
=& G(\gamma_1^R, \gamma_2^R, \gamma_3^R, u-u', \nu_0) \left\{ f \left[Tr(\Sigma p'_f p_f \Sigma^2) - \frac{(\eta \Sigma^2 p'_f)^2}{\eta \Sigma \eta'} \right] + k(p_f \Sigma^2 p'_f)^2 \right\} du' du \\
=& G(\gamma_1^R, \gamma_2^R, \gamma_3^R, u-u', \nu_0) \Delta_3 du' du,
\end{aligned} \tag{C.32}$$

where the fourth step follows because matrices inside a trace commute. Similarly (C.13) implies that

$$\begin{aligned}
T_{3a} =& f \left[E\left\{ Tr\left[Cov_u(d\hat{R}_u, d\hat{R}'_u)\Sigma \right] \right\} - \frac{\eta \Sigma E\left[Cov_u(d\hat{R}_u, d\hat{R}'_u)) \right] \Sigma \eta'}{\eta \Sigma \eta'} \right] \\
& + kp_f \Sigma E\left[Cov_u(d\hat{R}_u, d\hat{R}'_u)) \right] \Sigma p'_f \\
=& \Delta_6 du.
\end{aligned} \tag{C.33}$$

Eq. (C.12) implies that

$$\begin{aligned}
T_{2b} =& Cov\left(\gamma_1^R \hat{C}_{u'} + \gamma_2^R C_{u'} + \gamma_3^R y_{u'}, \gamma_1^R \hat{C}_u + \gamma_2^R C_u + \gamma_3^R y_u \right) \Delta_3 du' du \\
=& H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, u-u', \nu_0) \Delta_3 du' du,
\end{aligned} \tag{C.34}$$

where the second step follows from (C.6). Substituting (C.29)-(C.34) into (C.28), and integrating, we find the term inside the square root in the denominator in (5.4). ∎

Proof of Proposition 5.2: Proceeding as in the proof of Proposition 5.1, we find the following counterpart of (5.5):

$$\frac{1}{dt} \int_{t-\tau}^{t} \left\{ E\left[Cov_u(dR'_u, d\hat{R}_t) \right] + Cov\left[E_u(dR'_u), E_t(d\hat{R}_t) \right] + E(dR'_u)E(d\hat{R}_t) \right\}. \tag{C.35}$$

Since (C.15) shows that $Cov_t(dR'_t, d\hat{R}_{t'}) = Cov_t(d\hat{R}_t, d\hat{R}'_{t'})$, the first term inside the curly bracket in (C.35) is equal to its counterpart in (5.5). Eqs. (C.12) and (C.19) imply that

$$E_t(dR_t) = L_1 \Sigma \eta' dt + E_t(d\hat{R}_t) \tag{C.36}$$

$$\Rightarrow E(dR_t) = L_1 \Sigma \eta' dt + E(d\hat{R}_t)$$

$$\Rightarrow E(dR_t) = \left(L_1 \Sigma \eta' + L_2 \Sigma p'_f\right) dt, \tag{C.37}$$

where the third step follows from (C.24). Since (C.36) shows that $E_t(dR_t)$ is equal to $E_t(d\hat{R}_t)$ plus a constant, the second term inside the curly bracket in (C.35) is equal to its counterpart in (5.5). Moreover, (C.24) and (C.37) imply that the third term is equal to its counterpart in (5.5) plus $L_1 L_2 \eta \Sigma^2 p'_f du dt$. Integrating (C.35), we find that the numerator of the Sharpe ratio is as in the proposition.

Proceeding as in the proof of Proposition 5.1, we find (C.28), with (T_1, T_2, T_3) evaluated for raw returns $(dR_{u'}, dR_u)$ rather than index-adjusted returns $(d\hat{R}_{u'}, d\hat{R}_u)$. Since (C.15) shows that $Cov_t(dR_t, dR'_{t'}) = Cov_t(d\hat{R}_t, d\hat{R}'_{t'})$, (T_{2a}, T_{3a}) are the same as with index-adjusted returns. Since (C.36) shows that $E_t(dR_t)$ is equal to $E_t(d\hat{R}_t)$ plus a constant, T_{2b} is also the same as with index-adjusted returns. Eq. (3.8) implies that

$$\begin{aligned} T_{3a} =& f\left[E\left\{Tr\left[Cov_u(dR_u, dR'_u)\Sigma\right]\right\} - \frac{\eta \Sigma E\left[Cov_u(dR_u, dR'_u))\right]\Sigma \eta'}{\eta \Sigma \eta'}\right] \\ &+ k p_f \Sigma E\left[Cov_u(dR_u, dR'_u))\right]\Sigma p'_f \\ =& \left\{ f\left[Tr\left[(f\Sigma + k\Sigma p'_f p_f \Sigma)\Sigma\right] - \frac{\eta \Sigma (f\Sigma + k\Sigma p'_f p_f \Sigma)\Sigma \eta'}{\eta \Sigma \eta'}\right] + k p_f \Sigma (f\Sigma + k\Sigma p'_f p_f \Sigma)\Sigma p'_f \right\} du \\ =& \left(\Delta_6 + f\frac{\Delta_1}{\eta \Sigma \eta'}\right) du. \end{aligned}$$

Therefore, T_{3a} is equal to its counterpart under index-adjusted returns plus $f\frac{\Delta_1}{\eta\Sigma\eta'}du$. Eq. (C.37) implies that

$$\begin{aligned} T_1 =& \left\{ (L_1 \eta \Sigma + L_2 p_f \Sigma)\Sigma(L_1 \Sigma \eta' + L_2 \Sigma p'_f) - \frac{\left[(L_1 \eta \Sigma + L_2 p_f \Sigma)\Sigma \eta'\right]^2}{\eta \Sigma \eta'} \right. \\ &\left. + k\left[(L_1 \eta \Sigma + L_2 p_f \Sigma)\Sigma p'_f\right]^2 \right\} du' du \\ =& \left(L_1^2 \Delta_1 + 2 L_1 L_2 \Delta_2 + L_2^2 \Delta_3\right) du' du. \end{aligned}$$

Therfore, T_1 is equal to its counterpart under index-adjusted returns plus $\left(L_1^2\Delta_1 + 2L_1L_2\Delta_2\right) du'du$. Integrating (C.28), we find that the term inside the square root in the denominator of the Sharpe ratio is as in the proposition. ∎

Proof of Proposition 5.3: We first show that

$$\Sigma^i\theta' = \bar{\theta}\Sigma^i\eta' + v_i' \tag{C.38}$$

for all $i \in \mathbb{N}$, where

$$\begin{aligned}
\bar{\theta} \equiv & \frac{\theta_n + \theta_{n'}}{\eta_n + \eta_{n'}}, \\
v_i \equiv & (v_{i1}, .., v_{iN}), \\
v_{in} \equiv & (\Sigma_{nn} - \Sigma_{nn'})^i \epsilon_n, \\
\epsilon_n \equiv & \theta_n - \bar{\theta}\eta_n.
\end{aligned}$$

We proceed by induction on i. Eq. (C.38) holds for $i = 0$. To show that it holds for $i + 1$ if it holds for i, it suffices to show that

$$\Sigma v_i' = v_{i+1}'. \tag{C.39}$$

The n'th element of $\Sigma v_i'$ is

$$\begin{aligned}
(\Sigma v_i')_n = & \sum_{m \in \{1,..,N\}} \Sigma_{nm} v_{im} \\
= & \sum_{m \in \{1,..,N\} \setminus \{n,n'\}} \Sigma_{nm} v_{im} + \Sigma_{nn} v_{in} + \Sigma_{nn'} v_{in'}.
\end{aligned} \tag{C.40}$$

Eqs. $\Sigma_{nn} = \Sigma_{n'n'}$ and

$$\epsilon_n + \epsilon_{n'} = \theta_n + \theta_{n'} - \bar{\theta}(\eta_n + \eta_{n'}) = 0 \tag{C.41}$$

imply that

$$v_{in} + v_{in'} = (\Sigma_{nn} - \Sigma_{nn'})^i \epsilon_n + (\Sigma_{nn} - \Sigma_{nn'})^i \epsilon_{n'} = 0. \tag{C.42}$$

Eq. (C.42) implies that

$$\Sigma_{nn} v_{in} + \Sigma_{nn'} v_{in'} = (\Sigma_{nn} - \Sigma_{nn'}) v_{in} = (\Sigma_{nn} - \Sigma_{nn'})^{i+1} \epsilon_n = v_{i+1,n}. \tag{C.43}$$

Moreover, for each pair $(m, m') \neq (n, n')$, (C.42) and $\Sigma_{nm} = \Sigma_{nm'}$ imply that

$$\Sigma_{nm} v_{im} + \Sigma_{nm'} v_{im'} = \Sigma_{nm}(v_{im} - v_{im'}) = 0. \tag{C.44}$$

Eqs. (C.40), (C.43) and (C.44) imply (C.39).

Eqs. (C.38), (C.42) and $\eta_n = \eta_{n'}$ imply that

$$\eta \Sigma^i \theta' = \bar{\theta} \eta \Sigma^i \eta',$$

which in turn implies that

$$\eta \Sigma^i p'_f = \eta \Sigma^i \theta' - \frac{\eta \Sigma \theta'}{\eta \Sigma \eta'} \eta \Sigma^i \eta' = \eta \Sigma^i \theta' - \bar{\theta} \eta \Sigma^i \eta' = 0,$$

$$\Delta_1 = f \left[\eta \Sigma^3 \eta' - \frac{(\eta \Sigma^2 \eta')^2}{\eta \Sigma \eta'} \right],$$

and $\Delta_2 = 0$. Since $\eta \Sigma^2 p'_f = 0$, Proposition 5.2 implies that the numerator of the Sharpe ratio is equal under raw and index-adjusted returns. Moreover, the denominator is larger under raw returns because $\Delta_2 = 0$ and because the Cauchy-Schwarz inequality applied to the vectors $\Sigma^{\frac{1}{2}} \eta'$ and $\Sigma^{\frac{3}{2}} \eta'$ implies that $\Delta_1 \geq 0$. Therefore, if momentum yields a positive Sharpe ratio under raw returns, it yields a higher Sharpe ratio under index-adjusted returns. This inequality is strict when the Cauchy-Schwarz inequality is strict, which is the case when the vectors $\Sigma^{\frac{1}{2}} \eta'$ and $\Sigma^{\frac{3}{2}} \eta'$ are not collinear, i.e., $\Sigma^{\frac{1}{2}} \eta'$ is not an eigenvector of Σ. ∎

D Value

Proof of Lemma 6.1: Integrating the definition of returns, $dR_t = dD_t + dS_t - rS_t dt$, we find

$$S_t = \int_t^{t'} (dD_u - dR_u) e^{-r(u-t)} + S_{t'} e^{-r(t'-t)}.$$

Taking expectations as of time t, and letting $t' \to \infty$, we find the first equality in (6.1). The second equality follows from

$$E_t(dR_{t'}) = E_t\left[E_{t'}(dR_{t'})\right] = E_t\left[\frac{r\alpha\bar{\alpha}}{\alpha + \bar{\alpha}} \frac{\eta \Sigma \theta'}{\eta \Sigma \eta'} Cov_{t'}(dR_{t'}, \eta dR_{t'}) + \Lambda_{t'} Cov_{t'}(dR_{t'}, p_f dR_{t'})\right],$$

where the second step follows from (3.9). The second equality in each of (6.2) and (6.3) follows from (3.8). ∎

Proof of Proposition 6.1: The numerator in (4.3) is equal to

$$
\begin{aligned}
&\left(f+\frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(\Lambda_t\right)E\left(w_t^{\hat{V}}\Sigma p_f'\right)+Cov\left(\Lambda_t,w_t^{\hat{V}}\Sigma p_f'\right)\right]\\
=&\left(f+\frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(w_t^{\hat{V}}\right)E\left(\Lambda_t\Sigma p_f'\right)+Cov\left(w_t^{\hat{V}},\Lambda_t\Sigma p_f'\right)\right]\\
=&\left(f+\frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(\Gamma_{2t}'\right)E\left(\Lambda_t\Sigma p_f'\right)+Cov\left(\Gamma_{2t}',\Lambda_t\Sigma p_f'\right)-Cov\left(\frac{(1-\epsilon)(F_t-\bar{F})'}{r+\kappa},\Lambda_t\Sigma p_f'\right)\right],
\end{aligned}
\tag{D.1}
$$

where the third step follows from (6.8) and $E(F_t)=\bar{F}$ (implied by (C.4)). Eq. (D.1) coincides with (6.10) because of (C.12). Eqs. (3.10), (6.3) and (C.1)-(C.3) imply that

$$
E(\Gamma_{2t})=\left(f+\frac{k\Delta}{\eta\Sigma\eta'}\right)\left[\int_t^{\infty}E(\Lambda_{t'})e^{-r(t'-t)}dt'\right]\Sigma p_f'=\left(f+\frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{E(\Lambda_t)}{r}\Sigma p_f'=\frac{L_2}{r}\Sigma p_f'. \tag{D.2}
$$

Combining (C.24) and (D.2), we find

$$
E(\Gamma_{2t}')E(d\hat{R}_t)=\frac{L_2^2}{r}p_f\Sigma^2p_f'dt. \tag{D.3}
$$

Eqs. (3.1), (3.2), (6.1) and (6.2) imply that

$$
\Gamma_{2t}=a_0+(\gamma_1\hat{C}_t+\gamma_2C_t+\gamma_3y_t)\Sigma p_f'-\frac{\alpha\bar{\alpha}f}{\alpha+\bar{\alpha}}\frac{\eta\Sigma\theta'}{\eta\Sigma\eta'}\Sigma\eta'. \tag{D.4}
$$

Eqs. (C.12) and (D.4) imply that

$$
\begin{aligned}
Cov\left[\Gamma_{2t}',E_t(d\hat{R}_t)\right]=&Cov\left(\gamma_1\hat{C}_t+\gamma_2C_t+\gamma_3y_t,\gamma_1^R\hat{C}_t+\gamma_2^RC_t+\gamma_3^Ry_t\right)p_f\Sigma^2p_f'dt\\
=&H(\gamma_1,\gamma_2,\gamma_3,\gamma_1^R,\gamma_2^R,\gamma_3^R,0,\nu_0)p_f\Sigma^2p_f'dt,
\end{aligned}
\tag{D.5}
$$

where the second step follows from (C.6). Eq. (C.12) similarly implies that

$$
\begin{aligned}
Cov\left[F_t',E_t(d\hat{R}_t)\right]=&Cov\left(F_t',\gamma_1^R\hat{C}_t+\gamma_2^RC_t+\gamma_3^Ry_t\right)\Sigma p_f'dt\\
=&Cov\left(\gamma_1^R\hat{C}_t+\gamma_2^RC_t+\gamma_3^Ry_t,F_t\right)'\Sigma p_f'dt\\
=&K_1(\gamma_1^R,\gamma_3^R,0,\nu_0)p_f\Sigma^2p_f'dt,
\end{aligned}
\tag{D.6}
$$

where the third step follows from (C.7). Substituting (D.3)-(D.6) into (6.10), we find the numerator in (6.9).

The term inside the square root in the denominator in (4.3) is equal to

$$\begin{aligned}
& f\left[E\left(w_t^{\hat{V}}\right)\Sigma E\left(w_t^{\hat{V}\prime}\right)-\frac{\left[E\left(w_t^{\hat{V}}\right)\Sigma\eta'\right]^2}{\eta\Sigma\eta'}\right]+k\left[E\left(w_t^{\hat{V}}\right)\Sigma p_f'\right]^2 \\
& +f\left[Cov\left(w_t^{\hat{V}},\Sigma w_t^{\hat{V}\prime}\right)-\frac{Var\left(w_t^{\hat{V}}\Sigma\eta'\right)}{\eta\Sigma\eta'}\right]+kVar\left(w_t^{\hat{V}}\Sigma p_f'\right) \\
=& f\left[E\left(\Gamma_{2t}'\right)\Sigma E\left(\Gamma_{2t}\right)-\frac{\left[E\left(\Gamma_{2t}'\right)\Sigma\eta'\right]^2}{\eta\Sigma\eta'}\right]+k\left[E\left(\Gamma_{2t}'\right)\Sigma p_f'\right]^2 \\
& +f\left[Cov\left(\Gamma_{2t}',\Sigma\Gamma_{2t}\right)-\frac{Var\left(\Gamma_{2t}'\Sigma\eta'\right)}{\eta\Sigma\eta'}\right]+kVar\left(\Gamma_{2t}'\Sigma p_f'\right) \\
& -\frac{2(1-\epsilon)}{r+\kappa}\left\{f\left[Cov\left(\Gamma_{2t}',\Sigma F_t\right)-\frac{Cov\left(\Gamma_{2t}'\Sigma\eta',F_t'\Sigma\eta'\right)}{\eta\Sigma\eta'}\right]+kCov\left(\Gamma_{2t}'\Sigma p_f',F_t'\Sigma p_f'\right)\right\} \\
& +\frac{(1-\epsilon)^2}{(r+\kappa)^2}\left\{f\left[Cov\left(F_t',\Sigma F_t\right)-\frac{Var\left(F_t'\Sigma\eta'\right)}{\eta\Sigma\eta'}\right]+kVar\left(F_t'\Sigma p_f'\right)\right\},
\end{aligned} \tag{D.7}$$

where the second step follows from (6.8) and $E(F_t)=\bar{F}$. Eq. (D.2) implies that

$$f\left[E\left(\Gamma_{2t}'\right)\Sigma E\left(\Gamma_{2t}\right)-\frac{\left[E\left(\Gamma_{2t}'\right)\Sigma\eta'\right]^2}{\eta\Sigma\eta'}\right]+k\left[E\left(\Gamma_{2t}'\right)\Sigma p_f'\right]^2=\frac{L_2^2}{r^2}\Delta_3. \tag{D.8}$$

The remaining terms in (D.7) are

$$\begin{aligned}
& f\left[Cov\left(\Gamma_{2t}',\Sigma\Gamma_{2t}\right)-\frac{Var\left(\Gamma_{2t}'\Sigma\eta'\right)}{\eta\Sigma\eta'}\right]+kVar\left(\Gamma_{2t}'\Sigma p_f'\right) \\
=& Var(\gamma_1\hat{C}_t+\gamma_2 C_t+\gamma_3 y_t)\Delta_3 \\
=& H(\gamma_1,\gamma_2,\gamma_3,\gamma_1,\gamma_2,\gamma_3,0,\nu_0)\Delta_3,
\end{aligned} \tag{D.9}$$

where the first step follows from (D.4) and the second from (C.6),

$$
\begin{aligned}
& f\left[Cov\left(\Gamma'_{2t}, \Sigma F_t\right) - \frac{Cov\left(\Gamma'_{2t}\Sigma\eta', F'_t\Sigma\eta'\right)}{\eta\Sigma\eta'}\right] + kCov\left(\Gamma'_{2t}\Sigma p'_f, F'_t\Sigma p'_f\right) \\
=& f\left[Cov\left(\Gamma'_{2t}, \Sigma F_t\right) - \frac{Cov\left(\Gamma'_{2t}\Sigma\eta', \eta'\Sigma F_t\right)}{\eta\Sigma\eta'}\right] + kCov\left(\Gamma'_{2t}\Sigma p'_f, p_f\Sigma F_t\right) \\
=& \left\{f\left[p_f\Sigma^2 - \frac{p_f\Sigma^2\eta'\eta'\Sigma}{\eta\Sigma\eta'}\right] + kp_f\Sigma^2 p'_f p_f\Sigma\right\} Cov\left(\gamma_1\hat{C}_t + \gamma_2 C_t + \gamma_3 y_t, F_t\right) \\
=& K_1(\gamma_1, \gamma_3, 0, \nu_0)\Delta_3,
\end{aligned}
\tag{D.10}
$$

where the second step follows from (D.4) and the third from (C.7), and

$$
\begin{aligned}
& f\left[Cov\left(F'_t, \Sigma F_t\right) - \frac{Var\left(F'_t\Sigma\eta'\right)}{\eta\Sigma\eta'}\right] + kVar\left(F'_t\Sigma p'_f\right) \\
=& f\left[Tr\left[Cov\left(\Sigma F_t, F'_t\right)\right] - \frac{Cov\left(\eta\Sigma F_t, F'_t\Sigma\eta'\right)}{\eta\Sigma\eta'}\right] + kCov\left(p_f\Sigma F_t, F'_t\Sigma p'_f\right) \\
=& f\left[Tr\left[\Sigma Cov\left(F_t, F'_t\right)\right] - \frac{\eta\Sigma Cov\left(F_t, F'_t\right)\Sigma\eta'}{\eta\Sigma\eta'}\right] + kp_f\Sigma Cov\left(F_t, F'_t\right)\Sigma p'_f \\
=& \frac{\phi^2}{2\kappa}\left\{f\left[Tr\left(\Sigma^2\right) - \frac{\eta\Sigma^3\eta'}{\eta\Sigma\eta'}\right] + kp_f\Sigma^3 p'_f\right\} \\
=& \frac{\phi^2}{2\kappa}\Delta_4,
\end{aligned}
\tag{D.11}
$$

where the third step follows from (C.11). Substituting (D.8)-(D.11) into (D.7), we find the term inside the square root in the denominator in (6.9).

We finally show that for small s, the first and second terms in (6.10) yield positive value profits. From (6.9), this amounts to showing that $H(\gamma_1, \gamma_2, \gamma_3, \gamma_1^R, \gamma_2^R, \gamma_3^R, 0, \nu_0)$ is positive and $K_1(\gamma_1^R, \gamma_3^R, 0, \nu_0)$ is negative. Vayanos and Woolley (VW 2011) show that for small s, (b_1, b_2, γ_1) are of order one and positive, (γ_3, γ_3^R) of order one and negative, $(\gamma_2, \gamma_1^R, \gamma_2^R, \rho, \beta_1)$ of order s^2, and β_2 of order smaller than s^2. Hence,

$$
\begin{aligned}
& H(\gamma_1, \gamma_2, \gamma_3, \gamma_1^R, \gamma_2^R, \gamma_3^R, 0, \nu_0) \\
=& \left[\frac{\gamma_3^R b_1}{2\kappa(\kappa - b_2)}\left(\gamma_1 - \frac{\gamma_3 b_1}{\kappa + b_2}\right) - \frac{\gamma_3^R b_1}{(\kappa + b_2)(\kappa - b_2)}\left(\gamma_1 - \frac{\gamma_3 b_1}{2b_2}\right)\right]\beta_1^2\frac{\Delta}{\eta\Sigma\eta'} + o(s^4) \\
=& -\frac{\gamma_3^R\gamma_1 b_1\beta_1^2\Delta}{2\kappa(\kappa + b_2)\eta\Sigma\eta'} + o(s^4),
\end{aligned}
$$

and is positive for small s. Moreover,

$$K_1(\gamma_1^R, \gamma_3^R, 0, \nu_0) = \frac{\gamma_3^R b_1 \phi^2 \beta_2}{2\kappa(\kappa + b_2)(r + \kappa)} + o(\beta_2),$$

and is negative for small s. ∎

Proof of Proposition 6.2: Eqs. (6.10) and (D.7) hold under raw prices, but with $\Gamma_{1t}+\Gamma_{2t}$ instead of Γ_{2t}. Eq. (6.2) implies that Γ_{1t} is equal to the constant $\frac{L_1}{r}\Sigma\eta'$. Therefore, the additional terms in the case of (6.10), i.e., the numerator in (6.9), are

$$\frac{1}{dt}\Gamma_{1t}' E(d\hat{R}_t) = \frac{L_1 L_2}{r}\eta\Sigma^2 p_f',$$

where the second step follows from (C.24). Moreover, the additional terms in the case of (D.7), i.e., the term inside the square root in the denominator in (6.9), are

$$\begin{aligned} & f\left[\Gamma_{1t}'\Sigma\Gamma_{1t} - \frac{(\Gamma_{1t}'\Sigma\eta')^2}{\eta\Sigma\eta'}\right] + k\left(\Gamma_{1t}'\Sigma p_f'\right)^2 \\ & + 2\left\{f\left[\Gamma_{1t}'\Sigma E\left(\Gamma_{2t}\right) - \frac{\Gamma_{1t}'\Sigma\eta' E\left(\Gamma_{2t}'\right)\Sigma\eta'}{\eta\Sigma\eta'}\right] + k\Gamma_{1t}'\Sigma p_f' E\left(\Gamma_{2t}'\right)\Sigma p_f'\right\} \\ = & \frac{L_1^2}{r^2}\Delta_1 + 2\frac{L_1 L_2}{r^2}\Delta_2, \end{aligned}$$

where the second step follows from (D.2). ∎

Proof of Proposition 6.3: Proceeding as in the proof of Proposition 5.3, we find $\eta\Sigma^2 p_f' = 0$, $\Delta_1 \geq 0$ and $\Delta_2 = 0$. The numerator of the Sharpe ratio is thus equal under raw and index-adjusted prices, and the denominator is larger under raw prices. Therefore, if value yields a positive Sharpe ratio under raw prices, it yields a higher Sharpe ratio under index-adjusted prices. The condition for the inequality to be strict follows by the same argument as in the proof of Proposition 5.3. ∎

E Combining Momentum and Value

Lemma E.1 *The solution to the problem of maximizing (4.4) subject to (7.1) is (4.7) and*

$$\hat{y}^i = \frac{E(\hat{w}_t^i dR_t) Var(\hat{w}_t^j dR_t) - E(\hat{w}_t^j dR_t) Cov(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t)}{a\left[Var(\hat{w}_t^A dR_t) Var(\hat{w}_t^B dR_t) - Cov(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t)^2\right]}, \tag{E.1}$$

for $i, j \in \{A, B\}$ and $i \neq j$. The investor's maximum utility is given by (4.9), where SR_w is replaced by $SR_{w^{AB}}$.

Proof: Substituting (7.1) into (4.4), and noting that ηdR_t is orthogonal with $(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t)$, we can write the investor's maximization problem as

$$\max_{\hat{\hat{x}},\hat{y}} \left\{ \hat{\hat{x}} E(\eta dR_t) + \sum_{i=A,B} \hat{y}^i E(\hat{w}_t^i dR_t) \right.$$
$$\left. -\frac{a}{2} \left[\hat{\hat{x}}^2 Var(\eta dR_t) + \sum_{i=A,B} (\hat{y}^i)^2 Var(\hat{w}_t dR_t) + 2\hat{y}^A \hat{y}^B Cov(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t) \right] \right\}. \tag{E.2}$$

Eqs. (4.7) and (E.1) follow from (E.2). Substituting (4.7) and (E.1) into the investor's utility, which is the term in curly brackets in (E.2), we find (4.9) where SR_w is replaced by $SR_{w^{AB}}$. Comparison with (4.9) implies that $SR_{w^{AB}}$ is the Sharpe ratio of the optimal portfolio of (w_t^A, w_t^B). ∎

Proof of Proposition 7.1: The covariance between two general trading strategies (w_t^A, w_t^B) is

$$\begin{aligned}
& Cov\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right) \\
=& E\left[Cov_t\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right)\right] + Cov\left[E_t\left(\hat{w}_t^A dR_t\right), E_t\left(\hat{w}_t^B dR_t\right)\right] \\
=& E\left[Cov_t\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right)\right] \\
=& E\left[\left(w_t^A - \frac{w_t^A \Sigma \eta'}{\eta \Sigma \eta'}\eta\right)\left(f\Sigma + k\Sigma p_f' p_f \Sigma\right)\left(w_t^B - \frac{w_t^B \Sigma \eta'}{\eta \Sigma \eta'}\eta\right)'\right] dt \\
=& \left\{ f\left[E\left(w_t^A \Sigma w_t^{B'}\right) - \frac{E\left(w_t^A \Sigma \eta' w_t^B \Sigma \eta'\right)}{\eta \Sigma \eta'}\right] + kE\left(w_t^A \Sigma p_f' w_t^B \Sigma p_f'\right)\right\} dt \\
=& \left\{ f\left[E\left(w_t^A\right)\Sigma E\left(w_t^{B'}\right) - \frac{E\left(w_t^A\right)\Sigma\eta' E\left(w_t^B\right)\Sigma\eta'}{\eta\Sigma\eta'}\right] + kE\left(w_t^A\right)\Sigma p_f' E\left(w_t^B\right)\Sigma p_f' \right. \\
& \left. + f\left[Cov\left(w_t^A, \Sigma w_t^{B'}\right) - \frac{Cov\left(w_t^A \Sigma\eta', w_t^B \Sigma \eta'\right)}{\eta\Sigma\eta'}\right] + kCov\left(w_t^A \Sigma p_f', w_t^B \Sigma p_f'\right)\right\} dt,
\end{aligned} \tag{E.3}$$

where the second step follows because $E\left[Cov_t\left(\hat{w}_t^A dR_t, \hat{w}_t^B dR_t\right)\right]$ is of order dt and $Cov\left[E_t\left(\hat{w}_t^A dR_t\right),\right.$ $\left.E_t\left(\hat{w}_t^B dR_t\right)\right]$ of order $(dt)^2$, and the third step follows from (3.8) and (B.1).

Using (5.2), (6.8), $E(F_t) = \bar{F}$ and (E.3), we find

$$Cov\left(\hat{w}_t^{\hat{M}} dR_t, \hat{w}_t^{\hat{V}} dR_t\right) = \int_{t-\tau}^{t} \left(T_1 + T_2 - \frac{1-\epsilon}{r+\kappa} T_3\right) dt, \tag{E.4}$$

where

$$T_1 \equiv f\left[E(d\hat{R}_u')\Sigma E(\Gamma_{2t}) - \frac{E(d\hat{R}_u')\Sigma\eta' E(\Gamma_{2t}')\Sigma\eta'}{\eta\Sigma\eta'}\right] + kE(d\hat{R}_u')\Sigma p_f' E(\Gamma_{2t}')\Sigma p_f',$$

$$T_2 \equiv f\left[Cov(d\hat{R}_u', \Sigma\Gamma_{2t}) - \frac{Cov(d\hat{R}_u'\Sigma\eta', \Gamma_{2t}'\Sigma\eta')}{\eta\Sigma\eta'}\right] + kCov(d\hat{R}_u'\Sigma p_f', \Gamma_{2t}'\Sigma p_f'),$$

$$T_3 \equiv f\left[Cov(d\hat{R}_u', \Sigma F_t) - \frac{Cov(d\hat{R}_u'\Sigma\eta', F_t'\Sigma\eta')}{\eta\Sigma\eta'}\right] + kCov(d\hat{R}_u'\Sigma p_f', F_t'\Sigma p_f').$$

Eqs. (C.24) and (D.2) imply that

$$T_1 = \frac{L_2^2}{r}\Delta_3 du. \tag{E.5}$$

We can write (T_2, T_3) as

$$T_2 = T_{2a} + T_{2b}, \tag{E.6}$$

$$T_3 = T_{3a} + T_{3b}, \tag{E.7}$$

where

$$T_{2a} \equiv f\left[E\left[Cov_u(d\hat{R}_u', \Sigma\Gamma_{2t})\right] - \frac{E\left[Cov_u(d\hat{R}_u'\Sigma\eta', \Gamma_{2t}'\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Cov_u(d\hat{R}_u'\Sigma p_f', \Gamma_{2t}'\Sigma p_f')\right],$$

$$T_{2b} \equiv f\left[Cov\left[E_u(d\hat{R}_u'), \Sigma\Gamma_{2t}\right] - \frac{Cov\left[E_u(d\hat{R}_u')\Sigma\eta', \Gamma_{2t}'\Sigma\eta'\right]}{\eta\Sigma\eta'}\right] + kCov\left[E_u(d\hat{R}_u')\Sigma p_f', \Gamma_{2t}'\Sigma p_f'\right],$$

$$T_{3a} \equiv f\left[E\left[Cov_u(d\hat{R}_u', \Sigma F_t)\right] - \frac{E\left[Cov_u(d\hat{R}_u'\Sigma\eta', F_t'\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Cov_u(d\hat{R}_u'\Sigma p_f', F_t'\Sigma p_f')\right],$$

$$T_{3b} \equiv f\left[Cov\left[E_u(d\hat{R}_u'), \Sigma F_t\right] - \frac{Cov\left[E_u(d\hat{R}_u')\Sigma\eta', F_t'\Sigma\eta'\right]}{\eta\Sigma\eta'}\right] + kCov\left[E_u(d\hat{R}_u')\Sigma p_f', F_t'\Sigma p_f'\right].$$

(The terms $E_u(\Gamma_{2t})$ in T_{2b} and $E_u(F_t)$ in T_{3b} can be replaced by Γ_{2t} and F_t, respectively, because of (C.22).) Eq. (D.4) implies that

$$\begin{aligned}T_{2a} =& E\left[Cov_u(d\hat{R}'_u, \gamma_1\hat{C}_t + \gamma_2 C_t + \gamma_3 y_t)\right]\left\{f\left[\Sigma^2 p'_f - \frac{\Sigma\eta' p_f \Sigma^2\eta'}{\eta\Sigma\eta'}\right] + k\Sigma p'_f p_f \Sigma^2 p'_f\right\}\\ =& G(\gamma_1,\gamma_2,\gamma_3,t-u,\nu_0)\Delta_3 du, \qquad \text{(E.8)}\end{aligned}$$

where the second step follows from (C.14). Eqs. (C.12) and (D.4) imply that

$$\begin{aligned}T_{2b} =& E\left[Cov_u(\gamma_1^R\hat{C}_u + \gamma_2^R C_u + \gamma_3^R y_u, \gamma_1\hat{C}_t + \gamma_2 C_t + \gamma_3 y_t)\right]\Delta_3 du\\ =& H(\gamma_1^R,\gamma_2^R,\gamma_3^R,\gamma_1,\gamma_2,\gamma_3,t-u,\nu_0)\Delta_3 du, \qquad \text{(E.9)}\end{aligned}$$

where the second step follows from (C.6). Eq. (C.12) implies that

$$\begin{aligned}T_{3b} =& f\left[Cov\left[E_u(d\hat{R}'_u),\Sigma F_t\right] - \frac{Cov\left[E_u(d\hat{R}'_u)\Sigma\eta',\eta\Sigma F_t\right]}{\eta\Sigma\eta'}\right] + kCov\left[E_u(d\hat{R}'_u)\Sigma p'_f, p_f\Sigma F_t\right]\\ =& \left\{f\left[p_f\Sigma^2 - \frac{p_f\Sigma^2\eta'\eta\Sigma}{\eta\Sigma\eta'}\right] + kp_f\Sigma^2 p'_f p_f\Sigma\right\}E\left[Cov_u(\gamma_1^R\hat{C}_u + \gamma_2^R C_u + \gamma_3^R y_u, F_t)\right]du\\ =& K_1(\gamma_1^R,\gamma_3^R,t-u,\nu_0)\Delta_3 du, \qquad \text{(E.10)}\end{aligned}$$

where the third step follows from (C.7). Finally,

$$\begin{aligned}T_{3a} =& f\left[E\left\{Tr\left[Cov_u(d\hat{R}_u, F'_t\Sigma)\right]\right\} - \frac{E\left[Cov_u(\eta\Sigma d\hat{R}_u, F'_t\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Cov_u(p_f\Sigma d\hat{R}_u, F'_t\Sigma p'_f)\right]\\ =& f\left[E\left\{Tr\left[Cov_u(d\hat{R}_u, F'_t)\Sigma\right]\right\} - \frac{\eta\Sigma E\left[Cov_u(d\hat{R}_u, F'_t)\right]\Sigma\eta'}{\eta\Sigma\eta'}\right] + kp_f\Sigma E\left[Cov_u(d\hat{R}_u, F'_t)\right]\Sigma p'_f\\ =& \frac{\phi^2}{r+\kappa}\nu_0(\kappa,t-u)(\Delta_5 + \beta_2\gamma_1\Delta_3)du, \qquad \text{(E.11)}\end{aligned}$$

where the third step follows from (C.16) and because matrices inside a trace commute. Substituting (E.5)-(E.11) into (E.4), and integrating, we find (7.3).

We finally show that for small s, the temporal variation in weights generates negative covariance between momentum and value returns. The effects of temporal variation are the first, second, third and fifth term in (7.3). Using the asymptotics in the proof of Proposition 6.1, we find

$$G(\gamma_1,\gamma_2,\gamma_3,\tau,\nu_1) = -\left[\gamma_1\nu_1(\kappa,\tau) + \frac{\gamma_3 b_1}{\kappa - b_2}\left(\nu_1(\kappa,\tau) - \nu_1(b_2,\tau)\right)\right]\beta_1 + o(s^2),$$

$$H(\gamma_1^R,\gamma_2^R,\gamma_3^R,\gamma_1,\gamma_2,\gamma_3,\tau,\nu_1) = o(s^2),$$

$$\frac{1-\epsilon}{r+\kappa}K_1(\gamma_1^R,\gamma_3^R,\tau,\nu_1) = o(s^2),$$

$$\frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}\nu_1(\kappa,\tau)(\Delta_5 + \beta_2\gamma_1\Delta_3) = \frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}\nu_1(\kappa,\tau)\Delta_5 + o(1).$$

Therefore, if $\epsilon = 1$, the dominant term is $G(\gamma_1,\gamma_2,\gamma_3,\tau,\nu_1) < 0$, and if $\epsilon = 0$, the dominant term is $-\frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}\nu_1(\kappa,\tau)(\Delta_5 + \beta_2\gamma_1\Delta_3) < 0$. ∎

Proof of Proposition 7.2: Proceeding as in the proof of Proposition 7.1, we find (E.4), with (T_1, T_2, T_3) evaluated for raw returns dR_u and the total risk discount $\Gamma_{1t}+\Gamma_{2t}$, rather than for index-adjusted returns $d\hat{R}_u$ and the discount Γ_{2t}. Eq. (6.2) implies that Γ_{1t} is equal to the constant $\frac{L_1}{r}\Sigma\eta'$. Therefore, it does not affect $(T_{2a}, T_{2b}, T_{3a}, T_{3b})$. Since (C.5) and (C.14) imply that $Cov_t(d\hat{R}_t, \psi_1\hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'}) = Cov_t(dR_t, \psi_1\hat{C}_{t'} + \psi_2 C_{t'} + \psi_3 y_{t'})$, T_{2a} is the same as under index-adjustment. The variables (T_{2b}, T_{3b}) are also the same since (C.36) shows that $E_t(dR_t)$ is equal to $E_t(d\hat{R}_t)$ plus a constant. The variable T_{3a} is

$$\begin{aligned}
T_{3a} =& f\left[E\left\{Tr\left[Cov_u(dR_u, F_t'\Sigma)\right]\right\} - \frac{E\left[Cov_u(\eta\Sigma dR_u, F_t'\Sigma\eta')\right]}{\eta\Sigma\eta'}\right] + kE\left[Cov_u(p_f\Sigma dR_u, F_t'\Sigma p_f')\right] \\
=& f\left[E\left\{Tr\left[Cov_u(dR_u, F_t')\Sigma\right]\right\} - \frac{\eta\Sigma E\left[Cov_u(dR_u, F_t')\right]\Sigma\eta'}{\eta\Sigma\eta'}\right] + kp_f\Sigma E\left[Cov_u(dR_u, F_t')\right]\Sigma p_f' \\
=& \frac{\phi^2}{r+\kappa}\nu_0(\kappa, t-u)(\Delta_4 + \beta_2\gamma_1\Delta_3)du,
\end{aligned} \tag{E.12}$$

where the third step follows from (C.10). Therefore, T_{3a} is equal to its counterpart under index-adjustment plus $\frac{\phi^2}{r+\kappa}\nu_0(\kappa, t-u)\frac{\Delta_1}{\eta\Sigma\eta'}du$. Eqs. (C.37) and (D.2) imply that

$$\begin{aligned}
T_1 =& \frac{1}{r}\Bigg\{(L_1\eta\Sigma + L_2 p_f\Sigma)\Sigma(L_1\Sigma\eta' + L_2\Sigma p_f') - \frac{\left[(L_1\eta\Sigma + L_2 p_f\Sigma)\Sigma\eta'\right]^2}{\eta\Sigma\eta'} \\
&+ k\left[(L_1\eta\Sigma + L_2 p_f\Sigma)\Sigma p_f'\right]^2\Bigg\}du \\
=& \frac{1}{r}\left(L_1^2\Delta_1 + 2L_1L_2\Delta_2 + L_2^2\Delta_3\right)du.
\end{aligned}$$

Therfore, T_1 is equal to its counterpart under index-adjustment plus $\frac{1}{r}\left(L_1^2\Delta_1 + 2L_1L_2\Delta_2\right)du$. Integrating (E.4), we find the additional term (7.5). ∎

F Momentum and Value Over Long Horizons

The motivation for the use of the Sharpe ratio is as follows. Consider an investor who can invest in the riskless asset, the market index η and the strategy w_t, and has mean-variance preferences

$$E(\Delta W_{t,T}) - \frac{a}{2}Var(\Delta W_{t,T}), \tag{F.1}$$

over the increment $\Delta W_{t,T} \equiv W_{t+T}e^{-rT} - W_t$ in wealth. The investor chooses an investment $\hat{y}$ in the strategy and an overall exposure $\hat{\hat{x}}$ to the index at time t. We assume that these are chosen to grow at the riskless rate, so at time u the investment in the strategy is $\hat{y}e^{r(u-t)}$ and the overall exposure to the index is $\hat{\hat{x}}e^{r(u-t)}$. The investor's budget constraint is (4.5) and integrates to

$$\Delta W_{t,T} = \hat{\hat{x}} \int_t^{t+T} \eta dR_u + \hat{y} \int_t^{t+T} \hat{w}_u dR_u. \tag{F.2}$$

Substituting (F.2) into (F.1), and assuming that $(\int_t^{t+T} \eta dR_u, \int_t^{t+T} \hat{w}_u dR_u)$ are orthogonal, we can write the investor's maximization problem as

$$\max_{\hat{\hat{x}},\hat{y}} \left\{ \hat{\hat{x}} E\left(\int_t^{t+T} \eta dR_u\right) + \hat{y} E\left(\int_t^{t+T} \hat{w}_u dR_u\right) - \frac{a}{2}\left[\hat{\hat{x}}^2 Var\left(\int_t^{t+T} \eta dR_u\right) + \hat{y}^2 Var\left(\int_t^{t+T} \hat{w}_u dR_u\right)\right]\right\}. \tag{F.3}$$

The orthogonality of $(\int_t^{t+T} \eta dR_u, \int_t^{t+T} \hat{w}_u dR_u)$ does not generally follow from that of $(\eta dR_t, \hat{w}_t dR_t)$ because there can be lead-lag effects: ηdR_t can be correlated with $\hat{w}_u dR_u$ for $u > t$. Such a correlation can be ruled out under restrictions on the set of strategies and the asset structure. If $(\int_t^{t+T} \eta dR_u, \int_t^{t+T} \hat{w}_u dR_u)$ are orthogonal, then we can show that the investor's maximum utility depends on the characteristics of the strategy w_t only through the Sharpe ratio.

Lemma F.1 *The solution to the maximization problem (F.3) is*

$$\hat{\hat{x}} = \frac{E\left(\int_t^{t+T} \eta dR_u\right)}{aVar\left(\int_t^{t+T} \eta dR_u\right)}, \tag{F.4}$$

$$\hat{y} = \frac{E\left(\int_t^{t+T} \hat{w}_u dR_u\right)}{aVar\left(\int_t^{t+T} \hat{w}_u dR_u\right)}. \tag{F.5}$$

The investor's maximum utility is

$$\frac{E\left(\int_t^{t+T} \eta dR_u\right)^2}{2aVar\left(\int_t^{t+T} \eta dR_u\right)} + \frac{\left(SR_{w,T}\right)^2 T}{2a}. \tag{F.6}$$

Proof: The proof is the same as for Lemma 4.1. ∎

To prove Lemma 8.1, we first prove the following lemma:

Lemma F.2 *If the random variables* $\{X_i\}_{i=1,2,3,4}$ *are jointly normal, then*

$$Cov(X_1X_2, X_3) = E(X_1)Cov(X_2, X_3) + E(X_2)Cov(X_1, X_3) \tag{F.7}$$

$$\begin{aligned} Cov(X_1X_2, X_3X_4) =& E(X_1)E(X_3)Cov(X_2, X_4) + E(X_1)E(X_4)Cov(X_2, X_3) \\ &+ E(X_2)E(X_3)Cov(X_1, X_4) + E(X_2)E(X_4)Cov(X_1, X_3) \\ &+ Cov(X_1, X_3)Cov(X_2, X_4) + Cov(X_1, X_4)Cov(X_2, X_3). \end{aligned} \tag{F.8}$$

Proof: We set $\hat{X} \equiv X - E(X)$. To show (F.7), we note that

$$\begin{aligned} Cov(X_1X_2, X_3) =& Cov\left[(E(X_1) + \hat{X}_1)(E(X_2) + \hat{X}_2), X_3\right] \\ =& E(X_1)Cov(\hat{X}_2, X_3) + E(X_2)Cov(\hat{X}_1, X_3) + Cov(\hat{X}_1\hat{X}_2, X_3) \\ =& E(X_1)Cov(X_2, X_3) + E(X_2)Cov(X_1, X_3) + Cov(\hat{X}_1\hat{X}_2, \hat{X}_3). \end{aligned} \tag{F.9}$$

Eq. (F.9) implies (F.7) if

$$Cov(\hat{X}_1\hat{X}_2, \hat{X}_3) = 0. \tag{F.10}$$

To show (F.10), we write $\{\hat{X}_i\}_{i=1,2,4}$ as

$$\hat{X}_i = \frac{Cov(X_i, X_3)}{Var(X_3)}\hat{X}_3 + \epsilon_i, \tag{F.11}$$

where because of joint normality, ϵ_i is normal, mean zero and independent of X_3. Therefore,

$$\begin{aligned} Cov(\hat{X}_1\hat{X}_2, \hat{X}_3) =& \frac{Cov(X_1, X_3)Cov(X_2, X_3)}{Var(X_3)^2}Cov(\hat{X}_3^2, \hat{X}_3) + \frac{Cov(X_1, X_3)}{Var(X_3)}Cov(\hat{X}_3\epsilon_2, \hat{X}_3) \\ &+ \frac{Cov(X_2, X_3)}{Var(X_3)}Cov(\hat{X}_3\epsilon_1, \hat{X}_3) + Cov(\epsilon_1\epsilon_2, \hat{X}_3). \end{aligned} \tag{F.12}$$

The first term in (F.12) is zero because

$$Cov(\hat{X}_3^2, \hat{X}_3) = E(\hat{X}_3^3) = 0,$$

where the second step follows because $\hat{X}_3$ is normal with mean zero, and hence symmetrically distributed around zero. The second and third terms in (F.12) are zero because

$$Cov(\hat{X}_3\epsilon_i, \hat{X}_3) = E(\hat{X}_3^2\epsilon_i) = E(\hat{X}_3^2)E(\epsilon_i) = 0,$$

for $i = 1, 2$, where the second step follows because ϵ_i is independent of X_3, and the third because ϵ_i is mean zero. The fourth term in (F.12) is zero because (ϵ_1, ϵ_2) are independent of X_3. Therefore, (F.10) holds, and so does (F.7).

To show (F.8), we note that

$$\begin{aligned} Cov(X_1X_2, X_3X_4) =& Cov\left[(E(X_1) + \hat{X}_1)(E(X_2) + \hat{X}_2), X_3X_4\right] \\ =& E(X_1)Cov(\hat{X}_2, X_3X_4) + E(X_2)Cov(\hat{X}_1, X_3X_4) + Cov(\hat{X}_1\hat{X}_2, X_3X_4). \end{aligned} \tag{F.13}$$

Using (F.7), we find

$$\begin{aligned} E(X_1)Cov(\hat{X}_2, X_3X_4) =& E(X_1)\left[E(X_3)Cov(\hat{X}_2, X_4) + E(X_4)Cov(\hat{X}_2, X_3)\right] \\ =& E(X_1)E(X_3)Cov(X_2, X_4) + E(X_1)E(X_4)Cov(X_2, X_3), \end{aligned} \tag{F.14}$$

and similarly

$$E(X_2)Cov(\hat{X}_1, X_3X_4) = E(X_2)E(X_3)Cov(X_1, X_4) + E(X_2)E(X_4)Cov(X_1, X_3). \tag{F.15}$$

Moreover,

$$
\begin{aligned}
Cov(\hat{X}_1\hat{X}_2, X_3X_4) =& Cov\left[\hat{X}_1\hat{X}_2, (E(X_3)+\hat{X}_3)(E(X_4)+\hat{X}_4)\right] \\
=& E(X_3)Cov(\hat{X}_1\hat{X}_2, \hat{X}_4) + E(X_4)Cov(\hat{X}_1\hat{X}_2, \hat{X}_3) + Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\hat{X}_4) \\
=& Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\hat{X}_4), \qquad \text{(F.16)}
\end{aligned}
$$

where the second step follows because of (F.10). Using (F.11), we find

$$
Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\hat{X}_4) = \frac{Cov(X_3, X_4)}{Var(X_3)} Cov(\hat{X}_1\hat{X}_2, \hat{X}_3^2) + Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\epsilon_4), \qquad \text{(F.17)}
$$

$$
\begin{aligned}
Cov(\hat{X}_1\hat{X}_2, \hat{X}_3^2) =& \frac{Cov(X_1, X_3)Cov(X_2, X_3)}{Var(X_3)^2} Cov(\hat{X}_3^2, \hat{X}_3^2) + \frac{Cov(X_1, X_3)}{Var(X_3)} Cov(\hat{X}_3\epsilon_2, \hat{X}_3^2) \\
&+ \frac{Cov(X_2, X_3)}{Var(X_3)} Cov(\hat{X}_3\epsilon_1, \hat{X}_3^2) + Cov(\epsilon_1\epsilon_2, \hat{X}_3^2), \qquad \text{(F.18)}
\end{aligned}
$$

$$
\begin{aligned}
Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\epsilon_4) =& \frac{Cov(X_1, X_3)Cov(X_2, X_3)}{Var(X_3)^2} Cov(\hat{X}_3^2, \hat{X}_3\epsilon_4) + \frac{Cov(X_1, X_3)}{Var(X_3)} Cov(\hat{X}_3\epsilon_2, \hat{X}_3\epsilon_4) \\
&+ \frac{Cov(X_2, X_3)}{Var(X_3)} Cov(\hat{X}_3\epsilon_1, \hat{X}_3\epsilon_4) + Cov(\epsilon_1\epsilon_2, \hat{X}_3\epsilon_4). \qquad \text{(F.19)}
\end{aligned}
$$

Similar arguments as in the first part of the proof imply that the last three terms in (F.18) are zero, and so are the first and fourth terms in (F.19). To compute the first term in (F.18), we note that

$$
Cov(\hat{X}_3^2, \hat{X}_3^2) = E(\hat{X}_3^4) - E(\hat{X}_3^2) = 2E(\hat{X}_3^2)^2 = 2Var(X_3)^2,
$$

where the second step follows $\hat{X}_3$ is normal with mean zero. To compute the second and third terms in (F.19), we note that

$$
\begin{aligned}
Cov(\hat{X}_3\epsilon_2, \hat{X}_3\epsilon_4) =& E(\hat{X}_3^2\epsilon_2\epsilon_4) - E(\hat{X}_3\epsilon_2)E(\hat{X}_3\epsilon_4) \\
=& E(\hat{X}_3^2)E(\epsilon_2\epsilon_4) \\
=& E(\hat{X}_3^2)Cov(\epsilon_2, \epsilon_4) \\
=& Var(X_3)Cov(X_2, \epsilon_4)
\end{aligned}
$$

where the second, third and fourth steps follow because (ϵ_2, ϵ_4) are independent of X_3 and mean zero. Similarly, for the third term in (F.19),

$$
Cov(\hat{X}_3\epsilon_1, \hat{X}_3\epsilon_4) = Var(X_3)Cov(X_1, \epsilon_4).
$$

Therefore, (F.17)-(F.19) imply that

$$
\begin{aligned}
&Cov(\hat{X}_1\hat{X}_2, \hat{X}_3\hat{X}_4)\\
=&\frac{2Cov(X_1,X_3)Cov(X_2,X_3)Cov(X_3,X_4)}{Var(X_3)} + Cov(X_1,X_3)Cov(X_2,\epsilon_4) + Cov(X_1,X_3)Cov(X_1,\epsilon_4)\\
=&Cov(X_1,X_3)Cov\left[X_2, \frac{Cov(X_3,X_4)}{Var(X_3)}X_3 + \epsilon_4\right] + Cov(X_2,X_3)Cov\left[X_1, \frac{Cov(X_3,X_4)}{Var(X_3)}X_3 + \epsilon_4\right]\\
=&Cov(X_1,X_3)Cov(X_2,X_4) + Cov(X_2,X_3)Cov(X_1,X_4). \qquad \text{(F.20)}
\end{aligned}
$$

Eqs. (F.13)-(F.16) and (F.20) imply (F.8). ∎

Proof of Lemma 8.1: Eq. (8.5) follows from

$$
Cov(\hat{w}_t dR_t, \hat{w}_{t'} dR_{t'}) = E\left[Cov_t\left(\hat{w}_t dR_t, \hat{w}_{t'} dR_{t'}\right)\right] + Cov\left[E_t(\hat{w}_t dR_t), E_t(\hat{w}_{t'} dR_{t'})\right]
$$

and (C.22). We can write C_1 as

$$
\begin{aligned}
C_1 =& E\left[\hat{w}_t Cov_t\left(dR_t, \hat{w}_{t'} dR_{t'}\right)\right]\\
& E\left[\hat{w}_t Cov_t\left(dR_t, E_{t'}(\hat{w}_{t'} dR_{t'})\right)\right]\\
=& \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right) E\left[\hat{w}_t Cov_t\left(dR_t, \Lambda_{t'} w_{t'} \Sigma p_f'\right)\right] dt', \qquad \text{(F.21)}
\end{aligned}
$$

where the second step follows because the covariance $Cov_t\left(dR_t, \hat{w}_{t'} dR_{t'}\right)$ involves only the drift terms in $\hat{w}_{t'} dR_{t'}$ and not the Brownian terms, and the third step follows from (B.3). Since the strategy w_t is linear, the variable $w_{t'}\Sigma p_f'$ is normal. Applying (F.7) to the normal variables $(dR_t, \Lambda_{t'}, w_{t'}\Sigma p_f')$, we find

$$
Cov_t\left(dR_t, \Lambda_{t'} w_{t'} \Sigma p_f'\right) = E_t(\Lambda_{t'}) Cov_t(dR_t, w_{t'}\Sigma p_f') + E_t(w_{t'}\Sigma p_f') Cov_t(dR_t, \Lambda_{t'}). \qquad \text{(F.22)}
$$

Substituting (F.22) into (F.21), and noting that the conditional covariances in the right-hand side of (F.22) are constant because the variables $(dR_t, \Lambda_{t'}, w_{t'}\Sigma p_f')$ are normal, we find

$$
C_1 = \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\left[E\left(\hat{w}_t E_t(\Lambda_{t'})\right) Cov_t(dR_t, w_{t'}\Sigma p_f') + E\left(\hat{w}_t E_t(w_{t'}\Sigma p_f')\right) Cov_t\left(dR_t, \Lambda_{t'}\right)\right] dt',
$$

which implies (8.6). Using (B.3), we can write C_2 as

$$
C_2 = Cov\left(\Lambda_t w_t \Sigma p_f', \Lambda_{t'} w_{t'} \Sigma p_f'\right) dt dt'. \qquad \text{(F.23)}
$$

Applying (F.8) to the normal variables $(\Lambda_t, w_t\Sigma p_f', \Lambda_{t'}, w_{t'}\Sigma p_f')$, and noting that in steady state $E(\Lambda_t) = E(\Lambda_{t'})$ and $E(w_t\Sigma p_f') = E(w_{t'}\Sigma p_f')$, we find (8.7). ∎

Proof of Lemma 8.2: Eq. (8.8) follows from (3.10) and (C.1)-(C.3). Eq. (8.9) follows from (3.10) and (C.6). To derive (8.10), we note that

$$
\begin{aligned}
&E\left(\hat{w}_t w_{t'}\Sigma p_f'\right) Cov_t(dR_t, \Lambda_{t'})\\
=&\frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} E\left(\hat{w}_t w_{t'}\Sigma p_f'\right) G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t' - t, \nu_0)\Sigma p_f' dt\\
=&\frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} E\left(\hat{w}_t\Sigma p_f' w_{t'}\Sigma p_f'\right) \Lambda_{R,t'-t} dt\\
=&\frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}} E\left(w_t\Sigma p_f' w_{t'}\Sigma p_f'\right) \Lambda_{R,t'-t} dt,
\end{aligned}
\qquad \text{(F.24)}
$$

where the first step follows from (3.10) and (C.5), and the third from (B.1). Eq. (F.24) yields (8.10) because in steady state $E(w_t\Sigma p_f') = E(w_{t'}\Sigma p_f')$. ∎

Proof of Proposition 8.1: Eq. (8.11) follows from

$$
E\left(w_t^{\hat{M}}\Sigma^i v'\right) = E\left(w_t^{\hat{M}}\right)\Sigma^i v' = \left[\int_{t-\tau}^{t} E(d\hat{R}_u')\right]\Sigma^i v' = L_2\tau p_f \Sigma^{i+1} v',
$$

where the second step follows from (5.2), and the third from (C.24). To derive (8.12), we note that

$$
\begin{aligned}
Cov\left(w_t^{\hat{M}}\Sigma^i v', \Lambda_{t'}\right) =& Cov\left(w_t^{\hat{M}}, \Lambda_{t'}\right)\Sigma^i v'\\
=& \left[\int_{t-\tau}^{t} Cov(d\hat{R}_u', \Lambda_{t'})\right]\Sigma^i v',
\end{aligned}
\qquad \text{(F.25)}
$$

where the second step follows from (5.2). The term inside the integral in (F.25) can be written as

$$
\begin{aligned}
&E\left[Cov_u(d\hat{R}_u', \Lambda_{t'})\right] + Cov\left[E_u(d\hat{R}_u'), \Lambda_{t'}\right]\\
=&\frac{1}{f + \frac{k\Delta}{\eta\Sigma\eta'}}\left[G(\gamma_1^R, \gamma_2^R, \gamma_3^R, t' - u, \nu_0) + H(\gamma_1^R, \gamma_2^R, \gamma_3^R, \gamma_1^R, \gamma_2^R, \gamma_3^R, t' - u, \nu_0)\right] p_f\Sigma du,
\end{aligned}
\qquad \text{(F.26)}
$$

where the first step follows from (C.22), and the second from (3.10), (C.5), (C.6) and (C.12). Substituting (F.26) into (F.25), and integrating, we find (8.12). To derive (8.13), we use the counterpart of (F.25),

$$
Cov\left(\Lambda_t, w_{t'}^{\hat{M}}\Sigma p_f'\right) = \left[\int_{t'-\tau}^{t'} Cov(\Lambda_t, d\hat{R}_u')\right]\Sigma p_f', \qquad \text{(F.27)}
$$

and write the term inside the integral as

$$
\begin{aligned}
&E\left[Cov_u(\Lambda_t, d\hat{R}'_u)\right] + Cov\left[\Lambda_t, E_u(d\hat{R}'_u)\right] \\
=&\frac{1}{f+\frac{k\Delta}{\eta\Sigma\eta'}}\left[G(\gamma_1^R,\gamma_2^R,\gamma_3^R,t-u,\nu_0)1_{\{t>u\}} + H(\gamma_1^R,\gamma_2^R,\gamma_3^R,\gamma_1^R,\gamma_2^R,\gamma_3^R,|t-u|,\nu_0)\right]p_f\Sigma du,
\end{aligned} \tag{F.28}
$$

where the difference between (F.28) and (F.26) arises because u can exceed t. Substituting (F.28) into (F.27), and integrating, we find (8.13). To derive (8.14), we note that

$$
\begin{aligned}
&Cov\left(w_t^{\hat{M}}\Sigma^i p'_f, w_{t'}^{\hat{M}}\Sigma^i p'_f\right) \\
=&p_f\Sigma Cov\left(w^{\hat{M}'}, w_{t'}^{\hat{M}}\right)\Sigma p'_f \\
=&p_f\Sigma\left[\int_{t-\tau}^{t}\int_{t'-\tau}^{t'} Cov(d\hat{R}_u, d\hat{R}'_{u'})\right]\Sigma p'_f \\
=&p_f\Sigma\left[\int_{(u,u')\in[t-\tau,t]\times[t'-\tau,t'],u\neq u'} Cov(d\hat{R}_u, d\hat{R}'_{u'}) + \int_{u\in[t'-\tau,t]} Cov(d\hat{R}_u, d\hat{R}'_u)\right]\Sigma p'_f,
\end{aligned} \tag{F.29}
$$

where the second step follows from (5.2), and the third from separating non-diagonal from diagonal terms. The term inside the first integral is

$$
\begin{aligned}
&E\left[Cov_{\min\{u,u'\}}(d\hat{R}_u, d\hat{R}'_{u'})\right] + Cov\left[E_{\min\{u,u'\}}(d\hat{R}'_u), E_{\min\{u,u'\}}(d\hat{R}'_{u'})\right] \\
&E\left[Cov_{\min\{u,u'\}}(d\hat{R}_u, d\hat{R}'_{u'})\right] + Cov\left[E_u(d\hat{R}'_u), E_{u'}(d\hat{R}'_{u'})\right] \\
=&\left[G(\gamma_1^R,\gamma_2^R,\gamma_3^R,|u'-u|,\nu_0) + H(\gamma_1^R,\gamma_2^R,\gamma_3^R,\gamma_1^R,\gamma_2^R,\gamma_3^R,|u'-u|,\nu_0)\right]\Sigma p'_f p_f\Sigma du du',
\end{aligned} \tag{F.30}
$$

where the second step follows from (C.22), and the third from (C.5), (C.6) and (C.12). The term inside the second integral is

$$
\begin{aligned}
&E\left[Cov_u(d\hat{R}_u, d\hat{R}'_u)\right] + Cov\left[E_u(d\hat{R}'_u), E_u(d\hat{R}'_u)\right] \\
=&E\left[Cov_u(d\hat{R}_u, d\hat{R}'_u)\right] \\
=&\left[f\left(\Sigma - \frac{\Sigma\eta'\eta\Sigma}{\eta\Sigma\eta'}\right) + k\Sigma p'_f p_f\Sigma\right]du,
\end{aligned} \tag{F.31}
$$

where the second step follows because $Cov_u(d\hat{R}_u, d\hat{R}'_u)$ is of order du and $Cov\left[E_u(d\hat{R}'_u), E_u(d\hat{R}'_u)\right]$ of order $(du)^2$, and the third step follows from (C.13). Substituting (F.30) and (F.31) into (F.29), and integrating, we find (8.14). To derive (8.15), we note that

$$
\begin{aligned}
&Cov_t\left(dR_t, w_{t'}^{\hat{M}}\Sigma p_f'\right)\\
=&Cov_t\left(dR_t, w_{t'}^{\hat{M}}\right)\Sigma p_f'\\
=&\left[\int_{t'-\tau}^{t'} Cov_t(dR_t, d\hat{R}_u')\right]\Sigma p_f'\\
=&\left[\int_{t'-\tau}^{t'} Cov_t\left(dR_t, E_u(d\hat{R}_u')\right)1_{\{u>t\}} + Cov_t(dR_t, d\hat{R}_t')1_{\{t'-\tau<t\}}\right]\Sigma p_f'\\
=&\left[\int_{t'-\tau}^{t'} G(\gamma_1^R,\gamma_2^R,\gamma_3^R,u-t,\nu_0)\Sigma p_f' p_f\Sigma 1_{\{u>t\}}du + \left[f\left(\Sigma - \frac{\Sigma\eta'\eta\Sigma}{\eta\Sigma\eta'}\right) + k\Sigma p_f' p_f\Sigma\right]1_{\{t'-t<\tau\}}\right]\Sigma p_f' dt\\
=&\left[G(\gamma_1^R,\gamma_2^R,\gamma_3^R,\mathcal{T},\nu_4)\Sigma p_f' p_f\Sigma + \left[f\left(\Sigma - \frac{\Sigma\eta'\eta\Sigma}{\eta\Sigma\eta'}\right) + k\Sigma p_f' p_f\Sigma\right]1_{\{t'-t<\tau\}}\right]\Sigma p_f' dt\\
=&\left[w_{MR1,\mathcal{T}}\Sigma p_f' + w_{MR2,\mathcal{T}}\left(\Sigma^2 p_f' - \frac{\eta\Sigma^2 p_f'}{\eta\Sigma\eta'}\Sigma\eta'\right)\right]dt,
\end{aligned}
\tag{F.32}
$$

where the second step follows from (5.2), the third from separating non-diagonal from diagonal terms and noting that $Cov_t(dR_t, d\hat{R}_u')$ for $u>t$ involves only the drift terms in $d\hat{R}_u$ and not the Brownian terms, and the fourth from (C.5) and (C.13). Using (F.32), we find

$$
\begin{aligned}
&E(\hat{w}_t^{\hat{M}}\Lambda_{t'})Cov_t(dR_t, w_{t'}^{\hat{M}}\Sigma p_f')\\
=&E(\hat{w}_t^{\hat{M}}\Sigma p_f'\Lambda_{t'})w_{MR1,\mathcal{T}}dt + E\left[\hat{w}_t\left(\Sigma^2 p_f' - \frac{\eta\Sigma^2 p_f'}{\eta\Sigma\eta'}\Sigma\eta'\right)\Lambda_{t'}\right]w_{MR2,\mathcal{T}}dt\\
=&E(w_t^{\hat{M}}\Sigma p_f'\Lambda_{t'})w_{MR1,\mathcal{T}}dt + E\left[\left(w_t^{\hat{M}}\Sigma^2 p_f' - \frac{w_t^{\hat{M}}\Sigma\eta'}{\eta\Sigma\eta'}\eta\Sigma^2 p_f'\right)\Lambda_{t'}\right]w_{MR2,\mathcal{T}}dt,
\end{aligned}
\tag{F.33}
$$

where the second step follows from (B.1). Eq. (F.33) yields (8.15) because in steady state $E(\Lambda_t) = E(\Lambda_{t'})$. ∎

Proof of Proposition 8.2: Eq. (8.16) follows from

$$E\left(w_t^{\hat{V}}\Sigma^i v'\right) = E\left(w_t^{\hat{V}}\right)\Sigma^i v' = E(\Gamma_{2t}')\Sigma^i v' = \frac{L_2}{r}p_f\Sigma^{i+1}v',$$

where the second step follows from (6.8) and $E(F_t) = \bar{F}$ (implied by (C.4)), and the third from (D.2). Eq. (8.17) follows from

$$Cov\left(w_t^{\hat{V}}\Sigma^i v', \Lambda_{t'}\right) = Cov\left(w_t^{\hat{V}}, \Lambda_{t'}\right)\Sigma^i v' = \left[Cov\left(\Gamma'_{2t}, \Lambda_{t'}\right) - \frac{1-\epsilon}{r+\kappa}Cov\left(F'_t, \Lambda_{t'}\right)\right]\Sigma^i v',$$

(where the second step follows from (6.8)), (3.10), (C.6), (C.8) and (D.4). Eq. (8.18) follows similarly from

$$Cov\left(\Lambda_t, w_{t'}^{\hat{V}}\Sigma p'_f\right) = \left[Cov\left(\Lambda_t, \Gamma'_{2t'}\right) - \frac{1-\epsilon}{r+\kappa}Cov\left(\Lambda_t, F'_{t'}\right)\right]\Sigma p'_f,$$

(3.10), (C.6), (C.7) and (D.4). Eq. (8.19) follows from

$$\begin{aligned}
&Cov\left(w_t^{\hat{V}}\Sigma^i p'_f, w_{t'}^{\hat{V}}\Sigma^i p'_f\right)\\
=&p_f\Sigma Cov\left(w_t^{\hat{V}'}, w_{t'}^{\hat{V}}\right)\Sigma^i p'_f\\
=&p_f\Sigma Cov\left[Cov(\Gamma_{2t}, \Gamma'_{2t'}) - \frac{1-\epsilon}{r+\kappa}Cov(\Gamma_{2t}, F'_{t'}) - \frac{1-\epsilon}{r+\kappa}Cov(F_t, \Gamma'_{2t'}) + \frac{(1-\epsilon)^2}{(r+\kappa)^2}Cov(F_t, F'_{t'})\right]\Sigma p'_f,
\end{aligned}$$

(where the second step follows from (6.8)), (C.6), (C.7), (C.8), (C.11), and (D.4). To derive (8.20), we note that

$$\begin{aligned}
&Cov_t\left(dR_t, w_{t'}^{\hat{V}}\Sigma p'_f\right)\\
=&Cov_t\left(dR_t, w_{t'}^{\hat{V}}\right)\Sigma p'_f\\
=&\left[Cov_t(dR_t, \Gamma'_{2t'}) - \frac{1-\epsilon}{r+\kappa}Cov_t(dR_t, F'_{t'})\right]\Sigma p'_f\\
=&\left[G(\gamma_1, \gamma_2, \gamma_3, t'-t, \nu_0)\Sigma p'_f p_f\Sigma - \frac{(1-\epsilon)\phi^2}{(r+\kappa)^2}(\Sigma + \beta_2\gamma_1\Sigma p'_f p_f\Sigma)\nu_0(\kappa, t'-t)\right]\Sigma p'_f dt\\
=&\left[w_{VR1,t'-t}\Sigma p'_f + w_{VR2,t'-t}\left(\Sigma^2 p'_f - \frac{\eta\Sigma^2 p'_f}{\eta\Sigma\eta'}\Sigma\eta'\right)\right]dt, \qquad \text{(F.34)}
\end{aligned}$$

where the second step follows from (6.8), and the third from (C.5) and (C.10). Eq. (F.34) yields (8.20) through the same steps used to derive (8.15) from (F.32). ∎

G Calibration

Lemmas G.1-G.3 compute the quantities that we use in the calibration. These lemmas assume general (η, Σ), unless stated otherwise. Lemma G.4 specializes the results to $\eta = \mathbf{1}$ and $\Sigma = \hat{\sigma}^2(I + \omega \mathbf{1}'\mathbf{1})$.

Lemma G.1 *The Sharpe ratio of the market index η, expressed in annualized terms, is*

$$SR_\eta \equiv \frac{E(\eta dR_t)}{\sqrt{Var(\eta dR_t)dt}} = \frac{r\alpha\bar{\alpha}\sqrt{f}}{\alpha + \bar{\alpha}} \frac{\eta\Sigma\theta'}{\sqrt{\eta\Sigma\eta'}}. \tag{G.1}$$

The correlation between stock n and the index is

$$Corr(dR_{nt}, \eta dR_t) = \frac{\sqrt{f}(\Sigma\eta')_n}{\sqrt{\eta\Sigma\eta' \left[f\Sigma_{nn} + k[(\Sigma p_f')_n]^2\right]}}. \tag{G.2}$$

The fraction of stock n's variance that is generated by fund flows is

$$\frac{k[(\Sigma p_f')_n]^2}{f\Sigma_{nn} + k[(\Sigma p_f')_n]^2}. \tag{G.3}$$

Proof: Eq. (C.37) implies that

$$E(\eta dR_t) = L_1 \eta\Sigma\eta' dt. \tag{G.4}$$

Moreover,

$$Var(\eta dR_t) = E\left[Var_t(\eta dR_t)\right] = f\eta\Sigma\eta' dt, \tag{G.5}$$

where the first step follows by replacing $\hat{w}_t$ by η in the first two steps in (B.4), and the second step follows from (3.8). Substituting (G.4) and (G.5) into the definition of SR_η, we find (G.1). The correlation between stock n and the index is

$$Corr(dR_{nt}, \eta dR_t) = \frac{Cov(dR_{nt}, \eta dR_t)}{\sqrt{Var(dR_{nt})Var(\eta dR_t)}} = \frac{(f\Sigma\eta')_n}{\sqrt{f\eta\Sigma\eta'(f\Sigma + k\Sigma p_f' p_f \Sigma)_{nn}}}, \tag{G.6}$$

where the second step follows as in (G.5). Eq. (G.6) implies (G.2). Eq. (G.3) follows by separating fundamental and non-fundamental covariance terms in (3.8). ∎

Kacperczyk, Sialm, and Zheng (KSZ1 2005) compute the sum across industry sectors of squared deviations between the weight that an active fund gives to a sector and the sector's weight in a broad market index. We compute the same measure, but for simplicity construct weights using the number of shares times the covariance between one share and the index rather than times the dollar value of one share (see Footnote 3). When $\Sigma = \hat{\sigma}^2(I + \omega \mathbf{1}'\mathbf{1})$, the covariance is the same for all stocks. When, in addition, $\eta = 1$, the index fund gives weight $1/N$ to stock n, and the active fund gives

$$\frac{\theta_n - x_t}{\sum_{n=1}^N (\theta_n - x_t)} = \frac{\theta_n - x_t}{N(\bar{\theta} - x_t)},$$

since its portfolio in equilibrium is $\theta - x_t\eta$. The average deviation between weights is

$$\sqrt{\frac{\sum_{n=1}^N \left[\frac{\theta_n - x_t}{N(\bar{\theta} - x_t)} - \frac{1}{N}\right]^2}{N}} = \frac{\sigma(\theta)}{N(\bar{\theta} - x_t)} = \frac{\sigma(\theta)}{N\bar{\theta}} \frac{\bar{\theta}}{\bar{\theta} - x_t} \tag{G.7}$$

The ratio $\bar{\theta}/(\bar{\theta} - x_t)$ is the total number of shares held by the active and index funds, over the number of shares held by the active fund only. We set this to 11/10, since the holdings of active funds are about ten times those of index funds in KSZ1's sample period. Setting the left-hand side of (G.7) to 6.6%, N to 10, and $\bar{\theta}$ to one, we find $\sigma(\theta) = 0.6$.

Kacperczyk, Sialm, and Zheng (KSZ 2008) compute the difference in CAPM alphas between top and bottom return-gap deciles by evaluating the return gap over the year before time t and computing monthly CAPM alphas over the fourth month after time t. To replicate their calculation, we determine the distribution of the active fund's expected index-adjusted return (CAPM alpha) over an interval $[t + \tau_1, t + \tau_2]$ and conditional on C_t, which is minus the return gap at time t. We then compute the expectation over the top and bottom deciles of that distribution, and take the difference.

Lemma G.2 *The difference in the active fund's expected index-adjusted return over the interval* $[t + \tau_1, t + \tau_2]$ *between top and bottom return-gap deciles as of time* t *is*

$$-H\left(0, 1, 0, \gamma_1^R \frac{\Delta}{\eta \Sigma \eta'}, \gamma_2^R \frac{\Delta}{\eta \Sigma \eta'} - 1, \gamma_3^R \frac{\Delta}{\eta \Sigma \eta'}, \mathcal{T}, \nu_6\right) \frac{\sqrt{2\kappa}}{s} \left[E(z|z > z_9) - E(z|z < z_1)\right], \tag{G.8}$$

where $\mathcal{T} \equiv (\tau_1, \tau_2)$, *$z$ is a standardized normal variable, and* z_i, $i = 1, .., 9$, *is the boundary between the deciles* i *and* $i + 1$ *of* z.

Proof: The active fund's index-adjusted return is the discrepancy between the return of the fund and of its index benchmark. Since the fund's return in equilibrium is $(\theta - x_t\eta)dR_t - C_t dt$, its index-adjusted return is

$$
\begin{aligned}
&(\theta - x_t\eta)dR_t - C_t dt - \frac{Cov_t[(\theta - x_t\eta)dR_t - C_t dt, \eta dR_t]}{Var_t(\eta dR_t)}\eta dR_t \\
=&(\theta - x_t\eta)dR_t - C_t dt - \frac{Cov_t[(\theta - x_t\eta)dR_t, \eta dR_t]}{Var_t(\eta dR_t)}\eta dR_t \\
=&(\theta - x_t\eta)dR_t - C_t dt - \left(\frac{\theta\Sigma\eta'}{\eta\Sigma\eta'} - x_t\right)\eta dR_t \qquad \text{(G.9)}\\
=&p_f dR_t - C_t dt, \qquad \text{(G.10)}
\end{aligned}
$$

where the second step follows from (3.8). The fund's index-adjusted return over the interval $[t + \tau_1, t + \tau_2]$ is

$$
\int_{t+\tau_1}^{t+\tau_2} (p_f dR_u - C_u du).
$$

Because of normality, the expectation of this return conditional on C_t is ZC_t, where

$$
Z \equiv \frac{Cov\left[C_t, \int_{t+\tau_1}^{t+\tau_2}(p_f dR_u - C_u du)\right]}{Var(C_t)} = \frac{\int_{t+\tau_1}^{t+\tau_2} Cov\left[C_t, p_f dR_u - C_u du\right]}{Var(C_t)}. \qquad \text{(G.11)}
$$

Since the return gap is minus C_t, the difference in expected index-adjusted return between top and bottom return-gap deciles is

$$
-Z\sqrt{Var(C_t)}\left[E(z|z > z_9) - E(z|z < z_1)\right]. \qquad \text{(G.12)}
$$

The covariance inside the integral in (G.11) is

$$
\begin{aligned}
&Cov\left[C_t, p_f dR_u - C_u du\right] \\
=&Cov\left[C_t, p_f E_u(dR_u) - C_u du\right] \\
=&Cov\left[C_t, (\gamma_1^R \hat{C}_u + \gamma_2^R C_u + \gamma_3^R y_u)\frac{\Delta}{\eta\Sigma\eta'} - C_u\right] du \\
=&H\left(0, 1, 0, \gamma_1^R\frac{\Delta}{\eta\Sigma\eta'}, \gamma_2^R\frac{\Delta}{\eta\Sigma\eta'} - 1, \gamma_3^R\frac{\Delta}{\eta\Sigma\eta'}, u - t, \nu_0\right) du, \qquad \text{(G.13)}
\end{aligned}
$$

where the second step follows from (C.19) and the third from (C.6). Substituting (G.13) into (G.11), integrating, substituting into (G.12), and noting that (C.6) implies that $Var(C_t) = s^2/2\kappa$, we find (G.8). ∎

The difference in expected returns in Lemma G.2 is per share of the active fund, and one share coincides with the entire fund (Eq. (2.4)). To compare with KSZ, we need to express this difference per dollar, dividing by the active fund's dollar value. We assume that this dollar value is 10/11 of the total dollar value of the active and index funds, which is the value of the true market portfolio θ. To infer the latter, we divide the per-share annualized standard deviation of θ by the per-dollar one. The per-share standard deviation is

$$\sqrt{\frac{Var(\theta dR_t)}{dt}} = \sqrt{\theta(f\Sigma + k\Sigma p_f' p_f \Sigma)\theta'} = \sqrt{f\frac{(\eta\Sigma\theta')^2}{\eta\Sigma\eta'} + \left(f + \frac{k\Delta}{\eta\Sigma\eta'}\right)\frac{\Delta}{\eta\Sigma\eta'}},$$

where the first step follows as in (G.5). We assume that the per-dollar one is 15%, as for the index. To complete the comparison with KSZ, we set $(\tau_1, \tau_2) = (3/12, 4/12)$, and note that the term $[E(z|z > z_9) - E(z|z < z_1)]$ is approximately 3.4.

The investor's holdings of stock n at time t, through the index and active funds, are $(x_t\eta + y_t z_t)_n$. Lemma G.14 computes the standard deviation of the change in holdings between t and $t+\tau$. The change in holdings is the investor's signed volume. It can also be interpreted as the stock's flow-generated volume as a percent of assets managed by the active and index funds. This is because it is expressed in terms of number of shares, and the combined holdings of the active and index funds in the average stock are $\bar{\theta} = 1$ share.

Lemma G.3 *The standard deviation of the change in the investor's holdings of stock n between t and $t+\tau$ is*

$$\sqrt{2\left[H(0,0,1,0,0,1,0,\nu_0) - H(0,0,1,0,0,1,\tau,\nu_0)\right]}\,|(p_f)_n|\,. \tag{G.14}$$

Proof: Eq. (A.1) implies that the standard deviation of the change in the investor's holdings of stock n between t and $t+\tau$ is

$$\sqrt{Var\left[(x_{t+\tau}\eta + y_{t+\tau}z_{t+\tau})_n - (x_t\eta + y_t z_t)_n\right]} = \sqrt{Var(y_{t+\tau} - y_t)}\,|(p_f)_n|\,. \tag{G.15}$$

Moreover,

$$\begin{aligned} Var(y_{t+\tau} - y_t) =& Var(y_{t+\tau}) + Var(y_t) - 2Cov(y_t, y_{t+\tau}) \\ =& 2Var(y_t) - 2Cov(y_t, y_{t+\tau}), \end{aligned} \tag{G.16}$$

where the second step follows from stationarity. Combining (G.15) with (G.16), and using (C.6), we find (G.14). ∎

Lemma G.4 *If $\eta = \mathbf{1}$ and $\Sigma = \hat{\sigma}^2(I + \omega \mathbf{1}'\mathbf{1})$, then for all $i \in \mathbb{N}$,*

$$\eta\Sigma^i\eta' = \hat{\sigma}^{2i}(1+\omega N)^i N, \tag{G.17}$$

$$\eta\Sigma^i p_f' = 0, \tag{G.18}$$

$$p_f\Sigma^i p_f' = \hat{\sigma}^{2i}\sigma(\theta)^2 N, \tag{G.19}$$

$$Tr(\Sigma^i) = \hat{\sigma}^{2i}\left[(1+\omega N)^i + N - 1\right], \tag{G.20}$$

$$\Sigma^i p_f' = \hat{\sigma}^{2i}(\theta - \bar{\theta}\mathbf{1})'. \tag{G.21}$$

Proof: Using the binomial formula and $\eta = \mathbf{1}$, we find

$$\Sigma^i = \hat{\sigma}^{2i}\left(\sum_{i'=0}^{i} C^i_{i'}\omega^{i'}(\mathbf{1}'\mathbf{1})^{i'}\right) = \hat{\sigma}^{2i}\left(I + \sum_{i'=1}^{i} C^i_{i'}\omega^{i'}N^{i'-1}\mathbf{1}'\mathbf{1}\right). \tag{G.22}$$

Eq. (G.22) implies that

$$\Sigma^i\eta' = \hat{\sigma}^{2i}\left(1 + \sum_{i'=1}^{i} C^i_{i'}\omega^{i'}N^{i'}\right)\mathbf{1}' = \hat{\sigma}^{2i}(1+\omega N)^i\mathbf{1}', \tag{G.23}$$

$$\Sigma^i\theta' = \hat{\sigma}^{2i}\left[\theta + \left(\sum_{i'=1}^{i} C^i_{i'}\omega^{i'}N^{i'}\right)\bar{\theta}\mathbf{1}\right]' = \hat{\sigma}^{2i}\left[\theta - \bar{\theta}\mathbf{1} + (1+\omega N)^i\bar{\theta}\mathbf{1}\right]'. \tag{G.24}$$

Eqs. (G.23) and (G.24) imply that

$$\eta\Sigma\theta' = \bar{\theta}\eta\Sigma\eta',$$

and hence

$$p_f = \theta - \bar{\theta}\mathbf{1}. \tag{G.25}$$

Eq. (G.17) follows from (G.23). Eq. (G.21) follows from (G.23)-(G.25). Eq. (G.18) follows from (G.21). Eq. (G.19) follows from (G.21) and (G.25). Eq. (G.20) follows from (G.22). ∎

Lemma G.4 implies that the only characteristics of θ that affect Sharpe ratios are $\bar{\theta}$, which can be normalized to one, and $\sigma(\theta)$. The results in Lemmas (G.1) and (G.3) depend on the absolute values of the components of the vector $p_f = \theta - \bar{\theta}\mathbf{1} = \theta - \mathbf{1}$, i.e., the deviations between θ and $\mathbf{1}$. To compute the quantities in these lemmas, we set the deviations to their average value, which is $\sigma(\theta)$.

References

Albuquerque, Rui, and Jianjun Miao, 2010, Advance information and asset prices, working paper Boston University.

Ang, Andrew, and Joseph Chen, 2002, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* 63, 443–494.

Asness, Clifford, Tobias Moskowitz, and Lasse Pedersen, 2009, Value and momentum everywhere, working paper University of Chicago.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 261–199.

——— , and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Cespa, Giovani, and Xavier Vives, 2011, Higher order expectations, illiquidity, and short-term trading, working paper City University.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 475–512.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A theory of overconfidence, self-attribution, and security market under and over-reactions, *Journal of Finance* 53, 1839–1885.

Dasgupta, Amil, Andrea Prat, and Michela Verardo, 2011, The price impact of institutional herding, *Review of Financial Studies* 24, 892–925.

DeBondt, Werner, and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

Fama, Eugene, 1991, Efficient capital markets: II, *Journal of Finance* 46, 1575–1617.

——— , and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47(2), 427–465.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

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

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

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, Industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

———, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

Koijen, Ralph, Juan Carlos Rodriguez, and Alessandro Sbuelz, 2006, Momentum and mean-reversion in strategic asset allocation, working paper University of Chicago.

Lou, Dong, 2011, A flow-based explanation for return predictability, working paper London School of Economics.

Philippon, Thomas, 2008, The evolution of the us financial industry from 1860 to 2007: Theory and evidence, working paper New York University.

Schwert, William, 2003, Anomalies and market efficiency, in George Constantinides, Milton Harris, and Rene Stulz, ed.: *Handbook of the Economics of Finance* (Elsevier).

Vayanos, Dimitri, and Paul Woolley, 2011, An institutional theory of momentum and reversal, working paper London School of Economics.