3-23-2009

Forecast Accuracy Uncertainty and Momentum

Bing Han
University of Texas at Austin

Dong Hong
Singapore Management University

Mitch Warachka
Chapman University, warachka@chapman.edu

Recommended Citation

Bing Han, Dong Hong, Mitch Warachka. Forecast Accuracy Uncertainty and Momentum. Management Science 55 (6) 1035-1046 https://doi.org/10.1287/mnsc.1080.0992

Forecast Accuracy Uncertainty and Momentum *

Bing Han[†], Dong Hong[‡] and Mitch Warachka[§]

August 2008

*We thank Aydogan Alti, 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.

†University of Texas at Austin; Address: McCombs School of Business, 1 University Station - B6600 Austin, TX 78712; Email: bing.han@mccombs.utexas.edu, Phone: (512) 232-6822, Fax: (512) 471-5073.

‡Singapore Management University; Address: #4056 L.K.C. School of Business, 50 Stamford Road, 178899, Singapore; Email: donghong@smu.edu.sg, Phone: (65) 6828 0744, Fax: (65) 6828 0427.

§Singapore Management University; Address: #4055 L.K.C. School of Business, 50 Stamford Road, 178899, Singapore; Email: mitchell@smu.edu.sg, Phone: (65) 6828 0249, Fax: (65) 6828 0427.

Abstract

We demonstrate that stock price momentum and earnings momentum can result from uncertainty surrounding the accuracy of cashflow forecasts. Our model has multiple information sources issuing cashflow forecasts for a stock. The investor combines these forecasts into an aggregate cashflow estimate that has minimal mean-squared forecast error. This aggregate estimate weights each cashflow 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.

JEL Classification: G12, G14

I 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 cashflow growth but is *uncertain* about the *accuracy* of the information sources issuing these forecasts. The investor optimally combines the forecasts into an aggregate cashflow 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 cashflow estimate represents the investor's expectation of future cashflow 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 cashflow 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 cashflow. Investors understand the uncertainty inherent in measuring this skill and gradually update their assessment of each information source's accuracy.[2]

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 cashflow 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 cashflow growth forecasts are, on average, constant over short horizons.

Although our investor immediately incorporates newly issued or revised forecasts into their conditional cashflow expectation, the weights assigned to these forecasts are gradually updated.[3] The gradual updating of the information weights generates return predictability. In particular, earnings momentum and price momentum arise from learning about the relative

[1] Sinha, Brown, and Das (1997), Brown (2001), and Clement and Tse (2003) find that prior forecast errors predict the future forecast accuracy of analysts.

[2] Section A. contains additional justification for the gradual updating of the forecast accuracy estimates.

[3] This gradual updating is distinct from the slow diffusion of information in Hong and Stein (1999) and Hong, Torous, and Valkanov (2007).

forecast accuracy of the information sources. For example, after a series of positive cashflow innovations, hence price increases, 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 cashflow estimate. This shift in the information weights leads to higher expected cashflow growth and a higher stock price, although the individual forecasts remain unchanged (on average).

Momentum in our framework does not originate from a time-varying risk premium or from the behavioral biases assumed in Barberis, Shleifer, and Vishny (1998) and Daniel, Hirshleifer, and Subrahmanyam (1998). Interestingly, certain characteristics of the information weights mimic these biases, although our investor is not *assumed* to be influenced by behavioral biases. Furthermore, unlike agents in rational expectation models, our investor is not concerned with the impact of their learning on prices. Therefore, the statistical optimization underlying our aggregate cashflow estimate offers a middle ground between behavioral and rational perspectives on momentum.[4]

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 using analyst forecasts. In addition, we confirm our framework's prediction that momentum is stronger for stocks with greater cashflow uncertainty using analyst forecast dispersion as a proxy. A simulation study also verifies that under reasonable parameters, forecast accuracy uncertainty produces momentum profits whose magnitude is comparable to existing empirical studies.

Several other predictions from our model are consistent with the empirical evidence in Jiang, Lee, and Zhang (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

[4]Our representative investor holds the risky asset at every point in time and cannot profit from momentum. Limits to arbitrage and market frictions can prevent arbitrageurs from eliminating the momentum induced by their learning. Indeed, cashflow uncertainty may limit the willingness of arbitrageurs to implement momentum strategies.

firms, young firms, and those undergoing significant changes in their cashflow growth. Stronger momentum for stocks with higher return volatility and higher cashflow volatility are also predicted.

Timmermann (1993) and Lewellen and Shanken (2002) examine the ability of parameter uncertainty to generate return predictability. In contrast, our investor does not model time-varying cashflow dynamics and does not learn about a firm's cashflow growth parameters from realized cashflows. Moreover, cashflow growth uncertainty alone cannot generate return predictability. Instead, our investor's reliance on multiple cashflow forecasts with time-varying weights is crucial.

Our framework also differs from Hong, Stein, and Yu (2007)'s model that has a representative investor using simple univariate models to forecast cashflow when the true cashflow generating process is multivariate. The investor in Hong, Stein, and Yu (2007) 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 their cashflow expectation. Overall, the uncertainty that surrounds the relative accuracy of different cashflow forecasts has not appeared in the existing literature.

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

II The Model

Following Barberis, Shleifer, and Vishny (1998), our economy consists of a single risky security (stock) and a risk-neutral representative investor with an exogenous constant discount rate δ. All cashflows N_t are paid out as dividends. Under the objective probability, cashflow 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}:

$$y_{t+1} = \theta_{t+1} + \varepsilon_{t+1}, \tag{1}$$

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

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

The parameter $\bar{\theta}$ represents the unconditional average cashflow growth rate, and is set to zero without loss of generality. A nonzero unconditional mean cashflow growth rate adds a constant term to the average stock return but does not affect our conclusions regarding return predictability.[5] The parameter σ_θ captures the uncertainty surrounding expected cashflow growth, while 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 cashflow growth. In our model, realized cashflow growth is uninformative regarding future cashflow growth.[6] Instead, our investor receives multiple forecasts of future cashflow growth, with each forecast issued by a different information source (such as a sell-side analyst). Information sources issue cashflow growth forecasts for the next period.[7] Specifically, on date t, the investor observes the forecast μ_t^j for y_{t+1} issued by the j^{th} information source where $j = 1, \ldots, J$.

The investor forms their conditional expectation of future cashflow 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 cashflow growth forecasts to form their conditional ex-

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

[6] This property is consistent with Chan, Karceski, and Lakonishok (2003)'s finding that realized cashflows are poor predictors of future cashflows. Our investor does not learn about the parameters underlying cashflow growth from realized cashflows, as in Lewellen and Shanken (2002).

[7] Information sources can issue forecasts for a sequence of future cashflows without altering our predictions.

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

A. Optimal Weights and Forecast Accuracy Uncertainty

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

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

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

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

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

where $\mathbf{1}$ denotes a J-dimensional vector of ones, and $\Theta_t = \mathrm{E}_t[\epsilon_{t+1}\epsilon_{t+1}^T]$ is a J by J matrix. W_t summarizes the optimal weights assigned to each information source's cashflow growth forecast at time t. The aggregate cashflow estimate $\hat{\mu}_t$, which combines all available forecasts using their optimal weights, serves as the investor's conditional expectation of cashflow growth. By definition, this aggregate cashflow 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 $\mathrm{E}_t[\epsilon_{j,t+1}]$ may not be zero. With $\mathrm{E}_t\left[\epsilon_{t+1}^2\right] = \mathrm{Var}_t\left[\epsilon_{t+1}\right] + \left(\mathrm{E}_t\left[\epsilon_{t+1}\right]\right)^2$, the investor accounts for potential forecast biases when minimizing the 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 cashflow forecasts for a single stock into an aggregate cashflow estimate, rather than combining multiple stocks into a portfolio. The solution for the weights in (5) equals

$$W_t = \frac{\Theta_t^{-1}\mathbf{1}}{\mathbf{1}^T\Theta_t^{-1}\mathbf{1}} \,. \tag{6}$$

The corresponding aggregate cashflow 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 $\mu_t = (\mu_t^1, \ldots, \mu_t^j, \ldots, \mu_t^J)^T$ is the vector of cashflow growth forecasts.

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

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

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

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 cashflows.[8] Although the investor immediately incorporates newly released or revised forecasts into their aggregate cashflow 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 non-earnings-announcement dates are not accompanied by additional cashflow realizations. Hence, contemporaneous forecast errors are unavailable, and the investor has to rely on past forecast errors in these instances.

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

Appendix A shows that the optimal weights in (6), when Θ_t are estimated as in (7) and (8), coincide with the slope coefficients of the following regression

$$\mathbf{Y} = \mathbf{U}\, W + \varepsilon,$$

where $\mathbf{Y}$ is the vector of past realized cashflow innovations $\{y_{t-1}, \ldots, y_{t-n}\}$, and $\mathbf{U}$ 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 cashflow 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 their 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 cashflow 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.

B. Return Implications of Weight Updating

We now examine the asset pricing implications of our aggregate cashflow 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

$$P_t \;=\; \frac{\mathrm{E}_t^I[N_{t+1}]}{1+\delta} + \frac{\mathrm{E}_t^I[N_{t+2}]}{(1+\delta)^2} + \cdots\,, \tag{10}$$

where $E_t^I[-]$ denotes the *investor's* date t expectation conditional on the J cashflow growth forecasts. Specifically, $\mathrm{E}_t^I[y_{t+1}] = \hat{\mu}_t$. Recall that cashflow growth is forecasted for the next period. The conditional expectation $\mathrm{E}_t^I[y_{t+i}]$ for cashflow growth beyond this horizon, $i > 1$, is equal to zero, its unconditional mean. This implies

$$\mathrm{E}_t^I[N_{t+i}] \;=\; \mathrm{E}_t^I[N_t + y_{t+1} + \cdots + y_{t+i}] = N_t + \hat{\mu}_t\,. \tag{11}$$

The pricing formulation in (11) is similar to Barberis, Shleifer, and Vishny (1998) with a critical distinction. Our aggregate cashflow estimate $\hat{\mu}_t$ results from a combination of cashflow 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 cashflow estimate, $\hat{\mu}_{t+1}$.

The expected stock return under the objective probability equals

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

Thus, the expected return is determined by changes in the aggregate cashflow 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 that[9]

$$\mathrm{E}_t\left[\mu_{t+1}^j\right] = \mu_t^j. \tag{15}$$

Intuitively, cashflow growth uncertainty causes the information sources to maintain, on average, their existing cashflow growth forecasts over short horizons. This is consistent with the information sources being Bayesians with informative priors regarding the expected cashflow growth rate. It also parallels Hong, Stein, and Yu (2007)'s assumption that investors maintain

[9] 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.

their prevailing cashflow 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. 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

$$\mathrm{E}_t[R_{t+1}] \;=\; \frac{\sum_j (\omega^j_{t+1} - \omega^j_t)\,\mu^j_t}{\delta}\,. \tag{16}$$

Equation (16) implies that the investor's expected return is proportional to $\mathrm{Cov}(\Delta\omega, \mu)$, the covariance between the cashflow forecasts and changes in their information weights:[10]

$$\mathrm{E}_t[R_{t+1}] = \mathrm{Cov}(\Delta\omega, \mu)/\delta \;=\; \sigma_{\Delta\omega}\;\sigma_\mu\;\rho_{\Delta\omega,\mu}/\delta, \tag{17}$$

where $\mathrm{Cov}(\Delta\omega, \mu)$ 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.*

$$\mathrm{E}_t[R_{t+1}|y_t > \bar{y}] > 0$$
$$\mathrm{E}_t[R_{t+1}|y_t < \bar{y}] < 0\,.$$

Proof: For tractability, we consider two independent information sources. An optimistic information source issues a cashflow growth forecast $\mu^O_t > \bar{y}$ and a pessimistic information source issues a forecast $\mu^P_t < \bar{y}$. By symmetry, we prove Proposition 1 for the case of a positive earning surprise.

[10]The cross-sectional mean of $\Delta\omega$ is zero by definition since the information weights sum to one.

By (16), the expected stock return is

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

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

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

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

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

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:

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

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

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

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

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

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, since $\bar{y} < \mu_t^O$, 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, since $\bar{y} > \mu_t^P$, 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.■

Proposition 1 proves the existence of earnings momentum (or post-earnings announcement drift) in our model. A related empirical anomaly is price momentum. Chan, Jegadeesh, and Lakonishok (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 cashflows. With a constant discount rate, stocks returns are attributable to changes in expected cashflow. 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 cashflow N_0 of 1, zero unconditional mean cashflow growth ($\bar{\theta}$), and a 3 percent cashflow 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 percent per annum. The σ_θ parameter represents the uncertainty surrounding expected cashflow growth. This parameter is chosen to be 1 or 2 percent, which is reasonable in compared to σ_y. In each simulated economy, there are two cashflow forecasts. The optimistic forecast for period $t+1$'s cashflow is $N_t + Disp$, while 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 percent to 3 percent, which is quite conservative.

For each set of $(\sigma_\theta, Disp)$ parameters, we first simulate 2,000 cashflow 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, cashflow growth is equally likely to be positive or negative, and the information sources have identical true accuracies. The price momentum sort identifies paths where there is a recent trend in cashflow growth. For illustration, suppose a sequence of positive cashflow 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 cashflow 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 cashflow estimate and lowers the stock price. Thus, trends in realized cashflows 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 Shiller (1992)'s assertion that stock return volatility is excessive relative to the volatility of cashflow. However, excess volatility is induced by learning in our model, not irrationality. Lewellen and Shanken (2002)'s learning model also yields excess return volatility. From (13), the stock's realized return variance equals[11]

$$\mathrm{Var}_t(R_{t+1}) \;=\; \frac{1}{\delta^2}\left[\sigma_y^2 + \mathrm{Var}_t\left(\hat{\mu}_{t+1} - \hat{\mu}_t\right)\right] \,. \tag{23}$$

In the absence of forecast accuracy uncertainty, the stock's return variance reduces to $\frac{\sigma_y^2}{\delta^2}$. Consequently, forecast accuracy uncertainty leads to variability in the aggregate cashflow estimate that increases return volatility.

C. 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.

[11]Note that $Cov_t(y_{t+1}, \hat{\mu}_{t+1} - \hat{\mu}_t) = 0$ because $\mathrm{E}_t[y_{t+1}] = 0$ and (15).

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

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

The higher the correlation between the two information sources, the higher (lower) the information 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 invoked 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 cashflow forecast is negatively correlated with the consensus forecast of analysts, they can appear overconfident as in Daniel, Hirshleifer, and Subrahmanyam (1998). The appearance of overconfidence also arises when the estimated accuracy of the investor's private cashflow 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, Shleifer, and Vishny (1998). With the information weights being path-dependent, trends in realized cashflow growth change the estimated relative accuracy of information sources. Thus, the investor's aggregate cashflow appears to *extrapolate* from realized cashflows. 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.

III 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 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 2004. 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 multi-collinearity 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 towards the optimistic forecast, while a negative dW implies the representative pessimistic analyst has gained more influence on the investor's cashflow expectation.

A. 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. Since forecast dispersion measures cashflow uncertainty, our model predicts stronger momentum for stocks with high cashflow uncertainty.

B. 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 1.45% per month over the 1984 to 2004 sample period. In contrast, the average return of the bottom decile dW portfolio is 0.73% per month.[12] The difference is 0.72% with a t-statistic of 3.14. After adjusting for the Fama-French (1993) three factors and the Pástor-Stambaugh (2003) liquidity factor, the difference in the average returns of dW5 and dW1 portfolio is even larger at 0.93%, and statistically significant.

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 more weight while 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, 64.24% have a positive dW. Conversely, for stocks with negative earnings surprises, 71.1% 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

[12] Average return increases monotonically with dW. For example, the average return of the middle decile dW3 portfolio is 1.16% per month, smaller than that of the dW5 portfolio but larger than dW1.

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 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.[13] 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 0.91% and 1.15% for stocks in the dW1 and dW5 portfolios respectively, but only 0.18% for the dW3 portfolio. The difference in price momentum between dW1 and dW3 is 0.72% per month, with a t-statistic of 1.97. The difference in price momentum between dW1 and dW3 is 0.97% per month, with a t-statistic of 2.43. Therefore, consistent with Prediction 2, greater updating in the information weights leads to stronger price momentum. This pattern appears in unadjusted 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

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

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 towards more established firms and those with greater analyst coverage. Hong, Lim, and Stein (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 1.8%. 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.[14] 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.

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.

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

C. 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, Lee, and Zhang (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 cashflow 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 cashflow 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, Lee, and Zhang (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 while our framework offers an alternative interpretation.

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

IV Conclusions

We study stock prices in a simple model where the investor optimally combines multiple cashflow 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 or behavioral biases.

Empirical tests provide strong support for the model since 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 momentum whose magnitude is comparable with existing empirical evidence.

Our framework offers several interesting applications for future research. For example, our aggregate cashflow 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.

References

[1] Barberis, N., A. Shleifer, and R. Vishny, 1998. A model of investor sentiment. Journal of Financial Economics 49, 307-343.

[2] Brav, A., and J. B. Heaton, 2002. Competing theories of financial anomalies. Review of Financial Studies 15, 575-606.

[3] Brown, L. D., 2001. How important is past analyst forecast accuracy? Financial Analysts Journal 57, 44-49.

[4] Campbell, J. Y., and R. J. Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195-228.

[5] Chan, L. K. C., N. Jegadeesh, J. Lakonishok, 1996. Momentum strategies. Journal of Finance 51, 1681-1713.

[6] Chan, L. K. C., N. Jegadeesh, J. Lakonishok, 1999. The profitability of momentum strategies. Financial Analysts Journal 55, 80-90.

[7] Chan, L. K. C., J. Karceski, and J. Lakonishok, 2003. The level and persistence of growth rates. Journal of Finance 58, 643-684.

[8] Chordia, T., and L. Shivakumar, 2006. Earnings and price momentum. Journal of Financial Economics 80, 627-656.

[9] Clement, M. B., and S. Y. Tse, 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? The Accounting Review 78, 227-249.

[10] Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839-1885.

[11] Daniel, K., and S. Titman, 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605-1643.

[12] Fama, E., and K. R. French, 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

[13] Hirshleifer, D., and S. H. Teoh, 2003. Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics 36, 337-386.

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

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

[16] Hong, H., J. C. Stein, and J. Yu, 2007. Simple forecasts and paradigm shifts. Journal of Finance 62, 1207-1242.

[17] Hong H., W. Torous, and R. Valkanov, 2007. Do industries lead the stock market? Gradual diffusion of information and cross-asset return predictability. Journal of Financial Economics 83, 367-396.

[18] Jackson, A., and T. Johnson, 2006. Unifying underreaction anomalies. Journal of Business 79, 75-114.

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

[20] Jiang, G., C. M. C. Lee, and Y. Zhang, 2005. Information uncertainty and expected returns. Review of Accounting Studies 10, 185-221.

[21] Lee, C. M. C., and B. Swaminathan, 2000. Price momentum and trading volume. Journal of Finance 55, 2017-2069.

[22] Lewellen, J. and J. Shanken, 2002. Learning, asset pricing tests, and market efficiency. Journal of Finance 57, 1113-1145.

[23] Lim, T., 2001. Rationality and analysts' forecast bias. Journal of Finance 56, 369-385.

[24] Markowitz, H. M., 1952. Portfolio selection. Journal of Finance 7, 77-91.

[25] Pástor, L., and R. F. Stambaugh, 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642-685.

[26] Peng, L., and W. Xiong, 2006. Limited attention, overconfidence and category learning. Journal of Financial Economics 80, 563-602.

[27] Shiller, R. J., 1992. Market volatility, MIT Press, Cambridge, Mass.

[28] Sinha, P., L. D. Brown, and S. Das, 1997. A re-examination of financial analysts' differential earnings forecast accuracy. Contemporary Accounting Research 14, 1-42.

[29] Timmermann, A. G., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. Quarterly Journal of Economics 108, 1135-1145.

[30] Zhang, X. F., 2006. Information uncertainty and stock returns. Journal of Finance 61, 105-136.

APPENDICES

A. Regression Interpretation of the Optimal Weights

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

$$(\mathbf{Y} - \mathbf{U}W)^T (\mathbf{Y} - \mathbf{U}W) \ , \tag{25}$$

by selecting W. This $J \times 1$ matrix of information weights is often denoted β, while $\mathbf{U}$ represents a $n \times J$ matrix of cashflow 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, while the $n \times 1$ vector $\mathbf{Y}$ denotes the corresponding realized cashflows.

Inserting the constraint $\mathbf{1}^T W = 1$ into (25) yields

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

since $\mathbf{Y}\mathbf{1}^T - \mathbf{U}$ is an $n \times J$ matrix of forecast errors and $\frac{1}{n}\left(\mathbf{Y}\mathbf{1}^T - \mathbf{U}\right)^T \left(\mathbf{Y}\mathbf{1}^T - \mathbf{U}\right)$ 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.

B: Alternative Forecast Dynamics

This appendix relaxes the assumption that the cashflow forecasts are, on average, constant over short-horizons. Consider three scenarios defined by the realized cashflow y_{t+1}

1. $y_{t+1} > \mu_t^O$
2. $y_{t+1} < \mu_t^P$
3. $\mu_t^P < y_{t+1} < \mu_t^O$

where μ_t^O and μ_t^P denote the optimistic and pessimistic cashflow forecast respectively. In the first scenario, the optimistic and pessimistic information source both increase their cashflow forecasts at $t + 1$ since y_{t+1} is higher than μ_t^O and μ_t^P. These increases are consistent with both information sources being Bayesians. The updated forecasts, $\mu_{t+1}^O > \mu_t^O$ and $\mu_{t+1}^P > \mu_t^P$ imply the first scenario generates the $\hat{\mu}_{t+1} > \hat{\mu}_t$ relationship. Similarly, in the second scenario, the optimistic and pessimistic information source both decrease their cashflow forecasts since y_{t+1} is lower than μ_t^O and μ_t^P. This updating of the individual forecasts yields $\hat{\mu}_{t+1} < \hat{\mu}_t$.

For the third scenario, the individual cashflow forecasts are updated as follows

$$\begin{aligned} \mu_{t+1}^O &= \mu_t^O + \phi_t^O + \varepsilon_{t+1}^O \\ \mu_{t+1}^P &= \mu_t^P + \phi_t^P + \varepsilon_{t+1}^P, \end{aligned}$$

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

$$\frac{\phi_t^O}{\phi_t^P} = -\frac{\sigma_{O,t+1}^2}{\sigma_{P,t+1}^2}\,. \tag{26}$$

The condition in (26) states that the (absolute) amount by which the optimistic (pessimistic) information source updates their forecast downwards (upwards) is proportional to their estimated accuracy in (7). Therefore, the relatively less accurate information source updates their cashflow forecast more dramatically in the direction of the realized cashflow, with (26) being consistent with Bayesian updating.

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

$$\begin{aligned} & \left(\omega_t^O + \Delta\omega_t^O\right)\left(\mu_t^O + \phi_t^O\right) + \left(\omega_t^P + \Delta\omega_t^P\right)\left(\mu_t^P + \phi_t^P\right) - \omega_t^O\mu_t^O - \omega_t^P\mu_t^P \\ = \; & \Delta\omega_t^O\mu_t^O + \Delta\omega_t^P\mu_t^P + \omega_t^O\phi_t^O + \Delta\omega_t^O\phi_t^O + \omega_t^P\phi_t^P + \Delta\omega_t^P\phi_t^P\,. \end{aligned} \tag{27}$$

To demonstrate that momentum occurs in the third scenario, consider two cases. First, after positive cashflow growth realizations, the optimistic information source becomes more accurate, and is assigned a larger weight, $\Delta\omega_t^O > 0$. The first two terms on the right side of (27), $\Delta\omega_t^O\mu_t^O + \Delta\omega_t^P\mu_t^P$, are positive when combined since they equal

$$\begin{aligned} \Delta\omega_t^O\mu_t^P + \Delta\omega_t^P\mu_t^P + \Delta\omega_t^O\left(\mu_t^O - \mu_t^P\right) &= \mu_t^P\left(\Delta\omega_t^O + \Delta\omega_t^P\right) + \Delta\omega_t^O\left(\mu_t^O - \mu_t^P\right) \\ &= \Delta\omega_t^O\left(\mu_t^O - \mu_t^P\right) > 0\,. \end{aligned} \tag{28}$$

The above inequality follows from the optimistic information source being assigned a larger portfolio weight, $\Delta\omega_t^O > 0$, while the $\mu_t^O - \mu_t^P > 0$ and $\Delta\omega_t^O + \Delta\omega_t^P = 0$ properties hold by definition. The remaining terms in (27) sum to zero

$$\omega_t^O \phi_t^O + \Delta\omega_t^O \phi_t^O + \omega_t^P \phi_t^P + \Delta\omega_t^P \phi_t^P \;\; = \;\; \omega_{t+1}^O \phi_t^O + \omega_{t+1}^P \phi_t^P = 0\,,$$

due to the condition in (26) and $\omega_{t+1}^O/\omega_{t+1}^P = \sigma_{P,t+1}^2/\sigma_{O,t+1}^2$. Thus, $\mathrm{E}_t[\hat{\mu}_{t+1}] > \hat{\mu}_t$ after positive cashflow growth.

The second case involves negative cashflow growth realizations, which cause the weight of the pessimist to increase at the optimist's expense. In other words, $\Delta\omega_t^P > 0$ and $\Delta\omega_t^O < 0$. Consequently, the first two terms on the right side of (27), $\Delta\omega_t^O \mu_t^O + \Delta\omega_t^P \mu_t^P$, are negative when combined, while the remaining terms again sum to zero. Thus, $\mathrm{E}_t[\hat{\mu}_{t+1}] < \hat{\mu}_t$ after negative cashflow growth.

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

Table 1: Simulated Returns from Momentum Strategy

This table presents the average returns of momentum portfolios based on simulations with the following common inputs: initial cashflow $N_0 = 1$, unconditional cashflow growth $\bar{\theta} = 0$, and cashflow volatility $\sigma_y = 0.03$. There are two cashflow forecasts. The optimistic forecast for period $t+1$'s cashflow is $N_t + Disp$, while 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 cashflow growth) and $Disp$ parameters. For each set of parameters, we simulate 2000 cashflows and prices, with each time-series containing 120 observations. For each observation, the 2000 simulation paths are ranked cross-sectionally based on their cumulative returns over the last 6 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 1 to 6 periods. The mean returns for each holding period are recorded below with t-statistics in parentheses.

σ_θ	$Disp$	Formation period	Holding period returns		
		return (6 Periods)	1 Period	3 Periods	6 Periods
0.0100	0.0100	0.3200	0.0068	-0.0009	-0.0094
			(7.3081)	(-0.7042)	(-5.7803)
0.0200	0.0100	0.4269	0.0074	-0.0031	-0.0135
			(5.9231)	(-1.5030)	(-5.3187)
0.0100	0.0200	0.3321	0.0091	0.0027	-0.0041
			(9.6328)	(1.7454)	(-1.9609)
0.0200	0.0200	0.3821	0.0094	0.0014	-0.0077
			(7.4238)	(0.8075)	(-3.3060)

Table 2: Price Momentum Conditional on Information Weight Change

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 2004, stocks from the intersection of the CRSP and IBES datasets 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 towards the representative pessimistic analyst, while the dW5 portfolio contains stocks whose weight is shifting from the representative pessimistic analyst towards 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. Unadjusted returns and those adjusted by the Fama-French (1993) three factors and the Pástor-Stambaugh (2003) liquidity factor are both provided.

Panel A: sorted on MOM first, then on dW

	P5-P1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	P1	P2	P3	P4	P5	*P5-P1*	*t-stat*	*P5-P1*	*t-stat*
dW1	*91.50*	0.33	0.86	0.94	0.72	1.24	*0.91*	*1.60*	*1.00*	*1.58*
dW2	*81.17*	0.40	1.15	0.36	0.47	0.90	*0.50*	*0.82*	*0.45*	*0.65*
dW3	*75.12*	0.74	0.73	1.31	0.19	0.93	*0.18*	*0.30*	*-0.02*	*-0.03*
dW4	*78.77*	1.15	1.33	1.44	1.43	1.66	*0.51*	*0.90*	*0.42*	*0.68*
dW5	*88.46*	0.98	1.59	1.28	1.82	2.13	*1.15*	*1.96*	*1.15*	*1.72*
dW1 - dW3							*0.72*	*1.97*	*1.02*	*1.92*
dW5 - dW3							*0.97*	*2.43*	*1.17*	*2.39*

Panel B: sorted on dW first, then on MOM

	P5-P1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	P1	P2	P3	P4	P5	*P5-P1*	*t-stat*	*P5-P1*	*t-stat*
dW1	*87.91*	0.62	1.03	0.86	0.71	1.35	*0.73*	*1.49*	*1.00*	*1.83*
dW2	*77.41*	0.45	0.57	1.09	0.21	0.91	*0.46*	*0.78*	*0.18*	*0.28*
dW3	*73.61*	0.81	0.90	0.65	0.84	0.91	*0.10*	*0.16*	*0.21*	*0.30*
dW4	*73.26*	0.97	0.94	1.36	1.42	1.46	*0.49*	*1.04*	*0.31*	*0.60*
dW5	*89.41*	0.87	1.68	1.59	1.97	1.95	*1.09*	*1.75*	*1.15*	*1.64*
dW1 - dW3							*0.63*	*1.80*	*0.79*	*1.89*
dW5 - dW3							*0.98*	*2.47*	*0.94*	*2.45*

Table 3: Earnings Momentum Conditional on Information Weight Change

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 2004, stocks from the intersection of the CRSP and IBES datasets 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 towards the representative pessimistic analyst, while the dW5 portfolio contains stocks whose weight is shifting from the representative pessimistic analyst towards 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. Unadjusted returns and those adjusted by the Fama-French (1993) three factors and the Pástor-Stambaugh (2003) liquidity factor are both provided.

Panel A: sorted on MOM first, then on dW

	E5-E1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	E1	E2	E3	E4	E5	*E5-E1*	*t-stat*	*E5-E1*	*t-stat*
dW1	*29.79*	0.19	0.44	0.21	0.09	1.44	*1.25*	*2.33*	*1.43*	*2.39*
dW2	*30.19*	0.10	0.04	0.19	0.73	0.66	*0.57*	*1.10*	*0.88*	*1.62*
dW3	*25.97*	0.55	0.71	0.60	1.20	1.24	*0.69*	*1.46*	*0.92*	*1.75*
dW4	*25.16*	0.49	0.26	0.35	1.22	0.95	*0.46*	*1.04*	*0.59*	*1.19*
dW5	*25.41*	0.45	0.37	0.57	0.98	1.62	*1.17*	*2.63*	*1.36*	*2.68*
dW1 - dW3							*0.56*	*1.72*	*0.51*	*1.55*
dW5 - dW3							*0.48*	*1.60*	*0.44*	*1.43*

Panel B: sorted on dW first, then on MOM

	E5-E1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	E1	E2	E3	E4	E5	*E5-E1*	*t-stat*	*E5-E1*	*t-stat*
dW1	*29.57*	0.27	0.16	0.72	0.91	1.37	*1.10*	*2.89*	*1.08*	*2.99*
dW2	*28.68*	0.54	0.12	-0.13	0.67	1.29	*0.75*	*1.31*	*1.01*	*1.67*
dW3	*27.10*	0.31	0.12	0.26	0.23	0.89	*0.58*	*1.31*	*0.72*	*1.64*
dW4	*25.52*	0.41	0.45	0.01	1.33	1.41	*0.99*	*2.50*	*0.78*	*2.56*
dW5	*28.47*	0.63	0.59	0.61	0.57	1.54	*0.91*	*1.74*	*1.17*	*1.78*
dW1 - dW3							*0.52*	*1.69*	*0.36*	*1.53*
dW5 - dW3							*0.33*	*1.29*	*0.45*	*1.37*

Table 4: Price and Earnings Momentum Conditional on Forecast Dispersion

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 2004, stocks from the intersection of the CRSP and IBES datasets 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, while Panel B reverses the order of the double-sort. Panel C and Panel D record our results for earnings momentum rather than price momentum. Unadjusted returns and those adjusted by the Fama-French (1993) three factors and the Pástor-Stambaugh (2003) liquidity factor are both provided.

Panel A: price momentum, sorted on MOM first, then on uncertainty

	P5-P1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	P1	P2	P3	P4	P5	*P5-P1*	*t-stat*	*P5-P1*	*t-stat*
U1	*76.26*	0.83	1.34	1.38	1.44	1.44	*0.61*	*1.98*	*0.61*	*1.71*
U2	*75.30*	0.81	1.27	1.34	1.35	1.62	*0.81*	*2.65*	*0.77*	*2.21*
U3	*76.87*	0.74	1.41	1.36	1.26	1.72	*0.98*	*3.25*	*0.87*	*2.50*
U4	*79.25*	0.80	1.30	1.38	1.16	1.83	*1.03*	*3.26*	*0.92*	*2.53*
U5	*84.78*	0.35	1.06	1.28	1.25	1.98	*1.63*	*4.88*	*1.41*	*3.63*
U5 - U1							*1.02*	*4.46*	*0.80*	*3.19*

Panel B: price momentum, sorted on uncertainty first, then on MOM

	P5-P1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	P1	P2	P3	P4	P5	*P5-P1*	*t-stat*	*P5-P1*	*t-stat*
U1	*76.21*	1.25	1.31	1.39	1.53	1.90	*0.66*	*2.19*	*0.67*	*1.52*
U2	*70.14*	0.84	1.32	1.32	1.25	1.63	*0.78*	*2.67*	*0.59*	*1.77*
U3	*71.82*	0.91	1.45	1.37	1.22	1.82	*0.91*	*3.22*	*0.81*	*2.52*
U4	*76.41*	0.79	1.19	1.32	1.33	1.70	*0.91*	*3.10*	*0.65*	*1.93*
U5	*84.38*	0.23	0.92	1.13	1.26	1.81	*1.58*	*4.57*	*1.31*	*3.26*
U5 - U1							*0.92*	*4.34*	*0.64*	*2.70*

Panel C: earnings momentum, sorted on MOM first, then on uncertainty

	E5-E1 return spread	Holding period returns					*unadjusted*		*4-factor adjusted*	
	over past six months	E1	E2	E3	E4	E5	*E5-E1*	*t-stat*	*E5-E1*	*t-stat*
U1	*28.07*	0.99	1.10	1.32	1.39	1.65	*0.65*	*2.90*	*0.78*	*2.96*
U2	*25.79*	0.96	1.16	1.15	1.34	1.64	*0.68*	*2.84*	*0.71*	*2.57*
U3	*25.31*	0.90	1.20	1.21	1.35	1.72	*0.82*	*3.89*	*0.85*	*3.54*
U4	*26.22*	0.87	1.08	1.29	1.33	1.73	*0.86*	*3.53*	*0.84*	*2.91*
U5	*25.16*	0.50	1.14	1.27	1.42	1.66	*1.17*	*4.51*	*1.35*	*4.80*
U5 - U1							*0.52*	*3.05*	*0.57*	*3.32*