

Forecast Accuracy Uncertainty and Momentum

Bing Han

McCombs School of Business, University of Texas at Austin, Austin, Texas 78712,
bing.han@mcombs.utexas.edu

Dong Hong, Mitch Warachka

Lee Kong Chian School of Business, Singapore Management University, Singapore 178899, Singapore
{donghong@smu.edu.sg, mitchell@smu.edu.sg}

We demonstrate that stock price momentum and earnings momentum can result from uncertainty surrounding the accuracy of cash flow forecasts. Our model has multiple information sources issuing cash flow forecasts for a stock. The investor combines these forecasts into an aggregate cash flow estimate that has minimal mean-squared forecast error. This aggregate estimate weights each cash flow forecast by the estimated accuracy of its issuer, which is obtained from their past forecast errors. Momentum arises from the investor gradually learning about the relative accuracy of the information sources and updating their weights. Empirical tests validate the model's prediction of stronger momentum in stocks with large information weight fluctuations and high forecast dispersion. We also identify return predictability attributable to changes in the information weights.

Key words: momentum; uncertainty; learning

History: Received February 7, 2007; accepted December 21, 2008, by David A. Hsieh, finance. Published online in *Articles in Advance* March 23, 2009.

1. Introduction

This paper studies the theoretical and empirical implications of forecast accuracy uncertainty on stock returns. Our representative investor receives a disperse range of forecasts regarding a firm's future cash flow growth but is uncertain about the accuracy of the information sources issuing these forecasts. The investor optimally combines the forecasts into an aggregate cash flow estimate. To minimize the mean-squared forecast error of this aggregate estimate, the investor assigns more weight to forecasts issued by more accurate information sources. The corresponding aggregate cash flow estimate represents the investor's expectation of future cash flow growth and determines the firm's stock price.

The investor estimates the accuracy of each information source from their past forecast errors.¹ As additional cash flow realizations and forecast errors become available, the investor learns about their respective accuracy. Intuitively, an information source's true accuracy represents its unobservable "skill" at forecasting a firm's cash flow. Investors understand the uncertainty inherent in measuring this skill and gradually update their assessment of each information source's accuracy.²

¹ Sinha et al. (1997), Brown (2001), and Clement and Tse (2003) find that prior forecast errors predict analysts' future forecast accuracy.

² Section 2.1 contains additional justification for the gradual updating of the forecast accuracy estimates.

Our model features a risk-neutral representative investor, constant fundamental risk, and a constant discount rate. Expected stock returns are driven entirely by innovations in the investor's aggregate cash flow estimate. These innovations are determined by changes in the information weights and the dynamics of individual forecasts. We focus on the role of time-varying information weights, which has not been previously studied, by assuming that the individual cash flow growth forecasts are, on average, constant over short horizons.

Although the investor in our model immediately incorporates newly issued or revised forecasts into her conditional cash flow expectation, the weights assigned to these forecasts are gradually updated. The gradual updating of the information weights generates return predictability. In particular, earnings momentum and price momentum arise from learning about the relative forecast accuracy of the information sources. For example, after a series of unexpected positive cash flow innovations, the estimated accuracy of relatively optimistic information sources tends to improve. Thus, their information weights increase at the expense of pessimistic information sources. As a consequence, the optimistic information sources exert a greater influence on the aggregate cash flow estimate. This shift in the information weights leads to higher expected cash flow growth and a higher stock price, although the individual forecasts remain unchanged (on average).

Our framework offers several empirical predictions. Momentum is expected to be stronger for stocks with greater fluctuations in their information weights. This unique prediction is verified for both price momentum and earnings momentum using analyst earnings forecasts. In addition, we confirm our model's prediction that momentum is stronger for stocks with greater cash flow uncertainty using analyst forecast dispersion as a proxy. A simulation study also confirms that under reasonable parameters, forecast accuracy uncertainty produces momentum profits whose magnitudes are comparable with existing empirical studies.

Several other predictions from our model are consistent with the empirical evidence in Jiang et al. (2005), Daniel and Titman (2006), Jackson and Johnson (2006), and Zhang (2006), although we provide a new interpretation of their findings. For example, we predict stronger momentum in stocks with fewer available forecast errors; including small firms, young firms, and those undergoing significant changes in their cash flow growth. Stronger momentum for stocks with higher return volatility and higher cash flow volatility are also predicted.

This paper provides a middle ground between behavioral and rational perspectives on momentum. Momentum in our model does not originate from time-varying risk or a time-varying risk premium. Instead, our framework is based on a statistical optimization that combines multiple forecasts of uncertain accuracy into an aggregate cash flow estimate and a learning process that induces slow updating in the weights underlying this aggregate cash flow estimate. The gradual updating of these weights is distinct from the slow diffusion of information in Hong and Stein (1999) and Hong et al. (2007b). Unlike agents in rational expectation models, our investor is not concerned with the impact of her learning on prices. Although our investor is not *assumed* to be influenced by behavioral biases, §2.3 demonstrates that certain characteristics of the information weights mimic behavioral biases that have been invoked to explain momentum, including representativeness and conservatism (Barberis et al. 1998) as well as overconfidence (Daniel et al. 1998). We also demonstrate that information sources whose forecasts are positively correlated with more accurate forecasts are marginalized, which leads to the appearance of limited attention (e.g., Hirshleifer and Teoh 2003, Peng and Xiong 2006).

Our paper is related to recent studies on the ability of parameter uncertainty to generate return predictability (e.g., Timmermann 1993, Lewellen and Shanken 2002). In these studies, investors learn about an unknown parameter regarding a firm's cash flow dynamics. In contrast, our investor does not model cash flow dynamics and does not learn about a firm's cash flow growth parameters from realized cash flows.

Instead, our investor's reliance on multiple cash flow forecasts with time-varying weights is crucial.

Our framework also differs from Hong et al. (2007a). Their representative investor uses simple univariate models to forecast cash flow when the true cash flow generating process is multivariate. This investor is limited to a subset of available information and permanently alternates between two incorrect forecast procedures. In contrast, our investor conditions on all available forecasts when forming her cash flow expectation. We are the first to examine the uncertainty surrounding the relative accuracy of different cash flow forecasts.

The remainder of this paper is organized as follows. Section 2 introduces the optimal information weights, the learning mechanism regarding forecast accuracy, and the pricing implications of forecast accuracy uncertainty. Section 3 evaluates the implications of changes in the information weights on stock returns, earnings momentum, and price momentum. Section 4 summarizes and concludes the paper.

2. The Model

Following Barberis et al. (1998), our economy consists of a single risky security (stock) and a risk-neutral representative investor with an exogenous constant discount rate δ . All cash flows N_t are paid out as dividends. Under the objective probability, cash flow growth $y_{t+1} \equiv N_{t+1} - N_t$ is assumed to be independent over time, with an unknown and time-varying mean θ_{t+1} :

$$y_{t+1} = \theta_{t+1} + \varepsilon_{t+1}, \quad (1)$$

$$\theta_{t+1} \sim \mathcal{N}(\bar{\theta}, \sigma_{\theta}^2), \quad (2)$$

$$\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_y^2). \quad (3)$$

The parameter $\bar{\theta}$ represents the unconditional average cash flow growth rate, and is set to zero without loss of generality. A nonzero unconditional mean cash flow growth rate adds a constant term to the average stock return but does not affect our conclusions regarding return predictability.³ The parameter σ_{θ} captures the uncertainty surrounding expected cash flow growth, whereas σ_y measures the stock's fundamental risk. With risk-neutrality, fundamental risk does not influence stock prices.

The critical component of price formation is the investor's *conditional* expectation of future cash flow growth. In our model, realized cash flow growth is

³ The unconditional mean of cash flow growth $\bar{\theta}$ may also be uncertain. In this generalization, the investor learns about this unconditional mean and treats her expectation regarding its value as an additional forecast. We are grateful to an anonymous referee for suggesting this generalization.

uninformative regarding future cash flow growth.⁴ Instead, our investor receives multiple forecasts of future cash flow growth, with each forecast issued by a different information source (such as a sell-side analyst). Information sources issue cash flow growth forecasts for the next period.⁵ Specifically, on date t , the investor observes the forecast μ_t^j for y_{t+1} issued by the j th information source where $j = 1, \dots, J$.

The investor forms their conditional expectation of future cash flow growth by optimally combining the available forecasts into a single aggregate estimate that has the lowest mean-squared forecast error. Intuitively, the investor assigns more weight to forecasts issued by more accurate information sources. The crucial assumption is that the investor does not know the true forecast accuracy of the information sources but learns about their accuracy.

Although the investor uses the cash flow growth forecasts to form their conditional expectation of cash flow growth, they cannot directly evaluate the usefulness of these forecasts because the conditional mean of cash flow growth is unobservable and time-varying. We assume that their conditional expectation $E_t[y_{t+1} | \mu_t^j]$ under the objective probability coincides with the unconditional mean of cash flow growth, which is zero. This ensures that stock return predictability does not arise because observable state variables can predict cash flow growth.

2.1. Optimal Weights and Forecast Accuracy Uncertainty

The investor combines the J cash flow forecasts available at time t into a single aggregate estimate of future cash flow growth

$$\hat{\mu}_t = \sum_{j=1}^J \omega_t^j \mu_t^j, \quad \text{where } \sum_{j=1}^J \omega_t^j = 1. \quad (4)$$

The ω_t^j weights are chosen to minimize the mean-squared forecast error of the aggregate cash flow growth estimate, $E_t[(y_{t+1} - \hat{\mu}_t)^2]$.

Let W_t denote a column vector of weights ω_t^j and ϵ_{t+1} a column vector of forecast errors whose j th component is $\epsilon_{j,t+1} \equiv y_t - \mu_{t-1}^j$. With $y_{t+1} - \hat{\mu}_t = W_t^T \epsilon_{t+1}$ (superscript T denotes matrix transpose), the minimization of $E_t[(y_{t+1} - \hat{\mu}_t)^2]$ is equivalent to the following:

$$\begin{aligned} \min_{W_t} \quad & W_t^T \Theta_t W_t \\ \text{subject to} \quad & \mathbf{1}^T W_t = 1, \end{aligned} \quad (5)$$

⁴This property is consistent with the Chan et al. (2003) finding that realized cash flows are poor predictors of future cash flows. Our investor does not learn about the parameters underlying cash flow growth from realized cash flows, as in Lewellen and Shanken (2002).

⁵Information sources can issue forecasts for a sequence of future cash flows without altering our predictions.

where $\mathbf{1}$ denotes a J -dimensional vector of ones, and $\Theta_t = E_t[\epsilon_{t+1} \epsilon_{t+1}^T]$ is a J by J matrix. W_t summarizes the optimal weights assigned to each information source's cash flow growth forecast at time t . The aggregate cash flow estimate $\hat{\mu}_t$, which combines all available forecasts using their optimal weights, serves as the investor's conditional expectation of cash flow growth. By definition, this aggregate cash flow estimate has the lowest mean-squared forecast error among all other possible estimates.

The information sources are not assumed to issue unbiased forecasts. Thus, the conditional forecast bias $E_t[\epsilon_{j,t+1}]$ may not be zero. With $E_t[\epsilon_{t+1}^2] = \text{Var}_t[\epsilon_{t+1}] + (E_t[\epsilon_{t+1}])^2$, the investor accounts for potential forecast biases in their minimization of mean-squared forecast error. Lim (2001) argues that mean-squared forecast error is the appropriate metric for measuring analyst accuracy.

The minimization in (5) is reminiscent of the Markowitz (1952) minimum variance portfolio. However, our investor optimally combines multiple cash flow forecasts for a single stock into an aggregate cash flow estimate, rather than combining multiple stocks into a portfolio. The solution for the weights in (5) equals

$$W_t = \frac{\Theta_t^{-1} \mathbf{1}}{\mathbf{1}^T \Theta_t^{-1} \mathbf{1}}. \quad (6)$$

The corresponding aggregate cash flow growth estimate equals

$$W_t^T \mu_t = \frac{\mathbf{1}^T \Theta_t^{-1} \mu_t}{\mathbf{1}^T \Theta_t^{-1} \mathbf{1}},$$

where $\mu_t = (\mu_t^1, \dots, \mu_t^j, \dots, \mu_t^J)^T$ is the vector of cash flow growth forecasts.

The optimal weights in (6) are determined by the matrix $\Theta_t = E_t[\epsilon_{t+1} \epsilon_{t+1}^T]$, whose elements are unknown. The investor estimates these elements from the past n forecast errors as follows:

$$\sigma_{j,t}^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_{j,t-i}^2, \quad (7)$$

$$\sigma_{j,k,t} = \frac{1}{n} \sum_{i=1}^n \epsilon_{j,t-i} \epsilon_{k,t-i}. \quad (8)$$

Statistically, (7) equals the mean-squared forecast error of an information source. Intuitively, (7) represents the credibility of the j th forecast, with larger forecast errors reducing an information source's credibility.

The estimation of Θ_t utilizes past forecast errors but not the contemporaneous forecast error ϵ_t . This feature captures a gradual updating of an information source's estimated accuracy, and stems from the uncertainty associated with measuring their skill at forecasting future cash flows.⁶ Although the investor

⁶This parallels the uncertainty surrounding a fund manager's skill. Although fund returns are available daily, the assessment of manager skill is conducted less frequently.

immediately incorporates newly released or revised forecasts into her aggregate cash flow estimate, they revise the estimated accuracy of the information sources less frequently. Information processing costs would also slow the updating of the estimated forecast accuracies. Finally, forecast revisions between earning announcement dates are not accompanied by additional cash flow realizations. Hence, contemporaneous forecast errors are unavailable, and the investor has to rely on past forecast errors in these instances.

Appendix A demonstrates that the optimal weights in (6), when the elements of Θ_t are estimated according to (7) and (8), coincide with the slope coefficients of the following regression:

$$Y = UW + \varepsilon,$$

where Y is the vector of past realized cash flow innovations $\{y_{t-1}, \dots, y_{t-n}\}$, and U is a n by J matrix whose j th column is the vector of past forecasts from the j th information source. Observe that the above regression does not have an intercept and the slope coefficients are required to sum to one.

Our aggregate cash flow estimate can be understood in a Bayesian context. For simplicity, assume the forecasts μ_t^j represent uncorrelated signals regarding θ_{t+1} :

$$\mu_t^j = \theta_{t+1} + \eta_{t+1}^j, \tag{9}$$

where η_{t+1}^j is a mean zero error term. When the investor has a diffuse prior with mean zero for θ_{t+1} , the Bayesian posterior mean is a weighted average of the μ_t^j forecasts whose weights are proportional to each forecast's precision (inverse of variance). This feature is also apparent in the optimal weights defined by (6).

The regression interpretation of the information weights and the Bayesian interpretation of the aggregate cash flow growth estimate both depend on the normality assumption underlying (2) and (3). However, the minimization of mean-squared forecast error does not require any distributional assumptions. Thus, the weights in (6) are optimal without the normality assumption.

2.2. Return Implications of Weight Updating

We now examine the asset pricing implications of our aggregate cash flow estimate. The risk-neutrality of the representative investor and a discount rate equal to δ imply the (ex-dividend) stock price at time t equals

$$P_t = \frac{E_t^I[N_{t+1}]}{1 + \delta} + \frac{E_t^I[N_{t+2}]}{(1 + \delta)^2} + \dots, \tag{10}$$

where $E_t^I[-]$ denotes the investor's date t expectation conditional on the J cash flow growth forecasts. Specifically, $E_t^I[y_{t+1}] = \hat{\mu}_t$. Recall that cash flow growth

is forecasted for the next period. The conditional expectation $E_t^I[y_{t+i}]$ for cash flow growth beyond this horizon, $i > 1$, is equal to zero, its unconditional mean. This implies

$$E_t^I[N_{t+i}] = E_t^I[N_t + y_{t+1} + \dots + y_{t+i}] = N_t + \hat{\mu}_t. \tag{11}$$

The pricing formulation in (11) is similar to Barberis et al. (1998) with a critical distinction. Our aggregate cash flow estimate $\hat{\mu}_t$ results from a combination of cash flow forecasts rather than a single incorrect forecast.

By (10) and (11), the stock price is

$$P_t = \frac{N_t + \hat{\mu}_t}{\delta}, \tag{12}$$

which implies that the simple return between t and $t + 1$ equals

$$R_{t+1} \equiv P_{t+1} - P_t = \frac{y_{t+1} - \hat{\mu}_t}{\delta} + \frac{\hat{\mu}_{t+1}}{\delta}. \tag{13}$$

The realized return over the $(t, t + 1]$ horizon depends on two elements, the realized forecast error, $y_{t+1} - \hat{\mu}_t$, and next period's aggregate cash flow estimate, $\hat{\mu}_{t+1}$.

The expected stock return under the objective probability equals

$$E_t[R_{t+1}] = \frac{E_t[\hat{\mu}_{t+1} - \hat{\mu}_t]}{\delta}. \tag{14}$$

Thus, the expected return is determined by changes in the aggregate cash flow estimate. With $\hat{\mu}_t$ being a weighted average of the individual forecasts, its dynamics depend on changes in the information weights as well as the dynamics of individual forecasts. For expositional simplicity, we assume that⁷

$$E_t[\mu_{t+1}^j] = \mu_t^j. \tag{15}$$

Intuitively, cash flow growth uncertainty causes the information sources to maintain, on average, their existing cash flow growth forecasts over short horizons. This is consistent with the information sources being Bayesians with informative priors regarding the expected cash flow growth rate. It also parallels the Hong et al. (2007a) assumption that investors maintain their prevailing cash flow forecast procedure until there is convincing evidence of its inferiority compared to another forecast procedure.

Under the assumption in (15), stock returns follow a random walk when there is only one information source or the information weights are constant. The information weights are constant if the relative forecast accuracies of the information sources are known.

⁷ Appendix B demonstrates that the results in this section are unchanged when this assumption is relaxed and the individual forecasts are updated according to Bayesian principles.

However, uncertainty surrounding their forecast accuracy implies that the updating of the information weights generates return predictability. Interestingly, this predictability is not attributable to time-varying risk nor behavioral biases.

Specifically, by (14) and (15), expected stock returns in our model are

$$E_t[R_{t+1}] = \frac{\sum_j (\omega_{t+1}^j - \omega_t^j) \mu_t^j}{\delta}. \quad (16)$$

Equation (16) implies that the investor's expected return is proportional to $\text{Cov}(\Delta\omega, \mu)$, the covariance between the cash flow forecasts and changes in their information weights:⁸

$$E_t[R_{t+1}] = \text{Cov}(\Delta\omega, \mu) / \delta = \sigma_{\Delta\omega} \sigma_{\mu} \rho_{\Delta\omega, \mu} / \delta, \quad (17)$$

where $\text{Cov}(\Delta\omega, \mu)$ is computed across the J forecasts, σ_{μ} denotes the cross-sectional dispersion of the forecasts, and the $\sigma_{\Delta\omega}$ component represents the amount of updating in the information weights due to investor learning. As an application of (17), we demonstrate the presence of earnings momentum in our model.

PROPOSITION 1. *Stock prices drift after earnings announcements in the same direction as the earnings surprise. Specifically, the expected stock return next period is positive (negative) conditional on realized earnings growth being above (below) its mean:*

$$\begin{aligned} E_t[R_{t+1} | y_t > \bar{y}] &> 0, \\ E_t[R_{t+1} | y_t < \bar{y}] &< 0. \end{aligned}$$

PROOF. For tractability, we consider two independent information sources. An optimistic information source issues a cash flow growth forecast $\mu_t^O > \bar{y}$ and a pessimistic information source issues a forecast $\mu_t^P < \bar{y}$. By symmetry, we prove Proposition 1 for the case of a positive earning surprise.

By (16), the expected stock return is

$$E_t[R_{t+1}] = (\omega_{t+1}^O - \omega_t^O)(\mu_t^O - \mu_t^P) / \delta.$$

Thus, to prove Proposition 1, it is sufficient to show that

$$E[(\omega_{t+1}^O - \omega_t^O) | y_t > \bar{y}] > 0. \quad (18)$$

The optimal weights in (6), with two information sources, imply that $\omega_{t+1}^O - \omega_t^O$ is proportional to $\sigma_{P,t+1}^2 \sigma_{O,t}^2 - \sigma_{P,t}^2 \sigma_{O,t+1}^2$ where the estimated variances are defined in (7) and satisfy

$$\begin{aligned} \sigma_{O,t+1}^2 &= \sigma_{O,t}^2 + (\epsilon_{O,t}^2 - \epsilon_{O,t-n}^2) / n, \\ \sigma_{P,t+1}^2 &= \sigma_{P,t}^2 + (\epsilon_{P,t}^2 - \epsilon_{P,t-n}^2) / n. \end{aligned}$$

⁸ The cross-sectional mean of $\Delta\omega$ is zero by definition because the information weights sum to one.

We claim that after a positive earning surprise, the estimated accuracy of the optimistic information source on average improves relative to the pessimistic information source. Thus, the information weight for the optimistic information source tends to increase, $\omega_{t+1}^O - \omega_t^O > 0$. This property follows directly from the following inequalities:

$$E[(\epsilon_{O,t}^2 - \epsilon_{O,t-n}^2) | y_t > \bar{y}] < 0, \quad (19)$$

$$E[(\epsilon_{P,t}^2 - \epsilon_{P,t-n}^2) | y_t > \bar{y}] > 0. \quad (20)$$

A positive earnings surprise at date t does not provide useful information about the previous forecast error at $t - n$. Thus, $E[\epsilon_{j,t-n}^2 | y_t > \bar{y}]$ is simply the unconditional second moment of the j th information source's forecast error, and (19) and (20) are equivalent to

$$E[\epsilon_{O,t}^2 | y_t > \bar{y}] < E[\epsilon_{O,t}^2 | y_t < \bar{y}], \quad (21)$$

$$E[\epsilon_{P,t}^2 | y_t > \bar{y}] > E[\epsilon_{P,t}^2 | y_t < \bar{y}]. \quad (22)$$

To prove (21) and (22), observe that each realization $y_t > \bar{y}$ has a one-to-one correspondence with a $y'_t < \bar{y}$ having the same probability density. For the optimistic information source, because $\bar{y} < \mu_t^O$, the forecast errors $\epsilon_{O,t}$ and $\epsilon'_{O,t}$ corresponding to y_t and y'_t respectively satisfy $|\epsilon_{O,t}| < |\epsilon'_{O,t}|$, with (21) following immediately from this property. Conversely, for the pessimistic information source, because $\bar{y} > \mu_t^P$, the forecast errors $\epsilon_{P,t}$ and $\epsilon'_{P,t}$ corresponding to y_t and y'_t respectively satisfy $|\epsilon_{P,t}| > |\epsilon'_{P,t}|$, with (22) following immediately from this property. \square

Proposition 1 proves the existence of earnings momentum (or post-earnings announcement drift) in our model. A related empirical anomaly is price momentum. Chan et al. (1996, 1999) as well as Chordia and Shivakumar (2006) report that a large portion of price momentum occurs around earnings announcements. Campbell and Shiller (1988) demonstrate that stock returns are either attributable to changes in expected discount rates or expected cash flows. With a constant discount rate, stocks returns are attributable to changes in expected cash flow. Thus, price momentum and earnings momentum are closely related in our model.

Simulations investigate the magnitude of price momentum that can arise from changing information weights. The simulations have an initial cash flow $N_0 = 1$, zero unconditional mean cash flow growth ($\bar{\theta}$), and a 3% cash flow growth volatility (σ_y). This σ_y parameter is estimated as the standard deviation of the dividend growth rate for S&P 500 companies, which is historically 3.4% per annum. The σ_{θ} parameter represents the uncertainty surrounding expected cash flow growth. This parameter is chosen to be either 1% or 2%, which is reasonable in comparison to σ_y . In each simulated economy, there are two

cash flow forecasts. The optimistic forecast for period $t + 1$'s cash flow is $N_t + Disp$, whereas the pessimistic forecast is $N_t - Disp$, where $Disp$ measures the dispersion between the two forecasts. Our $Disp$ parameter implies forecast dispersion is approximately 1.5%–3%, which is quite conservative.

For each set of $(\sigma_\theta, Disp)$ parameters, we first simulate 2,000 cash flow and price paths according to (1), (2), (3), and (12). Each simulation path contains 120 monthly time-series observations. Then, for each time period, we rank the 2,000 simulation paths cross-sectionally based on stock returns over the last six periods. The top and the bottom deciles form zero-cost momentum portfolios (winners minus losers). Table 1 reports the average returns of the momentum portfolios during the formation period, as well as over three subsequent holding periods.

Table 1 shows that the price momentum strategy is profitable in our model. Under reasonable parameters, it yields significant profits ranging from 0.68% to 0.94% per month. The magnitude of these profits is consistent with the findings of Jegadeesh and Titman (1993). The reversal of momentum profits at longer holding periods in Table 1 is also consistent with the empirical evidence in Lee and Swaminathan (2000).

The intuition for the momentum results in Table 1 is as follows: In the simulated economy, cash flow growth is equally likely to be positive or negative as

the information sources have identical true accuracies. The price momentum sort identifies paths where there is a recent trend in cash flow growth. For illustration, suppose a sequence of positive cash flow growth realizations occurs by chance. This sequence enhances the optimistic information source's credibility. Therefore, the investor gradually assigns more weight to the optimistic information source. This shift in the information weights increases the aggregate cash flow estimate and leads to a further price increase. Eventually, forecast errors that contradict the earlier estimated accuracies are realized. The investor then updates the relative accuracy of the information sources and reduces the weight assigned to the optimistic information source. This updating causes a decline in the investor's aggregate cash flow estimate and lowers the stock price. Thus, trends in realized cash flows that are attributable to chance produce short-term momentum that reverts over the long-term.

Besides return predictability, fluctuations in the information weights produce additional return volatility. This feature is consistent with the Shiller (1992) assertion that stock return volatility is excessive relative to the volatility of cash flow. However, excess volatility is induced by learning in our model, not irrationality. The Lewellen and Shanken (2002) learning model also yields excess return volatility. From (13), the stock's realized return variance equals⁹

$$\text{Var}_t(R_{t+1}) = \frac{1}{\delta^2} [\sigma_y^2 + \text{Var}_t(\hat{\mu}_{t+1} - \hat{\mu}_t)]. \quad (23)$$

In the absence of forecast accuracy uncertainty, the stock's return variance reduces to σ_y^2/δ^2 . Consequently, forecast accuracy uncertainty leads to variability in the aggregate cash flow estimate that increases return volatility.

2.3. Appearance of Biases

Our framework can explain the *appearance* of behavioral biases that have previously been used to generate momentum, although behavioral biases are not assumed to influence investor expectations.

With two positively correlated information sources, investors focus their attention on the more accurate information source. Indeed, according to (6), the information weights are

$$\begin{aligned} \omega^1 &= \frac{\sigma_2^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}}, \\ \omega^2 &= \frac{\sigma_1^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}}. \end{aligned} \quad (24)$$

The higher the correlation between the two information sources, the higher (lower) the information

Table 1 Simulated Returns from Momentum Strategy

σ_θ	$Disp$	Formation-period return (six periods)	Holding-period returns		
			One period	Three periods	Six periods
0.0100	0.0100	0.3200	0.0068 (7.3081)	-0.0009 (-0.7042)	-0.0094 (-5.7803)
0.0200	0.0100	0.4269	0.0074 (5.9231)	-0.0031 (-1.5030)	-0.0135 (-5.3187)
0.0100	0.0200	0.3321	0.0091 (9.6328)	0.0027 (1.7454)	-0.0041 (-1.9609)
0.0200	0.0200	0.3821	0.0094 (7.4238)	0.0014 (0.8075)	-0.0077 (-3.3060)

Notes. This table presents the average returns of momentum portfolios based on simulations with the following common inputs: initial cash flow $N_0 = 1$, unconditional cash flow growth $\bar{\theta} = 0$, and cash flow volatility $\sigma_y = 0.03$. There are two cash flow forecasts. The optimistic forecast for period $t + 1$'s cash flow is $N_t + Disp$, whereas the pessimistic forecast is $N_t - Disp$, where $Disp$ measures the dispersion of the two forecasts. Each row corresponds to a set of simulations using these inputs along with the specified σ_θ (uncertainty surrounding expected cash flow growth) and $Disp$ parameters. For each set of parameters, we simulate 2,000 cash flows and prices, with each time-series containing 120 observations. For each observation, the 2,000 simulation paths are ranked cross-sectionally based on their cumulative returns over the last six periods. The momentum portfolio is the top decile minus and the bottom decile. The table reports the time-series average return of the momentum portfolios during the formation period, as well as over subsequent holding periods ranging from one to six periods. The mean returns for each holding period are recorded below with t -statistics in parentheses.

⁹ Note that $\text{Cov}_t(y_{t+1}, \hat{\mu}_{t+1} - \hat{\mu}_t) = 0$ because $E_t[y_{t+1}] = 0$ and (15).

weight assigned to the more (less) accurate information source. Therefore, a high positive covariance between the forecasts can effectively eliminate the less accurate information source. This feature leads to the appearance of limited attention, which is studied by Hirshleifer and Teoh (2003) as well as Peng and Xiong (2006).

Conversely, a negatively correlated forecast can receive a larger information weight than its accuracy alone justifies. When an investor's private cash flow forecast is negatively correlated with the consensus forecast of analysts, they can appear overconfident as in Daniel et al. (1998). The appearance of overconfidence also arises when the estimated accuracy of the investor's private cash flow forecast is superior to available public forecasts.

Path dependence in the estimated forecast accuracies is responsible for the appearance of representativeness and conservatism. These biases are utilized by Barberis et al. (1998). With the information weights being path-dependent, trends in realized cash flow growth change the estimated relative accuracy of information sources. Thus, the investor's aggregate cash flow forecast appears to extrapolate from realized cash flows. Furthermore, the impact of trends on the information weights persists beyond their termination due to the path dependence in (7). This property causes the information weights to exhibit conservatism.

3. Empirical Implementation

Our empirical implementation tests our model's implications regarding stock return predictability, earnings momentum, and price momentum using data on analysts' earnings forecasts and realized earnings from the Institutional Brokers' Estimate System (I/B/E/S). Estimation of the information weights, earnings surprises, and analyst forecast dispersion requires individual analyst forecasts. These forecasts begin in January 1984 of the I/B/E/S detail file. Thus, our sample period is from January 1984 to December 2007. The sample includes all domestic common stocks listed on the NYSE, AMEX, and NASDAQ that have at least two analyst forecasts, excluding REITs, ADRs, and stocks priced below \$5. We obtain daily and monthly stock returns as well as market capitalization data from CRSP, and book-to-market ratios and the earnings announcement dates from Compustat.

There is significant multicollinearity among the analyst forecasts. To circumvent this problem, we classify individual analysts into two groups: optimistic (those with forecasts above the median) and pessimistic (those with forecasts below the median). We then compute the average of each subset and refer to these averages as the representative optimistic

analyst and representative pessimistic analyst. The information weights of these representative analysts are computed using their forecast errors over the past eight quarters ($n = 8$). From these weights, we construct a variable denoted dW , measured quarterly for each stock, which represents changes in the weight of the representative optimistic analyst. A positive dW implies a shift toward the representative optimistic forecast, whereas a negative dW implies the representative pessimistic analyst has gained more influence on the investor's cash flow expectation.

3.1. Testable Hypotheses

We test the following predictions of our model.

PREDICTION 1. *A positive (negative) change in the optimistic analyst's weight dW is associated with higher (lower) stock returns next month.*

By definition, when the weight assigned to the optimistic analyst increases at the expense of the pessimistic analyst, the correlation $\rho_{\Delta\omega, \mu}$ is positive. Prediction 1 follows immediately from (17).

PREDICTION 2. *Momentum is stronger for stocks experiencing greater fluctuations in their information weights.*

This prediction follows from (17), which demonstrates that momentum profits increase with $\sigma_{\Delta\omega}$.

PREDICTION 3. *Momentum is stronger for stocks that have larger forecast dispersions.*

This prediction follows from (17), which demonstrates that momentum profits increase with σ_{μ} . Because forecast dispersion measures cash flow uncertainty, our model predicts stronger momentum for stocks with high cash flow uncertainty.

3.2. Empirical Results

To test Prediction 1, each quarter we sort stocks into five dW quintiles around earnings announcements. Stocks are held for one month after portfolio formation. For the top dW decile, where the optimistic analysts are gaining weight, the average stock return is 1.67% per month over our sample period. In contrast, the average return of the bottom decile dW portfolio is 0.99% per month. The difference of 0.68% is significant, and continues to be significant after adjusting for the Fama-French (1993) three factors as well as the Pástor-Stambaugh (2003) liquidity factor. This four-factor adjustment pertains to the January 1984 to December 2006 period because the liquidity factor is only available until December 2006.

As expected, there is more updating in the information weights after larger earnings surprises. As detailed in the proof of Proposition 1, on average, positive earnings surprises (hence positive returns) cause relatively optimistic information sources to receive

Table 2 Price Momentum Conditional on Information Weight Change

	P5 – P1 return spread over past six months	Holding-period returns					Raw returns		Four-factor adjusted	
		P1	P2	P3	P4	P5	P5 – P1	<i>t</i> -stat.	P5 – P1	<i>t</i> -stat.
Panel A: Sorted on MOM first, then on dW										
dW1	90.85	0.61	1.04	1.10	0.81	1.41	0.80	1.60	0.97	1.71
dW2	80.92	0.56	1.24	0.38	0.55	1.05	0.49	0.89	0.44	0.70
dW3	74.34	0.77	0.88	1.29	0.28	0.81	0.04	0.07	–0.04	–0.06
dW4	79.54	1.28	1.44	1.47	1.59	1.85	0.57	1.11	0.53	0.93
dW5	89.63	1.11	1.55	1.43	1.96	2.28	1.17	2.20	1.20	1.97
dW1 – dW3							0.76	2.05	1.01	1.98
dW5 – dW3							1.13	2.69	1.24	2.59
Panel B: Sorted on dW first, then on MOM										
dW1	84.19	0.93	1.20	1.08	0.80	1.56	0.63	1.34	0.93	1.67
dW2	75.36	0.57	0.71	1.06	0.25	0.91	0.35	0.66	0.27	0.20
dW3	74.59	0.83	1.09	0.52	0.68	0.94	0.12	0.20	0.27	0.40
dW4	72.19	1.13	1.19	1.41	1.54	1.64	0.51	1.20	0.40	0.84
dW5	87.76	1.03	1.71	1.56	2.06	2.08	1.15	2.03	1.21	2.05
dW1 – dW3							0.51	1.67	0.66	1.59
dW5 – dW3							1.03	2.60	0.94	2.41

Notes. This table summarizes price momentum conditional on the amount of updating in the information weights. This updating is denoted dW and equals the representative optimistic analyst's weight change from the previous earnings announcement. At the end of each month from January 1984 to December 2007, stocks from the intersection of the CRSP and IBES data sets are ranked on their returns over the past six months and dW. Stocks are then assigned to momentum quintiles (P1 to P5) and dW quintiles (dW1 to dW5) in ascending order. The dW1 portfolio contains stocks whose weight is shifting from the representative optimistic analyst toward the representative pessimistic analyst, whereas the dW5 portfolio contains stocks whose weight is shifting from the representative pessimistic analyst toward the representative optimistic analyst. Price momentum is the zero-cost portfolio that buys P5 and sells P1 every month, implemented within each dW portfolio. Panel A reports price momentum when stocks are first sorted on past returns, then on dW. Panel B reports the results when the sorting order is reversed. Raw returns and those adjusted by the Fama-French (1993) three factors and the Pastor-Stambaugh (2003) liquidity factor are both provided.

more weight whereas negative earnings surprises (hence negative returns) cause relatively pessimistic information sources to receive more weight. This feature is also supported by the data. For stocks with positive earnings surprises, 65.4% have a positive dW. Conversely, for stocks with negative earnings surprises, 72.6% have a negative dW. An earnings surprise is measured as the difference between a firm's actual earnings and the prevailing consensus analyst forecast, scaled by the consensus forecast. To exclude stale information, we include only the latest forecast issued by each analyst. These forecasts are required to be issued within one year prior to an earnings announcement.

Prediction 2 is tested using price momentum and earnings momentum. We implement a 6-1-1 price momentum strategy. At the end of every month, stocks are assigned to five quintiles (P1 to P5) in ascending order according to their returns over the prior six months. After skipping one month, the momentum portfolios are held for an additional month. Price momentum is computed as the average difference between the holding-period return of the P5 portfolio (past winners) and the P1 portfolio (past losers).

Earnings momentum parallels the price momentum strategy. Instead of sorting stocks according to

their past returns, earnings momentum portfolios (E1 to E5) are formed according to their most recent earnings surprise. Earnings momentum is the average return difference between E5 (positive surprises) and E1 (negative surprises) over the monthly holding periods.

To test Prediction 2, we compare the average returns of double-sorted portfolios formed using the past six-month returns or earnings surprises and dW. Over the cross-section of stocks, dW is positively correlated with formation-period returns and earnings surprises.¹⁰ Thus, we perform conditional double-sorts in both directions. For example, in panel A of Table 2, we first sort stocks according to their returns over the prior six months, with a second sort conditioning on dW. Conversely, in panel B of Table 2, the sorting order is reversed. The results for Prediction 2 do not depend on the order of the conditional double-sorts.

The empirical results in Table 2 indicate stronger price momentum in the dW1 portfolio and dW5 portfolio, relative to the dW3 portfolio. For example, panel A reports that the price momentum strategy generates an average monthly return of 0.80%

¹⁰ The time-series average of the monthly cross-sectional correlation between dW and earnings surprises is 0.20, whereas the correlation between dW and formation-period returns is 0.13.

Table 3 Earnings Momentum Conditional on Information Weight Change

	E5 – E1 return spread over past six months	Holding-period returns					Raw returns		Four-factor adjusted	
		E1	E2	E3	E4	E5	E5 – E1	<i>t</i> -stat.	E5 – E1	<i>t</i> -stat.
Panel A: Sorted on MOM first, then on dW										
dW1	27.83	-0.05	0.26	0.03	0.09	1.41	1.46	2.96	1.61	2.91
dW2	31.22	-0.04	-0.04	0.20	0.70	0.81	0.85	1.78	1.26	2.37
dW2	24.77	0.45	0.57	0.61	1.26	1.16	0.71	1.67	1.03	2.05
dW4	26.92	0.44	0.06	0.33	1.17	0.98	0.54	1.26	0.51	1.09
dW5	25.62	0.21	0.26	0.51	0.95	1.39	1.18	2.69	1.42	2.85
dW1 – dW3							0.75	1.92	0.58	1.46
dW5 – dW3							0.47	1.49	0.39	1.38
Panel B: Sorted on dW first, then on MOM										
dW1	30.63	0.12	0.06	0.73	0.93	1.32	1.19	3.39	0.97	2.41
dW2	26.75	0.34	0.16	-0.05	0.63	1.24	0.91	1.76	1.12	1.93
dW3	27.84	0.22	0.10	0.24	0.27	0.85	0.63	1.54	0.58	1.39
dW4	28.09	0.52	0.36	-0.03	1.34	1.22	0.69	1.88	0.46	1.08
dW5	26.28	0.45	0.41	0.52	0.78	1.39	0.93	1.86	1.21	2.44
dW1 – dW3							0.56	1.77	0.39	1.58
dW5 – dW3							0.30	1.19	0.63	1.29

Notes. This table summarizes earnings momentum conditional on the amount of updating in the information weights. This updating is denoted dW and equals the representative optimistic analyst's weight change from the previous earnings announcement. At the end of each month from January 1984 to December 2007, stocks from the intersection of the CRSP and IBES data sets are ranked according to their most recent earnings surprise and dW. Stocks are then assigned to earnings surprises quintiles (E1 to E5) and dW quintiles (dW1 to dW5) in ascending order. The dW1 portfolio contains stocks whose weight is shifting from the representative optimistic analyst toward the representative pessimistic analyst, whereas the dW5 portfolio contains stocks whose weight is shifting from the representative pessimistic analyst toward the representative optimistic analyst. Earnings momentum is the zero-cost portfolio that buys E5 and sells E1 every month, implemented within each dW portfolio. Panel A reports earnings momentum when the stocks are first sorted on earnings surprises, then on dW. Panel B reports the results when the sorting order is reversed. Raw returns and those adjusted by the Fama-French (1993) three factors and the Pástor-Stambaugh (2003) liquidity factor are both provided.

and 1.17% for stocks in the dW1 and dW5 portfolios respectively, but only 0.04% for the dW3 portfolio. The difference in price momentum between dW1 and dW3 is 0.76% per month, with a *t*-statistic of 2.05. The difference in price momentum between dW5 and dW3 is 1.13% per month, with a *t*-statistic of 2.69. Therefore, consistent with Prediction 2, greater updating in the information weights leads to stronger price momentum. This pattern appears in raw returns as well as risk-adjusted returns that account for the Fama-French (1993) three factors and the liquidity factor of Pástor and Stambaugh (2003).

The earnings momentum results in Table 3 exhibit a similar pattern as price momentum. In comparison to the dW3 portfolio, earnings momentum is stronger in the dW1 portfolio and the dW5 portfolio. Therefore, consistent with Prediction 2, earnings momentum also depends on the amount of updating in the information weights.

For emphasis, our study is limited to firms with at least two analysts. Thus, our sample is orientated toward more established firms and those with greater analyst coverage. Hong et al. (2000) document weaker price momentum in larger stocks and stocks with greater analyst coverage. Our results confirm this finding. However, even for firms with at least two analysts, our refined price momentum strategy that

conditions on fluctuations in the information weights can produce high profits. The results in Table 2 (panel A) indicate that past winners, which experience large increases in the optimistic analyst's weight, and past losers, which experience large decreases in the optimistic analyst's weight, produce an average monthly return spread of 1.67%. This return spread equals the difference between the average return of the (P5, dW5) portfolio and the (P1, dW1) portfolio. Consequently, our enhanced momentum return is larger than the 1% monthly return from the standard momentum strategy that does not condition on information weight fluctuations.

Finally, we test Prediction 3 by comparing the profitability of price momentum and earnings momentum across analyst forecast dispersion quintiles. At the end of each month, forecast dispersion is measured as the standard deviation of all forecasts issued during the past year for earnings in the current fiscal year. This standard deviation is then scaled by the consensus forecast.¹¹ Stocks are then sorted into quintiles (U1 to U5), with U1 containing stocks with the lowest forecast dispersion and U5 containing stocks with the largest forecast dispersion.

¹¹ Similar results are obtained when we scale by the stock price at the end of the prior year.

Table 4 Price and Earnings Momentum Conditional on Forecast Dispersion

	P5 – P1 return spread over past six months	Holding-period returns					Raw returns		Four-factor adjusted	
		P1	P2	P3	P4	P5	P5 – P1	<i>t</i> -stat.	P5 – P1	<i>t</i> -stat.
Panel A: Price momentum, sorted on MOM first, then on uncertainty										
U1	73.38	0.82	1.31	1.38	1.41	1.52	0.70	2.23	0.73	1.86
U2	76.47	0.77	1.26	1.33	1.33	1.55	0.77	2.77	0.82	2.53
U3	75.93	0.69	1.39	1.34	1.22	1.65	0.96	3.49	0.92	2.89
U4	80.53	0.72	1.27	1.33	1.12	1.77	1.05	3.67	1.04	3.14
U5	85.98	0.22	1.00	1.23	1.17	1.90	1.68	5.48	1.55	4.31
U5 – U1							0.98	4.28	0.82	3.31
Panel B: Price momentum, sorted on uncertainty first, then on MOM										
U1	78.55	1.18	1.28	1.34	1.42	1.94	0.76	2.76	0.61	1.97
U2	69.36	0.80	1.30	1.24	1.20	1.54	0.73	2.73	0.58	1.89
U3	72.18	0.86	1.41	1.32	1.15	1.72	0.85	3.31	0.83	2.80
U4	77.62	0.70	1.13	1.23	1.25	1.61	0.91	3.38	0.73	2.36
U5	83.22	0.10	0.83	1.06	1.19	1.75	1.65	5.19	1.42	3.83
U5 – U1							0.89	4.19	0.81	2.93
	E5 – E1 return spread over past six months	Holding-period returns					Raw returns		Four-factor adjusted	
		E1	E2	E3	E4	E5	E5 – E1	<i>t</i> -stat.	E5 – E1	<i>t</i> -stat.
Panel C: Earnings momentum, sorted on MOM first, then on uncertainty										
U1	29.39	0.95	1.08	1.32	1.51	1.66	0.71	3.26	0.82	2.88
U2	25.11	0.93	1.10	1.13	1.34	1.63	0.70	3.12	0.73	2.98
U3	26.42	0.83	1.19	1.17	1.31	1.71	0.87	3.57	0.90	3.31
U4	27.61	0.78	1.03	1.26	1.31	1.70	0.92	3.21	0.93	2.71
U5	25.22	0.39	1.06	1.25	1.38	1.68	1.29	4.50	1.44	4.71
U5 – U1							0.58	3.27	0.62	3.53
Panel D: Earnings momentum, sorted on uncertainty first, then on MOM										
U1	18.94	1.05	1.14	1.37	1.49	1.78	0.73	2.95	0.83	3.05
U2	21.33	0.93	1.06	1.24	1.29	1.51	0.59	2.44	0.65	2.31
U3	23.89	0.91	1.10	1.22	1.25	1.76	0.85	4.12	0.80	2.93
U4	24.10	0.84	1.13	1.04	1.35	1.71	0.87	3.67	0.92	3.12
U5	26.27	0.50	0.71	0.94	1.17	1.70	1.20	4.12	1.39	4.15
U5 – U1							0.47	2.75	0.56	3.09

Notes. This table describes the profitability of price momentum and earnings momentum conditional on analyst forecast dispersion. At the end of each month from January 1984 to December 2007, stocks from the intersection of the CRSP and IBES data sets are ranked on either their returns over the past six months or their most recent earnings surprises, along with their prevailing forecast dispersion. Stocks are then assigned to either past return quintiles (P1 to P5) or earnings surprise quintiles (E1 to E5) along with uncertainty quintiles (U1 to U5). Price momentum is the zero-cost portfolio that buys P5 and sells P1 every month. Earnings momentum replaces the past return quintiles with the most recent earnings surprises (E1 to E5). Panel A reports price momentum when stocks are first sorted on past returns, then dispersion, whereas panel B reverses the order of the double-sort. Panels C and D record our results for earnings momentum rather than price momentum. Raw returns and those adjusted by the Fama-French (1993) three factors and the Pastor-Stambaugh (2003) liquidity factor are both provided.

Table 4 shows that price momentum and earnings momentum both monotonically increase from the U1 portfolio to the U5 portfolio. Price momentum among stocks with high analyst forecast dispersion is about 1% higher per month than price momentum among stocks with low analyst forecast dispersion (U5 – U1). Similarly, earnings momentum among stocks with high analyst forecast dispersion is about 0.5% higher per month than earnings momentum among stocks with low analyst forecast dispersion. These differences are both statistically and economically significant. Therefore, the results in Table 4 support Prediction 3 as stocks with greater analyst forecast dispersion have stronger momentum.

3.3. Other Predictions

Our model also has a number of predictions for price momentum and earnings momentum that are consistent with previous empirical findings.

PREDICTION 4. *Momentum is stronger for small firms, young firms, and firms whose fundamentals are undergoing significant changes.*

Young firms (IPOs) and small firms have fewer available forecast errors to estimate the accuracy of each information source (small n in (7)). Thus, each additional forecast error exerts a greater impact on the investor's learning process and induces more dramatic fluctuations in the information weights (larger $\sigma_{\Delta\omega}$). Consistent with this prediction, Jiang

et al. (2005) and Zhang (2006) report that young firms and small firms exhibit stronger momentum.

More established firms have a larger number of forecast errors available for estimating each information source's accuracy. However, when significant firm-specific, industry, and macroeconomic shocks occur, their cash flow implications may not be immediately understood and agreed upon by market participants (Brav and Heaton 2002). Instead, these shocks can increase forecast accuracy uncertainty. Thus, we predict stronger momentum in stocks whose fundamentals are undergoing significant changes. Consistent with this prediction, Jackson and Johnson (2006) document that momentum is concentrated around seasoned equity offerings, stock re-purchases, equity-financed mergers, and dividend initiations as well as omissions.

PREDICTION 5. *Momentum is stronger for stocks with higher return volatility and for stocks with higher cash flow volatility.*

This prediction follows from the fact that fluctuations in the information weights create momentum and increase return volatility. Thus, stronger momentum coincides with periods of higher return volatility. Jiang et al. (2005) and Zhang (2006) report that stocks with more volatile returns have stronger momentum. However, they interpret their results as evidence that behavioral biases influence asset prices, whereas our framework offers an alternative interpretation.

Greater momentum for stocks with higher cash flow volatility is also predicted, although momentum profits are not compensation for cash flow risk in our model. As firms with high cash flow volatility (large σ_y) tend to have high forecast dispersion (large σ_μ), this prediction is subsumed by Prediction 2 and not tested.

4. Conclusions

We study stock prices in a simple model where the investor optimally combines multiple cash flow forecasts of unknown accuracy. The weights assigned to these forecasts depend on the accuracy of their issuer, which the investor estimates from past forecast errors. We demonstrate that earnings momentum and price momentum arise from these weights being updated as the investor gradually learns about the relative accuracy of the information sources issuing forecasts. Return predictability in our model is not caused by time-varying risk premia or behavioral biases.

Empirical tests provide strong support for the model because changes in the information weights predict stock returns and affect the profitability of earnings momentum and price momentum strategies. Simulation evidence confirms that under reasonable parameter values, our framework produces momen-

tum whose magnitude is comparable with existing empirical evidence.

Our framework offers several interesting applications for future research. For example, our aggregate cash flow estimate may improve upon the consensus earnings forecast (simple average) used in prior research. Our information weights can also be applied to better understand investor decisions to switch between investment styles and style momentum.

Acknowledgments

The authors thank Aydogan Altı, Doron Avramov, Michael Brennan, Bhagwan Chowdhry, Phil Dybvig, David Feldman, Lorenzo Garlappi, Ilan Guedj, John Griffin, Bruce Grundy, Jeffrey Hales, David Hirshleifer, Steve Hogan, Harrison Hong, Jennifer Huang, Chuan-Yang Hwang, Ravi Jagannathan, Paul Malatesta, Lubos Pástor, Francisco Perez-Gonzalez, Ramesh Rao, Mark Seasholes, Avanidhar Subrahmanyam, Rossen Valkanov, Jiang Wang, Wei Mun Wang, Guojun Wu, and Jun Yu for their helpful comments and suggestions as well as seminar participants at Cambridge University, National University of Singapore, Singapore Management University, University of Texas at Austin, and the annual meetings of the European Finance Association and the China International Conference in Finance.

Appendix A. Regression Interpretation of the Optimal Weights

The linear regression $\mathbf{Y} = \mathbf{U}W + \varepsilon$ minimizes the following sum of squared residuals

$$(\mathbf{Y} - \mathbf{U}W)^T(\mathbf{Y} - \mathbf{U}W), \quad (\text{A1})$$

by selecting W . This $J \times 1$ matrix of information weights is often denoted β , whereas \mathbf{U} represents a $n \times J$ matrix of cash flow forecasts that is usually denoted X when minimizing $(y - X\beta)^T(y - X\beta)$. In our context, the columns of \mathbf{U} are forecasts issued by an information source during the previous n periods, whereas the $n \times 1$ vector \mathbf{Y} denotes the corresponding realized cash flows.

Inserting the constraint $\mathbf{1}^T W = 1$ into (A1) yields

$$\begin{aligned} & (\mathbf{Y}\mathbf{1}^T W - \mathbf{U}W)^T(\mathbf{Y}\mathbf{1}^T W - \mathbf{U}W) \\ &= [(\mathbf{Y}\mathbf{1}^T - \mathbf{U})W]^T(\mathbf{Y}\mathbf{1}^T - \mathbf{U})W \\ &= W^T(\mathbf{Y}\mathbf{1}^T - \mathbf{U})^T(\mathbf{Y}\mathbf{1}^T - \mathbf{U})W \\ &= nW^T\Theta W \end{aligned}$$

because $\mathbf{Y}\mathbf{1}^T - \mathbf{U}$ is an $n \times J$ matrix of forecast errors and $(1/n)(\mathbf{Y}\mathbf{1}^T - \mathbf{U})^T(\mathbf{Y}\mathbf{1}^T - \mathbf{U})$ equals the Θ matrix whose elements are estimated using (7) and (8). Therefore, the objective function in (5) is identical to linear regression with a zero intercept and coefficients that sum to one.

Appendix B. Alternative Forecast Dynamics

This appendix relaxes the assumption that the cash flow forecasts are, on average, constant over short-horizons. Consider three scenarios defined by the realized cash flow y_{t+1} : (1) $y_{t+1} > \mu_t^O$, (2) $y_{t+1} < \mu_t^P$, and (3) $\mu_t^P < y_{t+1} < \mu_t^O$, where μ_t^O and μ_t^P denote the optimistic and pessimistic cash flow forecast respectively. In the first scenario, the optimistic and pessimistic information source both increase their cash flow forecasts at $t + 1$ because y_{t+1} is higher

than μ_i^O and μ_i^P . These increases are consistent with both information sources being Bayesians. The updated forecasts, $\mu_{i+1}^O > \mu_i^O$ and $\mu_{i+1}^P > \mu_i^P$ imply the first scenario generates the $\hat{\mu}_{i+1} > \hat{\mu}_i$ relationship. Similarly, in the second scenario, the optimistic and pessimistic information source both decrease their cash flow forecasts because y_{i+1} is lower than μ_i^O and μ_i^P . This updating of the individual forecasts yields $\hat{\mu}_{i+1} < \hat{\mu}_i$.

For the third scenario, the individual cash flow forecasts are updated as follows

$$\begin{aligned} \mu_{i+1}^O &= \mu_i^O + \phi_i^O + \varepsilon_{i+1}^O, \\ \mu_{i+1}^P &= \mu_i^P + \phi_i^P + \varepsilon_{i+1}^P, \end{aligned}$$

where ε_{i+1}^O and ε_{i+1}^P are both mean zero, and

$$\frac{\phi_i^O}{\phi_i^P} = -\frac{\sigma_{O,t+1}^2}{\sigma_{P,t+1}^2}. \tag{B1}$$

The condition in (B1) states that the (absolute) amount by which the optimistic (pessimistic) information source updates their forecast downward (upward) is proportional to their estimated accuracy in (7). Therefore, the relatively less accurate information source updates their cash flow forecast more dramatically in the direction of the realized cash flow, with (B1) being consistent with Bayesian updating.

According to (16), the expected stock return $E_t[R_{t+1}]$ is determined by the expected change in the aggregate cash flow $E_t[\hat{\mu}_{t+1}] - \hat{\mu}_t$ which equals

$$\begin{aligned} &(\omega_i^O + \Delta\omega_i^O)(\mu_i^O + \phi_i^O) + (\omega_i^P + \Delta\omega_i^P)(\mu_i^P + \phi_i^P) - \omega_i^O \mu_i^O - \omega_i^P \mu_i^P \\ &= \Delta\omega_i^O \mu_i^O + \Delta\omega_i^P \mu_i^P + \omega_i^O \phi_i^O + \Delta\omega_i^O \phi_i^O + \omega_i^P \phi_i^P + \Delta\omega_i^P \phi_i^P. \end{aligned} \tag{B2}$$

To demonstrate that momentum occurs in the third scenario, consider two cases. First, after positive cash flow growth realizations, the optimistic information source becomes more accurate, and is assigned a larger weight, $\Delta\omega_i^O > 0$. The first two terms on the right side of (B2), $\Delta\omega_i^O \mu_i^O + \Delta\omega_i^P \mu_i^P$, are positive when combined because they equal

$$\begin{aligned} &\Delta\omega_i^O \mu_i^O + \Delta\omega_i^P \mu_i^P + \Delta\omega_i^O (\mu_i^O - \mu_i^P) \\ &= \mu_i^P (\Delta\omega_i^O + \Delta\omega_i^P) + \Delta\omega_i^O (\mu_i^O - \mu_i^P) \\ &= \Delta\omega_i^O (\mu_i^O - \mu_i^P) > 0. \end{aligned} \tag{B3}$$

The above inequality follows from the optimistic information source being assigned a larger portfolio weight, $\Delta\omega_i^O > 0$, whereas the $\mu_i^O - \mu_i^P > 0$ and $\Delta\omega_i^O + \Delta\omega_i^P = 0$ properties hold by definition. The remaining terms in (B2) sum to zero

$$\omega_i^O \phi_i^O + \Delta\omega_i^O \phi_i^O + \omega_i^P \phi_i^P + \Delta\omega_i^P \phi_i^P = \omega_{i+1}^O \phi_i^O + \omega_{i+1}^P \phi_i^P = 0,$$

due to the condition in (B1) and $\omega_{i+1}^O/\omega_{i+1}^P = \sigma_{P,t+1}^2/\sigma_{O,t+1}^2$. Thus, $E_t[\hat{\mu}_{i+1}] > \hat{\mu}_i$ after positive cash flow growth.

The second case involves negative cash flow growth realizations, which cause the weight of the pessimist to increase at the optimist's expense. In other words, $\Delta\omega_i^O > 0$ and $\Delta\omega_i^P < 0$. Consequently, the first two terms on the right side of (B2), $\Delta\omega_i^O \mu_i^O + \Delta\omega_i^P \mu_i^P$, are negative when combined, whereas the remaining terms again sum to zero. Thus, $E_t[\hat{\mu}_{i+1}] < \hat{\mu}_i$ after negative cash flow growth.

To summarize, expected returns are higher (lower) following a series of positive (negative) cash flow growth realizations when the individual cash flow forecasts are updated according to Bayesian principles.

References

Barberis, N., A. Shleifer, R. Vishny. 1998. A model of investor sentiment. *J. Financial Econom.* **49** 307–343.

Brav, A., J. B. Heaton. 2002. Competing theories of financial anomalies. *Rev. Financial Stud.* **15** 575–606.

Brown, L. D. 2001. How important is past analyst forecast accuracy? *Financial Analysts J.* **57** 44–49.

Campbell, J. Y., R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Rev. Financial Stud.* **1** 195–228.

Chan, L. K. C., N. Jegadeesh, J. Lakonishok. 1996. Momentum strategies. *J. Finance* **51** 1681–1713.

Chan, L. K. C., N. Jegadeesh, J. Lakonishok. 1999. The profitability of momentum strategies. *Financial Analysts J.* **55** 80–90.

Chan, L. K. C., J. Karceski, J. Lakonishok. 2003. The level and persistence of growth rates. *J. Finance* **58** 643–684.

Chordia, T., L. Shivakumar. 2006. Earnings and price momentum. *J. Financial Econom.* **80** 627–656.

Clement, M. B., S. Y. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *Accounting Rev.* **78** 227–249.

Daniel, K., S. Titman. 2006. Market reactions to tangible and intangible information. *J. Finance* **61** 1605–1643.

Daniel, K., D. Hirshleifer, A. Subrahmanyam. 1998. Investor psychology and security market under- and overreactions. *J. Finance* **53** 1839–1885.

Fama, E., K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *J. Financial Econom.* **33** 3–56.

Hirshleifer, D., S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *J. Accounting Econom.* **36** 337–386.

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

Hong, H., T. Lim, J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *J. Finance* **55** 265–295.

Hong, H., J. C. Stein, J. Yu. 2007a. Simple forecasts and paradigm shifts. *J. Finance* **62** 1207–1242.

Hong, H., W. Torous, R. Valkanov. 2007b. Do industries lead the stock market? Gradual diffusion of information and cross-asset return predictability. *J. Financial Econom.* **83** 367–396.

Jackson, A., T. Johnson. 2006. Unifying underreaction anomalies. *J. Bus.* **79** 75–114.

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

Jiang, G., C. M. C. Lee, Y. Zhang. 2005. Information uncertainty and expected returns. *Rev. Accounting Stud.* **10** 185–221.

Lee, C. M. C., B. Swaminathan. 2000. Price momentum and trading volume. *J. Finance* **55** 2017–2069.

Lewellen, J., J. Shanken. 2002. Learning, asset pricing tests, and market efficiency. *J. Finance* **57** 1113–1145.

Lim, T. 2001. Rationality and analysts' forecast bias. *J. Finance* **56** 369–385.

Markowitz, H. M. 1952. Portfolio selection. *J. Finance* **7** 77–91.

Pástor, L., R. F. Stambaugh. 2003. Liquidity risk and expected stock returns. *J. Political Econom.* **111** 642–685.

Peng, L., W. Xiong. 2006. Limited attention, overconfidence and category learning. *J. Financial Econom.* **80** 563–602.

Shiller, R. J. 1992. *Market Volatility*. MIT Press, Cambridge, MA.

Sinha, P., L. D. Brown, S. Das. 1997. A re-examination of financial analysts' differential earnings forecast accuracy. *Contemporary Accounting Res.* **14** 1–42.

Timmermann, A. G. 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *Quart. J. Econom.* **108** 1135–1145.

Zhang, X. F. 2006. Information uncertainty and stock returns. *J. Finance* **61** 105–136.