Return Uncertainty and Biases in Expected Returns

January 2005

Abstract

We study the relationship between return uncertainty and behavioral finance by aggregating multiple return forecasts for a single asset into an estimate of its unknown expected return. The combination of forecasts which minimizes the uncertainty of the estimated expected return is determined by the optimal *information portfolio*. This minimization provides an alternative explanation for biases in expected returns that have previously been attributed to psychology. Specifically, biases which appear similar to overconfidence, biased self-attribution, representativeness, conservatism and limited attention arise from the information portfolio weights assigned to return forecasts. Higher dispersion across the return forecasts increases return predictability and the magnitude of these biases. However, our optimal information portfolio yields testable implications distinct from psychology, which we verify empirically using revisions in analyst earnings forecasts.

1 Introduction

The empirical asset pricing literature defines expected returns ex-post. For example, the Fama-French (1993) three factor model computes market, SMB and HML sensitivities from historical returns. However, a consensus regarding the correct formulation of expected returns remains elusive since the number of factors required to generate expected returns is controversial. Indeed, dispersion in the ex-ante return forecasts of market participants parallels academic disagreements surrounding the ex-post calibration of expected returns. Moreover, randomness in the dynamics of the underlying factors as well as uncertainty regarding an asset's factor loadings next period imply expected returns are unknown.

Our framework has a combination of return forecasts determining an individual asset's unknown expected return. Return forecasts are issued by information sources after interpreting a state variable such as a firm's projected earnings or industry conditions. Public information sources include analysts and the firm itself, while the investor generates private return forecasts. The historical accuracy of an information source is measured according to its prior forecast errors, defined as the difference between realized and forecasted returns. Historical covariances between the forecast errors of different information sources are also analyzed.

The information portfolio combines the return forecasts for an individual asset into an estimate of its expected return. This is accomplished by assigning each information source a portfolio weight. In contrast to existing portfolio theory for multiple assets with known expected returns, our information portfolio applies to multiple return forecasts for a single asset whose expected return is unknown. Specifically, by invoking a mathematical formulation similar to standard portfolio theory, our optimal information portfolio minimizes the aggregate uncertainty of a single asset's estimated return by assigning higher portfolio weights to information sources with greater historical accuracy.[1] The most accurate estimate of an asset's expected return follows directly from the information portfolio, and is labeled the investor's *perceived* return. Consequently, the influence of a return forecast on the investor's perceived return is determined by its information portfolio weight, which depends on the historical

[1] After imposing a common distributional assumption on the set of return forecasts, this minimization is equivalent to solving for the *best linear unbiased estimate* (BLUE) of an asset's true expected return. However, since our optimal information portfolio is independent of any distributional assumptions, we refrain from referring to the information portfolio weights as linear regression coefficients.

accuracy of the information source issuing the return forecast.

With regards to behavioral finance, the optimal information portfolio aggregates return forecasts in a manner which mimics psychological biases. In particular, when information sources have heterogenous beliefs regarding an asset's expected return, the perceived return exhibits biases that appear similar to overconfidence and biased self-attribution as well as representativeness and conservatism. These two pairs of biases have been previously incorporated into the finance literature by Daniel, Hirshleifer and Subrahmanyam (1998) and Barberis, Shleifer and Vishny (1998) respectively. However, information portfolio theory explains the appearance of these biases in an asset's perceived return using optimal information portfolio weights rather than psychology. For example, the optimal information portfolio emphasizes the investor's private return forecasts which are accurate, while downplaying their less accurate counterparts. In addition, even in the absence of any theoretical justification, state variables with trends in their dynamics or stable return implications receive larger information portfolio weights. A bias which mimics limited attention is also instilled into the perceived return since return forecasts which are positively correlated with those from more accurate information sources are ignored by the information portfolio.

Furthermore, unlike Bayesian frameworks which assume investors commit psychological biases by imposing assumptions on the prior distribution, we examine the optimal combination of multiple return forecasts when estimating an asset's expected return. Therefore, biases in the perceived return are outputs from our optimal information portfolio. Brav and Heaton (2002) demonstrate the difficulty of distinguishing between behavioral and rational explanations for anomalies using Bayesian techniques.[2]

In addition, the optimal information portfolio also generates return predictability. Higher return uncertainty, characterized by more disparate return forecasts, also increases the apparent strength of the perceived return biases. Thus, dispersion in the return forecasts is crucial to the magnitude of our biases, and reflects ex-ante uncertainty regarding an asset's expected return rather than investor psychology. For example, changes in a firm's capital structure or investment strategy can alter its future earnings dynamics. However, our framework does not assume the return implications of such events are immediately understood and agreed upon by all market participants. Indeed, every return forecast would be identical and without error under this simplifying assumption. Instead, dispersion between the return forecasts may result from parameter uncertainty as well as disagreement regarding

[2] Section 4 contains further details on the distinction between information portfolio theory and the Bayesian approach.

the appropriate interpretation of limited information. Lewellen and Shanken (2002) investigate the repercussions of parameter uncertainty on asset pricing tests and return predictability.

To illustrate our notion of return uncertainty, a BusinessWeek survey conducted at the end of 2005 reported year-end 2006 forecasts for the S&P 500 ranging between 880 and 1,635 with a standard deviation of 95 points. Provided the 76 forecasters included in the survey have unequal historical accuracies, the average forecast of 1,347 is not an optimal estimate for the S&P 500 at the end of 2006. Finding the optimal weight to assign each forecast is one motivation for our information portfolio.

Empirically, Jackson and Johnson (2006) document that momentum and post-earnings announcement drift both coincide with firm-specific events that alter a firm's earnings, while the composite share issuance variable of Daniel and Titman (2005) also indicates return predictability. In addition, Kumar (2005) and Zhang (2005) both report that behavioral biases appear stronger in periods of higher uncertainty. Besides being event and time dependent, Baker and Wurgler (2005) report that characteristics such as size and age explain a firm's sensitivity to investor sentiment.

These empirical regularities are consistent with information portfolio theory as well as psychological biases. However, several testable implications unique to information portfolio theory are available. Most importantly, investors find accurate information sources more credible when forming their perceived return. Even in periods of high return uncertainty, a relative ranking of the information sources by their historical accuracies is equivalent to the existence of an information portfolio. Prior returns of the asset are not excluded from being the most accurate source of public information. The historical accuracy of an information source arises from two components; uncertainty in the dynamics of an underlying state variable and uncertainty regarding its return implications. As a consequence, after controlling for state variable uncertainty, information portfolio theory asserts that investors focus their attention on state variables which have the highest correlation with returns. Furthermore, predictability in a state variable and its return implications increases the accuracy of an information source, granting its return forecasts more influence over the investor's perceived return.

The first aspect of our empirical study pertains to the return implications of earnings. Specifically, we measure the *sensitivity* of returns to earnings revisions by computing firm-specific correlations between these variables. We find earnings momentum profits increase monotonically from low to high sensitivity stocks by 50%, evidence consistent with investors focusing more attention on this information source when its corresponding return forecast has been more accurate. The second aspect of our empirical study considers the role of earnings *uncertainty*. As documented in Zhang (2005), momentum profits are larger for stocks with higher earnings dispersion. More importantly, portfolios derived from double sorts on the sensitivity and uncertainty measures continue to display both relationships with earnings momentum. Consequently, even after controlling for earnings uncertainty, firms whose returns are more sensitive to earnings revisions experience greater earnings momentum. This finding is consistent with more accurate information sources having more influence over the perceived return, which is the central prediction of information portfolio theory. Several robustness checks verify that our sensitivity and uncertainty measures are not driven by factors such as book-to-market, size and analyst coverage.

Nonetheless, in theory, if the most accurate information sources incorporate investor psychology into their return forecasts, then information portfolio theory and psychology are compatible as both influence the investor's perceived return. Indeed, the exact decomposition of the perceived return into the effects of psychology versus information portfolio theory is ultimately an empirical question. With this caveat in mind, empirical evidence in Section 5 reports that the optimal information portfolio cannot be ignored when analyzing historical returns.

The remainder of this paper begins with the introduction of the optimal information portfolio in Section 2. Section 3 illustrates the impact of state variable predictability on the historical accuracy of an information source and examines its ability to induce return predictability, while Section 4 links the optimal information portfolio with biases previously attributed to psychology. Testable implications of information portfolio theory are provided in Section 5 along with an empirical implementation. Our conclusions and suggestions for further research are contained in Section 6.

2 Information Portfolio Theory

As in Daniel, Hirshleifer and Subrahmanyam (1998) as well as Barberis, Shleifer and Vishny (1998), we consider a single-investor, single-asset model. Thus, we restrict our attention to an investor functioning as a price-setter who does not "free-ride" on market prices.

Underlying our framework are state variables, examples of which include forecasts for the earnings or sales of an individual firm as well as macroeconomic and industry conditions among many other possibilities. Each state variable, or its forecast for next period, is interpreted by an information source who expresses its estimated return implications.[3] For simplicity, each return forecast is generated by a single state variable. In practice, an individual analyst can issue quarterly earnings forecasts and long term growth rate projections for earnings along with price targets and buy versus sell recommendations, while firms often disclose their earnings and sales figures in conjunction with "guidance" for these state variables. Therefore, multiple information sources may originate from an individual analyst or firm depending on the amount of information they release. Moreover, different forecasts for a state variable can yield different return forecasts. For example, cross-sectional dispersion in earnings forecasts would imply return uncertainty. The information portfolio is responsible for aggregating across the various return interpretations and forecasts of state variables.

Furthermore, certain return forecasts are *hybrids* possessing both private and public characteristics. For example, public announcements such as analyst earnings forecasts are not expressed as returns and therefore require additional interpretation by the investor. In contrast, the conversion of analyst price targets into return forecasts is immediate.

To examine the impact of return uncertainty on the investor's perceived returns, we consider J return forecasts for a single asset. The magnitude of J accounts for the multitude of available information sources and their potentially disparate forecasts for an asset's expected return. Thus, our information portfolio aggregates J return forecasts originating from J unique information sources who evaluate the return implications of J state variable forecasts, with every return forecast expressing the return implications of a single state variable forecast. When state variable dynamics are random, there would likely exist more state variable forecasts than state variables but this relationship is not crucial to our framework.

For emphasis, information sources are only assumed to issue return forecasts. The mechanism for computing their historical accuracy is addressed in the next subsection. Whether an information source issues also forecasts for the state variables is immaterial to our framework since the historical accuracy of an information source pertains entirely to its time series of prior return forecasts.

[3] Although sales are usually reported in millions of dollars and earnings stated on a per share basis, we abstract from these scale complications by aggregating across their return implications. The state variable to return forecast transformation is another source of return uncertainty and addressed in the next section.

2.1 Historical Forecast Accuracy

The historical accuracy of each return forecast is critical to the information portfolio's solution and is computed from the previous forecast errors of an information source. Specifically, at time $t-1$, the time series of forecast errors for the j^{th} information source consists of the following vector

$$\begin{bmatrix} \epsilon^{j}_{t-1} \\ \vdots \\ \epsilon^{j}_{t-n} \end{bmatrix} = \begin{bmatrix} y_{t-1} \\ \vdots \\ y_{t-n} \end{bmatrix} - \begin{bmatrix} \mu_{j,t-1} \\ \vdots \\ \mu_{j,t-n} \end{bmatrix} \quad \text{for } j = 1, 2, \ldots, J \tag{1}$$

where $\mu_{j,t-i}$ denotes the return forecast issued at time $t-i-1$ for the asset over the previous $i = 1, \ldots, n$ periods, while y_{t-i} represents the corresponding realized return. Therefore, at time $t-1$, the j^{th} information source issues the return forecast $\mu_{j,t}$ for the $(t-1, t]$ horizon, while y_t denotes the asset's realized return at time t. The calendar time corresponding to the $(t-1, t]$ interval is arbitrary.

At $t-1$, the historical accuracy of the j^{th} information source equals

$$\sigma^2_{j,t} = \frac{1}{n} \sum_{i=1}^{n} \left(\epsilon^{j}_{t-i} \right)^2 , \tag{2}$$

according to their forecast errors ϵ^{j}_{t-i} over the last n periods. Let $\sigma^2_{j,*}$ denote the unknown true variance associated with the return forecasts of the j^{th} information source. The historical accuracy in equation (2) serves as the investor's estimate of $\sigma^2_{j,*}$ at $t-1$, and represents the accuracy of $\mu_{j,t}$ from the investor's perspective. Statistically, equation (2) calculates the mean-squared error (MSE) of the previous n forecast errors.[4] Observe that equation (2) allows information sources to employ Bayesian methods when generating their return forecasts with the usual tradeoff between variance and bias arising from an informative prior.

Similarly, the covariance between the time series of forecast errors for the j^{th} and k^{th} information source is estimated as

$$\sigma_{j,k,t} = \frac{1}{n} \sum_{i=1}^{n} \epsilon^{j}_{t-i} \, \epsilon^{k}_{t-i} , \tag{3}$$

for $j \neq k$. Equation (3) represents the investor's estimate of the true but unknown covariance $\sigma_{j,k,*}$ between the return forecasts of two information sources at $t-1$. For emphasis, since the estimates in

[4] This property follows from $E\left[\epsilon^2\right] = Var\left[\epsilon\right] + \left(E\left[\epsilon\right]\right)^2$ with the bias in a forecast equaling $E[\epsilon]$.

equations (2) and (3) are calculated using realized forecast errors over the last n periods, they should be denoted as $\hat{\sigma}^2_{j,t}$ and $\hat{\sigma}_{j,k,t}$ respectively. The hats are omitted for notational simplicity.

Furthermore, the value of n in equations (2) and (3) is specific to an individual asset.[5] Intuitively, established firms in stable industries have a large n. Conversely, initial public offerings, firms undergoing a significant corporate restructuring or undertaking a large investment and those operating in industries that experience major technological innovations have a small n. In the terminology or Brav and Heaton (2002), major corporate events result in a *change point* which reduces n in our framework.

Overall, the μ_t vector of return forecasts at time $t-1$ equals

$$\mu_t = \begin{bmatrix} \mu_{1,t} \\ \mu_{2,t} \\ \vdots \\ \mu_{J,t} \end{bmatrix} . \tag{4}$$

A time series of these vectors $\mu_{t-1}, \ldots, \mu_{t-n}$ over the last n periods yields a Θ_t matrix summarizing the historical accuracies of the J information sources as well as their historical covariances, described by equations (2) and (3) respectively. The Θ_t matrix is a historical estimate of the true but unknown variance-covariance matrix for the J return forecasts in equation (4). For notational simplicity, we suppress the t subscripts for μ and Θ in the remainder of the paper with the understanding that they pertain to the beginning of each period.

The cross-sectional dispersion across the J forecasts of the μ vector at a single point in time,

$$\sigma^2_{\mu,t} = \frac{1}{J-1} \sum_{j=1}^{J} \left(\mu_{j,t} - \bar{\mu}_t\right)^2 , \tag{5}$$

where $\bar{\mu}_t$ is defined as the average return forecast at time $t-1$

$$\bar{\mu}_t = \frac{1}{J} \sum_{j=1}^{J} \mu_{j,t} , \tag{6}$$

does not have an explicit role in our solution for the information portfolio. Nonetheless, $\sigma^2_{\mu,t}$ has an important economic interpretation by offering another proxy for return uncertainty which is likely to be inversely related to n.

[5] When n is information source dependent, a j subscript would be added to form n_j. For example, n could proxy for the experience of an information source. However, for ease of exposition, a common value is written for all J information sources as our initial focus is on a firm-specific information environment.

2.2 Optimal Information Portfolio

Unlike classical portfolio theory, an asset's true expected return is not assumed to be known. Instead, the investor minimizes aggregate forecast to obtain the most accurate estimate of an asset's expected return.[6] The optimization problem responsible for producing the optimal information portfolio W is

$$\begin{aligned} \min_{W} \quad & \frac{1}{2} W^T \Theta W \\ \text{subject to:} \quad & W^T \mathbf{1} = 1 \,, \end{aligned} \tag{7}$$

where $\mathbf{1}$ denotes a J-dimensional vector of ones. As proven later in this section, after imposing a common distributional assumption on every return forecast, the objective function in equation (7) is equivalent to finding the best linear unbiased estimator (BLUE) of the asset's expected return given available forecasts. Therefore, equation (7) is consistent with linear regression models used throughout the empirical finance literature. The optimal information portfolio is solved in the following proposition whose proof is contained in Appendix A.

Proposition 1. *The solution for the optimal information portfolio W in equation (7) equals*

$$W = \frac{\Theta^{-1}\mathbf{1}}{\mathbf{1}^T \Theta^{-1} \mathbf{1}} \,. \tag{8}$$

Proposition 1 immediately generates an estimate for the asset's expected return which is the subject of the next proposition. For emphasis, when private return forecasts are evaluated, the firm-specific information portfolio is also investor-specific.

Extending the optimization problem in equation (7) by constraining the investor to obtain a *target* expected return, as in classical portfolio theory, allows the investor to "manipulate" the portfolio weights. In particular, denote the target return as r^* and observe that the constraint $W^T \mu = r^*$ enables the investor to overweight optimistic forecasts when seeking a high target return. Consequently, this extension is left for further research.

[6] This approach is related to Peng and Xiong (2004)'s minimization for the variance of beliefs regarding subsequent dividends, while Hong, Scheinkman and Xiong (2005) invoke mean-variance preferences when analyzing different information sources.

2.3 Regression Interpretation of Optimal Information Portfolio

After imposing a common distributional assumption on every return forecast, the objective function in equation (7) is equivalent to finding the best linear unbiased estimate of the asset's true expected return. Specifically, assume the vector μ of return forecasts has the following multivariate normal distribution

$$\mu \stackrel{d}{\sim} \mathcal{N}(\eta \mathbf{1}, \Theta) \,. \tag{9}$$

To solve for the optimal conditional expectation, $E[\eta | \mu]$, consider the linear estimator

$$E[\eta | \mu] = W^T \mu \,, \tag{10}$$

whose *coefficients* W minimize the conditional variance of the residuals

$$\begin{aligned} Var[\eta - E[\eta|\mu]] &= Var\left[\eta - W^T \mu\right] \\ &= Var\left[\eta - W^T \mathcal{N}(\eta \mathbf{1}, \Theta)\right] \\ &= W^T \Theta W \,, \end{aligned} \tag{11}$$

after substituting in the distribution for μ from equation (9). The W coefficients are also required to produce an unbiased estimator

$$\begin{aligned} E\left[\eta - W^T \mu\right] &= 0 \\ E\left[\eta - W^T \mathcal{N}(\eta \mathbf{1}, \Theta)\right] &= 0 \\ \eta - \eta W^T \mathbf{1} &= 0 \,, \end{aligned} \tag{12}$$

which leads to the $W^T \mathbf{1} = 1$ constraint. Therefore, the best linear unbiased estimate of the asset's expected return is an equivalent formulation of the objective function in equation (7) where the investor minimizes $W^T \Theta W$ subject to the $W^T \mathbf{1} = 1$ constraint.[7] Consequently, the statistical justification underlying linear regression models also applies to our optimal information portfolio.

However, time-varying information portfolio weights are crucial to our interpretation in Section 4 of the biases manifested in the investor's perceived return. More importantly, the objective function in equation (7) which defines the optimal information portfolio is independent of the distributional assumption in equation (9). Therefore, we refrain from referring to our information portfolio weights as linear regression coefficients.

[7] Minimizing $W^T \Theta W$ in equation (11) is equivalent to minimizing $\frac{1}{2} W^T \Theta W$ in equation (7).

2.4 Perceived Return and Aggregate Return Uncertainty

By aggregating across the return forecasts, the optimal information portfolio immediately generates an estimate for the asset's expected return. This estimate is referred to as the investor's perceived return, and summarizes the information provided by the J return forecasts. Proposition 2 below computes the perceived return and its aggregate uncertainty using the optimal information portfolio.

Proposition 2. *The perceived return implied by the investor's optimal information portfolio weights in Proposition 1 equals*

$$W^T\mu = \frac{\mathbf{1}^T\Theta^{-1}\mu}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}, \tag{13}$$

while the aggregate uncertainty of this estimate is

$$W^T\Theta W = \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}. \tag{14}$$

Proof: The perceived return follows immediately from equation (8) while the aggregate forecast error is computed as[8]

$$\begin{aligned} W^T\Theta W &= \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\mathbf{1}^T\Theta^{-1}\,\Theta\,\Theta^{-1}\mathbf{1}\frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}} \\ &= \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}. \end{aligned} \tag{15}$$

Denote the true return distribution of the asset as $\mathcal{N}(\eta,\nu)$. Ex-ante, the investor is unaware of the asset's true expected return denoted η. As a consequence, the investor is compelled to aggregate the J return forecasts and rely on the perceived return in equation (13). To clarify, $W^T\Theta W$ is not an estimate of ν. Even if $W^T\Theta W$ equals zero or η is known (as in classical portfolio theory), the asset is not necessarily riskless.

Under the assumption in equation (9), a hierarchal structure describes the asset's return distribution, which has an unknown expectation η distributed $\mathcal{N}\left(W^T\mu, W^T\Theta W\right)$. In particular, the distributional assumption imposed on the return forecasts by equation (9) implies that $W^T\mu \overset{d}{\sim} \mathcal{N}\left(\eta, W^T\Theta W\right)$. Therefore, according to Berger (1985), the return distribution of the asset equals[9]

$$\mathcal{N}\left(W^T\mu, \nu + W^T\Theta W\right). \tag{16}$$

[8] A negative portfolio weight implies the investor reverses the sign of this information source's return forecast when the perceived return is computed.

[9] See Lemma 1 in Section 4.2 of Berger (1985).

Thus, ex-ante return uncertainty reflects the asset's true variability as well as the aggregate uncertainty of the J forecasts. Consequently, equation (16) provides further justification for the minimization of $W^T \Theta W$ in equation (7).

2.5 Important Information Portfolio Properties

We begin with the following corollary of Proposition 2 which offers an explicit expression for the information portfolio between two independent information sources.

Corollary 1. *For $J = 2$ and Θ being the diagonal matrix*

$$\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix},$$

the information portfolio W equals

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \left(\frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \right) \begin{bmatrix} \frac{1}{\sigma_1^2} \\ \frac{1}{\sigma_2^2} \end{bmatrix} = \frac{1}{\sigma_2^2 + \sigma_1^2} \begin{bmatrix} \sigma_2^2 \\ \sigma_1^2 \end{bmatrix}. \tag{17}$$

Therefore, the investor's perceived return equals

$$\frac{\sigma_2^2 \mu_1 + \sigma_1^2 \mu_2}{\sigma_1^2 + \sigma_2^2}, \tag{18}$$

while the aggregate uncertainty of this estimate is

$$\frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \tag{19}$$

According to equation (18), the return forecasts of the more accurate information source has a larger portfolio weight and greater influence on the investor's perceived return. Intuitively, the information source with greater historical accuracy has more credibility. The next corollary of Proposition 2 extends Corollary 1 by examining correlated return forecasts.

Corollary 2. *For $J = 2$, let Θ equal*

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

Under this structure, the portfolio weights are

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} \begin{bmatrix} \sigma_2^2 - \sigma_{12} \\ \sigma_1^2 - \sigma_{12} \end{bmatrix}, \tag{20}$$

while the resulting perceived return equals

$$\frac{\sigma_2^2\,\mu_1 + \sigma_1^2\,\mu_2 - \sigma_{12}\,(\mu_1 + \mu_2)}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}. \tag{21}$$

The aggregate uncertainty associated with this estimate of the asset's expected return is

$$\frac{\sigma_1^2\sigma_2^2 - (\sigma_{12})^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}. \tag{22}$$

Appendix B proves that a negative historical covariance, $\sigma_{12} < 0$, reduces the perceived return's aggregate uncertainty in equation (22), a property which parallels its contribution in classical portfolio theory for multiple assets rather than multiple return forecasts for a single asset. In our framework, negative correlation between two information sources represents "offsetting" forecast errors.

3 Historical Accuracy and Return Predictability

Several empirical studies link firm characteristics and periods of uncertainty with behavioral biases originating from the psychology literature.[10] In the context of information portfolio theory, biases in the perceived return are strongest when n is small, which indicates the firm has entered a new "regime" for which little data is available. This setting would also be characterized by high cross-sectional dispersion in equation (5) amongst the return forecasts. Recall that a information source's historical accuracy encompasses forecast variability associated with a state variable as well as uncertainty regarding its return implications, with both of these components being functions of n.

For simplicity, we begin by examining one information source to investigate the impact of n. We then return to a two-information source environment when return predictability is studied. It is important to emphasize that information sources may interpret distinct state variables, have different

[10] This literature includes Zhang (2005) and Kumar (2005). Baker and Wurgler (2005) report that young, small firms are more sensitive to investor sentiment, while Jackson and Johnson (2006) find that momentum and post-earnings announcement drift both result from events that significantly alter a stock's earnings. The composite share issuance variable of Daniel and Titman (2005) also indicates return predictability, while Vassalou and Apedjinou (2004) report that momentum strategies are most profitable for firms with high levels of corporate innovation.

forecasting techniques for these state variables, and utilize unique transformations of these state variables when generating their ex-ante return forecasts. None of these elements underlying a forecast's composition are necessarily disclosed by an information source. Instead, at each point in time, the investor observes a collection of disparate return forecasts issued by various information sources including their own private forecasts. Previous forecast errors for each information source define the Θ matrix underlying our optimal information portfolio.

3.1 Uncertainty in the Return Implications of State Variables

Assume the j^{th} information source utilizes a linear model for converting a known state variable V_t into its return forecast

$$\mu_{j,t} = \hat{\alpha} + \hat{\beta} V_t \, . \tag{23}$$

The hats signify the unknown coefficients of the transformation, while the state variable V_t in equation (23) is not random. Other information sources may or may not utilize this state variable when issuing their return forecast.

According to equation (24) below, the α and β coefficients in equation (23) are calibrated from historical data on realized returns and state variables

$$y_{t-i} = \alpha + \beta V_{t-i} + \xi_{t-i} \, , \tag{24}$$

over the previous $i = 1, \ldots, n$ periods where ξ_{t-i} is an i.i.d. error term distributed $\mathcal{N}\left(0, \sigma_\xi^2\right)$. After estimating equation (24) to obtain $\hat{\alpha}$ and $\hat{\beta}$, the information source utilizes equation (23) to convert V_t into $\mu_{j,t}$.

To illustrate the importance of n, the forecast error ϵ_t^j below contributes another observation to the time series of historical forecast errors in equation (1) at time t. In particular, the $Var\left[\epsilon_t^j\right]$ term in equation (25) below augments the historical accuracy computation in equation (2) at time t[11]

$$
\begin{aligned}
E\left[\epsilon_t^j\right]^2 &= Var\left[y_t - \mu_{j,t}\right] \\
&= Var\left[\xi_t\right] + \left\{Var\left[\hat{\alpha}\right] + (V_t)^2\,Var\left[\hat{\beta}\right] + 2\,V_t\,Cov\left[\alpha - \hat{\alpha}, \beta - \hat{\beta}\right]\right\} \qquad (25) \\
&= \text{Transformation Uncertainty} + \text{Estimation Error in Transformation}\,.
\end{aligned}
$$

For large n, the $\hat{\alpha}$ and $\hat{\beta}$ estimates converge to α and β respectively, implying equation (25) reduces to $Var\left[\xi_t\right]$. As a consequence, when $n \to \infty$, the uncertainty of the j^{th} information source converges to σ_ξ^2. This quantity represents the inherent variability of the state variable to return transformation, and equals the true but unknown $\sigma_{j,*}^2$ variance for the j^{th} information source.[12]

However, when n is small, estimation error in $\hat{\alpha}$ and $\hat{\beta}$ is more severe. Lewellen and Shanken (2002) examine the asset pricing implications of parameter uncertainty and demonstrate that return predictability cannot necessarily be exploited by investors. Indeed, a small n obscures a forecast's true accuracy, undermining the credibility of a knowledgeable information source or a truly relevant state variable. For example, Jagannathan and Wang (2005) find that consumption explains the role of the SMB and HML factors in cross-sectional returns. However, SMB and HML dominate consumption in empirical applications due to the limitations of consumption data.

Equation (25) also illustrates the importance of predictability in the return implications of a state variable to reducing $Var\left[\epsilon_t^j\right]$. Indeed, if the conversion of V_t into $\mu_{j,t}$ is perfectly predictable, then the coefficients in equation (23) would be known rather than estimates, implying the ξ_{t-i} error terms in equation (24) are identically zero. For example, transforming an analyst's price target into a return forecast involves a perfectly predictable function (although not the linear relationship in equation (23)). Conversely, when the relationship between a stock's return and a state variable is unpredictable, this instability causes a forecast to have higher uncertainty.

[11] The $\mu_{j,t}$ return forecast is unbiased since $E\left[y_t - \mu_{j,t}\right] = E\left[\alpha - \hat{\alpha} + V_t\left[\beta - \hat{\beta}\right] + \xi_t\right]$ is zero provided $E\left[\hat{\alpha}\right]$ and $E\left[\hat{\beta}\right]$ equal α and β respectively. These equalities follow from equation (24) providing unbiased estimates of the coefficients according to the (BLUE) theory underlying linear regression models.

[12] To clarify, the investor cannot compute $Var\left[\epsilon_t^j\right]$ in equation (25) at time $t-1$ when the transformation in equation (23) is not disclosed by the information source issuing the j^{th} return forecast. Instead, the investor relies on the j^{th} information sources's historical accuracy which is computed from its previous forecast errors according to equation (2).

Finally, even the idealized environment in equation (25) has two important complications. First, the α and β parameters may be time-varying, complicating their estimation for any value of n. Second, as discussed in the next subsection, V_t could be a forecast for the state variable with its own random evolution. For example, ex-ante usage of the Fama-French (1993) model requires forecasts for the SMB and HML factors.

3.2 Uncertainty in State Variable Dynamics

Jackson and Johnson (2006) document a *post-event drift* in analyst forecasts following seasoned equity offerings, stock re-purchases, equity-financed mergers and dividend initiations as well as omissions. Persistent analyst forecasts indicate that the full implications of such events are not immediately understood.

Suppose the j^{th} return forecast is derived from an information source's forecast for a state variable, denoted $\tilde{V}_t$, which is a linear function of its previous realization

$$\tilde{V}_t = \hat{a} + \hat{b} V_{t-1} , \tag{26}$$

while the true dynamics of V_t are described by

$$V_{t-i} = a + b V_{t-i-1} + \zeta_{t-i} , \tag{27}$$

over the previous $i = 1, \ldots, n$ periods where ζ_{t-i} is another i.i.d. error term whose distribution is $\mathcal{N}\left(0, \sigma_\zeta^2\right)$. Equation (27) is utilized to estimate the a and b coefficients, while the ζ_{t-i} error terms signify the state variable's random evolution. The $\tilde{V}_t$ notation contains a tilde to emphasize that the information source is forecasting the state variable, in contrast to equation (23) where V_t is known.

When equation (23) with known α and β parameters is combined with equation (26), the following return forecast is generated by the j^{th} information source,

$$\begin{aligned} \mu_{j,t} &= \alpha + \beta \tilde{V}_t \\ &= \alpha + \beta \left[\hat{a} + \hat{b} V_{t-1} \right] . \end{aligned} \tag{28}$$

For clarification, the conversion of the state variable into a return forecast continues to be specified by equation (23). However, for simplicity, the α and β coefficients are assumed to be known since our attention is currently focused on the contribution of state variable uncertainty to the j^{th} information source's historical accuracy.[13]

Inserting the true dynamics of the state variable in equation (27) into equation (24) implies that returns evolve as

$$y_t = \alpha + \beta \left[a + bV_{t-1} + \zeta_t\right] + \xi_t \,. \tag{29}$$

When combined, equations (29) and (28) imply[14]

$$\begin{aligned} E\left[\epsilon_t^j\right]^2 &= Var\left[y_t - \mu_{j,t}\right] \\ \\ &= Var\left[\xi_t\right] + \beta^2 Var\left[\zeta_t\right] \\ &+ \left\{\beta^2 Var\left[\hat{a}\right] + \beta^2 \left(V_{t-1}\right)^2 Var\left[\hat{b}\right] + 2\beta^2 V_{t-1} Cov\left[a - \hat{a}, b - \hat{b}\right]\right\} \\ \\ &= \text{Transformation Uncertainty} + \text{State Variable Uncertainty} \\ &+ \text{Estimation Error in State Variable Dynamics}\,. \end{aligned} \tag{30}$$

To clarify, the term $Var\left[\zeta_t\right]$ corresponds to state variable uncertainty, while $Var\left[\xi_t\right]$ represents randomness in the return implications of the state variable.

The estimation error in equation (30) tends toward zero as $n \to \infty$, implying $\sigma_{j,t+1}^2$ converges to $\sigma_\xi^2 + \beta^2\sigma_\zeta^2$ which equals the j^{th} information source's true accuracy $\sigma_{j,*}^2$ in this economy. This information environment involves an infinite amount of relevant return and state variable time series data. Furthermore, if there is no uncertainty regarding the evolution of the state variable or its return implications, then σ_ξ^2 and σ_ζ^2 are both zero. In the limit of this no uncertainty environment, $\sigma_{j,t+1}^2$ converges to zero and the asset's true expected return η is eventually revealed.

An important property of equation (30) is that predictability in V_t reduces $Var\left[\epsilon_t^j\right]$. As a consequence, after controlling for $Var\left[\xi_t\right]$, return forecasts derived from a predictable state variable are more accurate. Conversely, conditioning a return forecast on an unpredictable state variable yields less

[13]The j superscript also applies to the α, β, a and b coefficients as well as the ξ and ζ error terms but is omitted for notational simplicity.

[14]The linearity of equations (26), (28) and (29) imply $\mu_{j,t}$ is an unbiased estimate of y_t. Specifically, $E\left[\epsilon_t^j\right]$ equals zero since the unbiased estimates produced by equation (27) ensure $E\left[\hat{a}\right] = a$ and $E\left[\hat{b}\right] = b$.

accurate information source. More formally, consider two state variables; the first being unpredictable and the second highly predictable. Given identical $Var[\xi_t]$ terms, return forecasts derived from the highly predictable state variable are more accurate. As an extreme example, suppose V_t is perfectly predictable and follows a known deterministic process with $\hat{V}_t \equiv V_t$ as in the previous subsection. For this special case, $Var[\zeta_t]$ equals zero while the a and b coefficients are known, implying equation (30) reduces to $Var[\xi_t]$.

In general, state variables are not necessarily normally distributed, nor are their return implications required to be linear transformations. Indeed, no assumptions are imposed on the conversion of a state variable into a return forecast when solving for the information portfolio.

3.3 Return Predictability

Consider an array of return forecasts ranging from optimistic to pessimistic in response to a firm initiating an investment. Thus, the profitability (earnings / cashflow) of the investment may be thought of as the relevant state variable.[15]

We study a simple two-period, two information source example. At the initial timepoint t_1, the firm announces a large investment. Over the subsequent $(t_1, t_2]$ interval, the high return forecast is $\mu_{H,1}$ while its low return counterpart is $\mu_{L,1}$. At t_1, assume the two return forecasts are equally credible, implying $\sigma^2_{H,1}$ equals $\sigma^2_{L,1}$. For simplicity, the two information sources are independent with both information sources issuing simultaneous forecasts. According to Corollary 1, the investor's perceived return over the $(t_1, t_2]$ horizon is $\mu_1 = \frac{1}{2}[\mu_{H,1} + \mu_{L,1}]$ while the return realized at t_2 equals $r_{1,2}$.

By time t_2, partial information regarding the success of the investment is revealed. In particular, there are two scenarios, the first indicating success and the second failure. The ex-ante probability attached to these scenarios is irrelevant when the ex-post return sequence is studied at t_3. Furthermore, assume $\mu_{H,2}$ continues to exceed $\mu_{L,2}$ although the disparity between these forecasts at time t_2 may be lower than at t_1 depending on the prevailing uncertainty regarding the investment's profitability.

[15] Berk, Naik and Green (1999) as well as Gomes, Kogan and Zhang (2003) and Carlson, Fisher and Giammarino (2004) investigate the influence of investment decisions on cross-sectional return characteristics.

Success Revealed:

$$\begin{cases} r_{1,2} \text{ is high} \\ \sigma^2_{H,2} < \sigma^2_{L,2}, \text{ high return forecast is more accurate since } r_{1,2} \text{ is high} \\ \text{Perceived return } \mu_2 \text{ over } (t_2, t_3] \text{ horizon is closer to } \mu_{H,2} \\ \hline \mu_2 \text{ reflects extrapolation from high } r_{1,2} \end{cases}$$

Failure Revealed:

$$\begin{cases} r_{1,2} \text{ is low} \\ \sigma^2_{L,2} < \sigma^2_{H,2}, \text{ low return forecast is more accurate since } r_{1,2} \text{ is low} \\ \text{Perceived return } \mu_2 \text{ over } (t_2, t_3] \text{ horizon is closer to } \mu_{L,2} \\ \hline \mu_2 \text{ reflects extrapolation from low } r_{1,2} \end{cases}$$

A small n would increase the disparity between $\sigma^2_{H,2}$ and $\sigma^2_{L,2}$, causing either $\mu_{H,2}$ or $\mu_{L,2}$ to have a greater influence on μ_2 for a given realized return $r_{1,2}$. Thus, μ_2 is closer to either the high or low return forecast at t_2 when n is small.

One may argue that the empirical evidence concerning long term reversals motivates a mean-reverting prior distribution when issuing return forecasts. However, when the optimistic (pessimistic) information source at t_1 decreases (increases) their return forecasts at t_2, uncertainty is reduced. Indeed, if $\mu_{H,1}$ and $\mu_{L,1}$ converge to a common return forecast μ_* at t_2, then the return uncertainty created by the investment is resolved and equation (5) is zero. Therefore, extrapolation *attributable* to information portfolio theory continues as long as there is return uncertainty.

Intuitively, an sequence of returns may exhibit predictability as a result of events whose return implications are not immediately understood and agreed upon by all information sources. This finding does not prevent the return forecasts from being updated by the information sources. Instead, only the continuation of forecast dispersion beyond one period is required.[16] Furthermore, greater predictability in the profitability of the investment implies the disparity between the return forecasts would narrow more rapidly. In particular, our discussion in the previous subsection for a general state variable

[16]The horizon between the issuance of forecasts is important. Long time intervals allow more uncertainty to be resolved before the investor's perceived return is adjusted.

implies predictability in the investment's profitability reduces return uncertainty.

In summary, high (low) return forecasts are more accurate following high (low) realized returns. Therefore, according to information portfolio theory, return forecasts are assigned larger information portfolio weights when they are similar to previous return realizations. As a consequence, the perceived return appears to be *extrapolated* from past returns. Intuitively, many return sequences exist ex-ante, with the investment's success determining a particular realized return sequence. Similarly, Bondarenko and Bossaerts (2000) provide an excellent description of the return bias induced by conditioning on an option's eventual in-the-money or out-of-the-money status.

3.4 Conditional Expectations and Forecast Heterogeneity

The law of iterated expectations is usually invoked to conclude that the "error" separating an expected conditional return and its realization has mean zero. However, information portfolio theory allows different information sources to utilize distinct statistical methodologies when forecasting state variables or modeling their return implications. Disparate return forecasts also originate from information sources analyzing different state variables. Consequently, the law of iterated expectations does not ensure homogenous return forecasts across the J information sources.

Therefore, the potential for disagreement regarding future state variables and their impact on the asset's return is not ignored by our framework since a non-zero value for equation (5) enables forecast heterogeneity to justify the existence of multiple information sources. From a practical perspective, we assume the asset's expected return is sufficiently complex to prevent market participants from fully agreeing upon the η parameter. This assumption would also generate transactions between market participants.

Conversely, the standard econometric approach when testing market efficiency restricts itself to a single return forecast, which implicitly assumes that σ_μ^2 in equation (5) is zero. Specifically, a single multifactor model is calibrated to generate conditional expected returns, a procedure which fails to account for uncertainty surrounding the interpretation of available information. Indeed, this methodology assumes there is no disagreement regarding the forecasts of the factors next period. Therefore, if earnings are relevant to factor dynamics, the reality of earnings forecast dispersion is ignored. Furthermore, the beta coefficient for every factor is constrained to be identical across all information sources, who are further assumed to employ the same multifactor model.

As a final clarification, our model's structure allows the return forecasts to be determined by an asset's correlation with factor portfolios. Under this formulation, the asset's true expected return is unknown for at least one of three reasons; randomness in the dynamics of the factors, estimation error in the factor sensitivities, and uncertainty regarding the number of required factors. The factors can be interpreted as state variables or as intermediaries between state variables and return forecasts. As an example, the market return could serve as a state variable. Alternatively, information sources could utilize state variables to predict the market factor, whose return implications are obtained from calibrating a firm's beta coefficient.

4 Information Portfolios and Perceived Return Biases

This section connects our optimal information portfolio with several return biases previously attributed to investor psychology. In particular, we demonstrate that the *appearance* of overconfidence, biased self-attribution, representativeness, conservatism and limited attention are instilled into the perceived return by the optimal information portfolio. However, none of the information sources nor the investor are assumed to be influenced by psychological biases in our analysis.

4.1 Appearance of Overconfidence and Biased Self-Attribution

To analyze the appearance of overconfidence in the perceived return, we examine with two information sources; one private and one public whose moments are denoted with pr and pb subscripts respectively. A bias in the investor's perceived return which mimics overconfidence is produced whenever the information portfolio weight w_{pr} for a private information source exceeds the information portfolio weight w_{pb} of a public information source.

Interpretation 1. *Appearance of Overconfidence*

Corollary 1 implies the following information portfolio weights for private and public information

$$\begin{bmatrix} w_{pb} \\ w_{pr} \end{bmatrix} = \frac{1}{\sigma_{pr}^2 + \sigma_{pb}^2} \begin{bmatrix} \sigma_{pr}^2 \\ \sigma_{pb}^2 \end{bmatrix} . \tag{31}$$

Consequently, private information is overweighted with w_{pr} exceeding w_{pb} whenever $\sigma_{pr}^2 < \sigma_{pb}^2$. Furthermore, the perceived return would equal

$$\frac{1}{\sigma_{pr}^2 + \sigma_{pb}^2} \left[\sigma_{pr}^2 \, \mu_{pb} + \sigma_{pb}^2 \, \mu_{pr}\right] , \tag{32}$$

which emphasizes μ_{pr} more than μ_{pb}.

According to equation (32), *whenever* a private information source is more accurate than its public counterpart, the investor's perceived return appears to exhibit overconfidence. Recall that several private information sources can originate from the investor. Indeed, forecasts for state variables such as earnings require further interpretation to become return forecasts, and create additional private information sources. In contrast, when there is no uncertainty surrounding the return implications of publically available information, the number of private information sources is reduced. Overall, let the return forecast μ_{pr} in equation (32) be associated with the investor's most accurate source of private information. Provided this private source is more accurate than the public information source over the last n periods, the investor appears overconfident.

The next interpretation of our optimal information portfolio extends the above analysis to include the appearance of biased self-attribution. In the context of information portfolio theory, the investor exhibits the appearance of biased self-attribution when one of their private information sources is more accurate than a public information source, while another private information source is less accurate according to equation (2). These private information sources are denoted by c and d subscripts respectively and supplement the original public information source with J increasing to three.[17]

Interpretation 2. *Appearance of Overconfidence with Biased Self-Attribution*
Consider the variance-covariance matrix

$$\begin{bmatrix} \sigma_c^2 & 0 & 0 \\ 0 & \sigma_d^2 & 0 \\ 0 & 0 & \sigma_{pb}^2 \end{bmatrix} ,$$

with the property that $\sigma_d^2 > \sigma_{pb}^2 > \sigma_c^2$. The corresponding information portfolio equals

$$[w_c, \, w_d, \, w_{pb}] \;=\; \frac{1}{D} \left[\sigma_d^2 \, \sigma_{pb}^2, \, \sigma_c^2 \, \sigma_{pb}^2, \, \sigma_c^2 \, \sigma_d^2\right] ,$$

[17] With respect to terminology in the psychology literature, confirming private information sources are more accurate than a public information source, while disconfirming private information sources have been less accurate than all public information sources over the last n periods.

where D is defined as $D = \sigma_d^2 \, \sigma_c^2 + \sigma_d^2 \, \sigma_{pb}^2 + \sigma_c^2 \, \sigma_{pb}^2$. Therefore, the perceived return $W^T \mu$ equals

$$\frac{\left[\sigma_c^2 \, \mu_d + \sigma_d^2 \, \mu_c\right] \sigma_{pb}^2 + \sigma_c^2 \, \sigma_d^2 \, \mu_{pb}}{D}\,, \tag{33}$$

and is influenced more by μ_c than μ_d.

The $\sigma_d^2 > \sigma_c^2$ property ensures the information portfolio weight for μ_c exceeds the information portfolio weight of μ_d. Thus, more accurate private information sources have more influence over the investor's perceived return. Interestingly, the investor may appear overconfident even if their private information sources are inaccurate on average since inaccurate private information sources receive small information portfolio weights. For example, if the investor successfully predicts the return implications of industry characteristics, but performs poorly on other dimensions such as a firm's individual earnings, then the importance of industry data is accentuated by the information portfolio. By implication, the investor pursues trading strategies derived from private information sources which have provided them with *individual* success, regardless of the technique's generality.

Overall, the investor's perceived return gravitates towards the most accurate private information sources and away from those which are less accurate. This tendency causes the investor's perceived return to exhibit the appearance of overconfidence and biased self-attribution as a result of the optimal information portfolio.

These implications of the optimal information portfolio are formalized below with two private information sources whose historical accuracies are denoted $\sigma_{pr,1}^2$ and $\sigma_{pr,2}^2$ respectively, along with a public information source whose accuracy equals σ_{pb}^2. Let p_j represent the probability that the j^{th} private information source is less accurate than its public counterpart, $\sigma_{pr,j}^2 > \sigma_{pb}^2$, after n periods according to equation (2). These probabilities capture the skill of the information sources. Recall from the previous section that assessing an information source's skill is more difficult when n is small, implying $p_j \approx \frac{1}{2}$. The following four scenarios summarize the comparative historical accuracies of the three information sources after n periods.

Scenario	Historical Accuracies					Probability	Investor Appears to Exhibit
A	$\sigma^2_{pr,1}$, $\sigma^2_{pr,2}$	$<$	σ^2_{pb}	$<$	Neither	$(1-p_1)\,(1-p_2)$	Overconfidence from both private sources
B	$\sigma^2_{pr,1}$	$<$	σ^2_{pb}	$<$	$\sigma^2_{pr,2}$	$(1-p_1)\,p_2$	Overconfidence, 1^{st} private source is confirming
C	$\sigma^2_{pr,2}$	$<$	σ^2_{pb}	$<$	$\sigma^2_{pr,1}$	$p_1\,(1-p_2)$	Overconfidence, 2^{nd} private source is confirming
D	Neither	$<$	σ^2_{pb}	$<$	$\sigma^2_{pr,1}$, $\sigma^2_{pr,2}$	$p_1\,p_2$	No Overconfidence

Therefore, a bias which mimics overconfidence arises in scenarios A, B and C. These scenarios have a cumulative probability of

$$(1-p_1)\,(1-p_2) + (1-p_1)\,p_2 + p_1\,(1-p_2) \quad = \quad 1 - p_1\,p_2\,. \tag{34}$$

Suppose private and public information sources are equally accurate with $p_1 = p_2 = \frac{1}{2}$. In this situation, the probability of a bias mimicking overconfidence is 75%. The investor appears to exhibit the greatest amount of overconfidence in scenario A where both private information sources are more accurate than the public information source. Furthermore, in scenarios B and C, a historically accurate (inaccurate) private information source is assigned a larger (smaller) portfolio weight than the public information source. Therefore, the probability that biased self-attribution appears to influence the investor's perceived return equals 50%.

Finally, our results continue to apply when there are more public than private information sources.[18] When two public information sources and one private information source are available, the following four scenarios are available after n periods.

[18] The relationship between the number of *private* information sources and their accuracy is ambiguous. More private information sources could increase the likelihood of at least one private information source being more accurate than the public information source. Conversely, additional private information sources may diminish the resources allocated to generating each return forecast and thereby decrease their accuracy.

Scenario	Historical Accuracies					Probability	Investor Appears to Exhibit
A	$\sigma^2_{pb,1}\,,\,\sigma^2_{pb,2}$	$<$	σ^2_{pr}	$<$	Neither	$p_1\,p_2$	No Overconfidence
B	$\sigma^2_{pb,1}$	$<$	σ^2_{pr}	$<$	$\sigma^2_{pb,2}$	$p_1\,(1-p_2)$	Limited Overconfidence from 2^{nd} public source
C	$\sigma^2_{pb,2}$	$<$	σ^2_{pr}	$<$	$\sigma^2_{pb,1}$	$p_2\,(1-p_1)$	Limited Overconfidence from 1^{st} public source
D	Neither	$<$	σ^2_{pr}	$<$	$\sigma^2_{pb,1}\,,\,\sigma^2_{pb,2}$	$(1-p_1)\,(1-p_2)$	Overconfidence from both public sources

The concept of limited overconfidence in scenarios B and C reflects the private information source's larger portfolio weight relative to one of the public information sources. Only in scenario A when the private information source is less accurate than both public information sources is there no evidence of overconfidence.

4.2 Appearance of Representativeness and Conservatism

Recall from the previous section that accuracy increases as a result of predictability in state variable dynamics as well as predictability in their return implications. These properties imply that *trends*, whether in the state variables or their return implications, are capable of increasing an information source's portfolio weight since trends imply predictability. For example, a strong trend in a binomial sequence consists predominately of either up or down movements, implying the estimated binomial probability over the last n observations would be near 0 or 1.

Consider two information sources, labeled consistent and inconsistent, with the former arising from predictable state variables or stability in the return implications of the state variables. The moments for the consistent and inconsistent information sources are denoted μ_C and σ^2_C as well as μ_I and σ^2_I respectively. The property $\sigma^2_I > \sigma^2_C$ is induced by predictability underlying the consistent information source.

Interpretation 3. *Appearance of Representativeness*

Let $\mu = \begin{bmatrix} \mu_C \\ \mu_I \end{bmatrix}$ *and* $\Theta = \begin{bmatrix} \sigma_C^2 & 0 \\ 0 & \sigma_I^2 \end{bmatrix}$. *From Corollary 1, the information portfolio equals*

$$\begin{bmatrix} w_C \\ w_I \end{bmatrix} = \frac{1}{\sigma_I^2 + \sigma_C^2} \begin{bmatrix} \sigma_I^2 \\ \sigma_C^2 \end{bmatrix}, \tag{35}$$

and implies the perceived return

$$\frac{1}{\sigma_C^2 + \sigma_I^2} \left[\sigma_C^2 \, \mu_I + \sigma_I^2 \, \mu_C \right], \tag{36}$$

is influenced more by μ_C *than* μ_I.

Hence, consistent information sources have more influence on the investor's perceived return than their inconsistent counterparts. However, trends can produce consistent information sources in the absence of any superior knowledge regarding the true dynamics of a state variable or its relationship with future returns. Thus, an information source's consistency may be temporary, especially when n is small. For example, prior returns can generate a consistent information source for an IPO over a short horizon until the dynamics of a firm's earnings are able to be reliably estimated.

Assume the return forecasts from two information sources both emanate from the true return distribution, which is further assumed to be stationary. These two information sources have identical levels of theoretical accuracy, implying that predictability in a state variable or stability in its return implications does not offer an advantage to one information source versus the other. Therefore, any trend that causes one of the two information sources to be more accurate after n periods is statistically insignificant. Nonetheless, for a finite n, one of the information sources is always more accurate.[19]

Scenario	Consistent		Inconsistent	Probability	Investor Appears to Exhibit
A	σ_1^2	$<$	σ_2^2	$\frac{1}{2}$	Representativeness; 1^{st} source consistent, 2^{nd} inconsistent
B	σ_2^2	$<$	σ_1^2	$\frac{1}{2}$	Representativeness; 2^{nd} source consistent, 1^{st} inconsistent

[19] The return forecasts originate from a (normal) continuous distribution. Consequently, the probability that σ_1^2 equals σ_2^2 after n periods is zero.

Observe that the appearance representativeness occurs in both scenarios, while its apparent magnitude is proportional to the disparity $|\sigma_1^2 - \sigma_2^2|$. With both historical accuracies computed according to equation (2), this distance decreases as n increases since the return forecasts from both information sources arise from the same distribution.

Furthermore, the investor's perceived return may appear insensitive to the release of new information. As illustrated below, even for the simplest case where two information sources are available, the perceived return has four degrees of freedom.

Interpretation 4. *Appearance of Conservatism*
According to Corollary 1, an infinite number of μ_C, μ_I, σ_C^2 and σ_I^2 combinations that result in the same perceived return,

$$\frac{\sigma_C^2\,\mu_I + \sigma_I^2\,\mu_C}{\sigma_C^2 + \sigma_I^2}\,.$$

Therefore, conservatism cannot be established without evaluating multiple sources of information since the investor's perceived return is an aggregate quantity.

For example, the announcement of poor earnings coupled with strong sales may occur after a large investment or price discounting. On the announcement date, the investor's perceived return need not fluctuate as much as the individual return forecasts derived from earnings and sales due to their *negative* correlation. As demonstrated in the next subsection when a bias which mimics limited attention is investigated, negatively correlated forecasts receive larger information portfolio weights. Consequently, the investor's perceived return does not necessarily respond significantly to a *single* news item, unless all of the return forecasts are updated in a similar direction after its release.

4.3 Appearance of Limited Attention

The $\sigma_{1,2}$ covariance term in Corollary 2 incorporates the appearance of limited attention bias into the perceived return. Barber, Odean and Zhu (2003) present empirical evidence of this bias for individual investors.

Figure 1 illustrates the response of the perceived return and its uncertainty in equations (21) and (22) respectively as a function of the correlation between two forecasts. Observe that forecast correlation has a dramatic impact on the investor's aggregate uncertainty but less influence on their perceived return. Appendix B formalizes this assertion by computing the partial derivatives of the perceived return and its aggregate uncertainty in Corollary 2 with respect to σ_{12}. Intuitively, the investor ignores an information source whose return forecasts are positively correlated with more accurate information sources. This behavior parallels the removal of independent variables in linear regression models due to multicollinearity.

For example, if two analysts are simultaneously optimistic or pessimistic, then the investor may limit their attention to a single *representative* information source.[20] In contrast, if these information sources offer alternative perspectives on the asset's expected return, then the investor benefits from analyzing both forecasts since their errors have historically been negatively correlated. More formally, consider the portfolio weights in Corollary 2

$$\begin{aligned} w_1 &= \frac{\sigma_2^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} > 0 \\ w_2 &= \frac{\sigma_1^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} > 0\,. \end{aligned} \tag{37}$$

If the two forecasts are independent, then σ_{12} equals zero and both portfolio weights are positive. However, suppose that σ_{12} equals σ_1^2. The Cauchy-Schwartz inequality implies that $\sigma_1^2 \leq \sigma_2^2$ when $\sigma_{12} = \sigma_1^2$. Thus, the first information source is more accurate than the second.[21] More importantly, when the covariance equals σ_1^2, the portfolio weights in equation (37) become

$$w_1 = \frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2 + \sigma_1^2 - 2\sigma_1^2} = \frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2 - \sigma_1^2} = 1 \tag{38}$$

$$w_2 = \frac{\sigma_1^2 - \sigma_1^2}{\sigma_2^2 + \sigma_1^2 - 2\sigma_1^2} = 0\,. \tag{39}$$

Consequently, positive covariance between the two information sources can eliminate the second return forecast from the perceived return, with a visual illustration in Figure 2. Thus, investors tend to ignore return forecasts which are positively correlated with more accurate information sources, while negatively correlated forecasts have the greatest influence over the investor's perceived return.

[20] Optimism (pessimism) is defined by positive (negative) forecast errors.

[21] The Cauchy-Schwartz inequality provides an upper bound on the covariance, $(\sigma_{12})^2 \leq \sigma_1^2\,\sigma_2^2$. Therefore, when $\sigma_{12} = \sigma_1^2$, this inequality is interpreted as $\left(\sigma_1^2\right)^2 \leq \sigma_1^2\,\sigma_2^2$ which implies $\sigma_1^2 \leq \sigma_2^2$.

Therefore, when attempting to detect conservatism, it is essential to evaluate the aggregate impact of contradictory information.

In summary, the number of return forecasts the investor processes depends on their correlation structure. Thus, the investor may rely on broadly defined sector information rather than firm-specific characteristics if the latter are positively correlated within an industry. For example, during the Internet bubble, the returns of dot-com firms appear to have been driven by industry characteristics. In addition, earnings forecasts issued by analysts who herd are less likely to influence an asset's perceived return.

4.4 Rational versus Behavioral Interpretations

Although the perceived return is derived from the optimal information portfolio, this estimate of the asset's expected return is not referred to as being *rational* since the return forecasts may or may not incorporate investor psychology. In particular, the most accurate information sources could be those which incorporate investor psychology into their return forecasts. As a consequence, information portfolio theory does not preclude behavioral biases from influencing the perceived return. For example, suppose all J return forecasts are identical and equal to μ_*, with this common expectation further *assumed* to be the result of at least one psychological bias. In this situation, the investor's perceived return μ_* contains this behavioral bias regardless of the information portfolio. Conversely, the use of psychology could increase forecast dispersion due to disagreements over the exact nature of the biases committed by investors. As a result, the relevance of information portfolio theory is enhanced by differences of opinion regarding investor psychology. Overall, decomposing the perceived return into the effects of psychology versus the optimal information portfolio is ultimately an empirical question. Our objective in this paper is to demonstrate that return uncertainty can instill biases into perceived returns which mimic those in the psychology literature.

In addition, information portfolio theory facilitates the computation of an investor's expected return, which enhances rather than contradicts utility maximization. The extent to which information uncertainty is priced by the market is not addressed in this paper since additional assumptions on investor utility are not invoked to convert the results of Proposition 2 into a dollar-denominated price. Instead, we limit our attention to the investor's perceived return, a critical input for any pricing application. However, Proposition 3, below whose proof is in Appendix C, provides a utility maximizing application of Proposition 2 after imposing the distributional assumption in equation (9) on the return forecasts.

Proposition 3. *Suppose the investor has initial wealth M and a negative exponential utility function whose coefficient of absolute risk aversion is denoted a, $U(M) = 1 - e^{-aM}$. Given the return distribution in equation (16), the optimal fraction of wealth f invested in the risky asset equals*

$$f \;=\; \frac{\mathbf{1}^T \Theta^{-1} (\mu - r_f \mathbf{1})}{a\, M \; [1 + \nu \mathbf{1}^T \Theta^{-1} \mathbf{1}]} \,, \tag{40}$$

where r_f represents the riskfree interest rate.

As an explicit illustration, consider two correlated identical return forecasts issued by information sources with identical historical accuracies ($\mu_1 = \mu_2 = \mu_*$, $\sigma_1^2 = \sigma_2^2 = \sigma_*^2$) with σ_{12} describing the off-diagonal sample covariance element of Θ as in Corollary 2. Under these specifications, the solution for f in Proposition 3 reduces to

$$f \;=\; \frac{2\,(\mu_* - r_f)}{a\, M} \left(\frac{1}{2\nu + \sigma_*^2 + \sigma_{12}} \right) . \tag{41}$$

Observe that when information sources forecast higher returns or are more accurate historically, the investor increases their exposure to the risky asset. When the return forecasts are negatively correlated, the investor also purchases a larger amount of the risky asset. According to equation (41), accurate return forecasts offset the investor's risk aversion. Therefore, it is difficult to distinguish the influence of time-varying risk aversion on the investor's portfolio allocation from the quality of available forecasts. In particular, more return uncertainty proxies for higher risk aversion. Consequently, inaccurate information sources cause the investor to reduce their exposure to the risky asset.

As a special case of equation (41), the fraction of wealth allocated to the risky asset equals $\frac{\eta - r_f}{a\, M\, \nu}$ when there is no uncertainty regarding the asset's expected return since σ_*^2 and σ_{12} are zero while $\mu_* = \eta$.

4.5 Contrast with Bayesian Methods

When an asset's expected return is *unstable*, Brav and Heaton (2002) demonstrate that representativeness and conservatism result from Bayesian priors which underweight past or recent observations respectively. In their model, these biases arise from uncertainty regarding a random *change point* which initiates a different economic regime. Therefore, they highlight the difficulty posed by different possible priors when attempting to disentangle rational from behavioral explanations of return patterns. Furthermore, overconfidence may be inserted directly into the prior distribution of a private return forecast by assuming the investor underestimates its uncertainty.

In contrast, our biases arise from aggregating across multiple return forecasts. Indeed, biases are outputs of information portfolio theory, enabling us to provide testable implications independent of any prior distribution. Information portfolio theory also requires at least two return forecasts. Although Bayesian updating is applicable to multiple forecasts, a prior distribution(s) remains an integral part of the posterior and therefore the investor's expected return. In contrast, the historical accuracies of our information sources are derived entirely from return forecasts and realizations.

5 Empirical Implementation

Information portfolio theory does not assume that investor beliefs are characterized by behavioral biases in the psychology literature. Instead, uncertainty surrounding an asset's expected return increases the potential for the perceived return to manifest these biases. In this section, testable implications of information portfolio theory are discussed and verified empirically.

5.1 Testable Implications

There are several testable implications of information portfolio theory. These include hypotheses that enable us to distinguish information portfolio theory from psychology.

First, biases are more pronounced when an asset's expected return is uncertain. Therefore, in the aftermath of events which undermine the relevance of previous return forecasts, biases in the perceived return are stronger. Corporate restructurings, significant investments as well as technological innovations all reduce the relevance of previous observations. Therefore, these events reduce n and are likely to increase return forecast dispersion. Second, the investor may condition their beliefs on forecasts derived from state variables which are correlated with returns, regardless of their theoretical justification, provided they lower forecast errors. This tendency is aggravated when n is small. Third, negatively correlated return forecasts reduce the investor's aggregate forecast error. In contrast, a return forecast is assigned a lower portfolio weight if its forecast errors are positively correlated with another more accurate information source. Fourth, assuming two state variables can be transformed into returns with comparable accuracy, the more predictable state variable has a greater influence on the perceived return.

Two testable hypotheses involving private return forecasts are also available. First, investors overweight their (confirming) accurate private information sources at the expense of their (disconfirming) failures. Consequently, the trading strategies implemented by an investor are determined by the success of their private return forecasts. Second, less experienced investors have a greater propensity to exhibit overconfidence and biased self-attribution since they have produced fewer forecast errors to assess their ability. This implication once again corresponds to a small n proxying for a poor information environment.[22]

To distinguish between the implications of behavioral theories versus information portfolio theory, the accuracy associated with each information source is crucial. When investors focus their attention on accurate information sources, our theory (at least partially) describes their perceived return. Thus, ranking information sources by their historical accuracy facilitates empirical tests of information portfolio theory.

However, psychology and information portfolio theory are not incompatible. The extent to which they both influence the perceived return is ultimately an empirical question. For example, if the most accurate information sources incorporate investor psychology into their return forecasts, then the role of psychology in the perceived return is undeniable. In this economy, return uncertainty causes the optimal information portfolio to augment, rather than completely explain, the presence of biases in the perceived return.

5.2 Hypotheses and Data

For testing the two dimensions of information portfolio theory, we examine the transformation of earnings to returns (sensitivity) and the variability (uncertainty) of earnings. In our empirical exercise, there are implicitly two information sources; earnings and everything unexplained by earnings.

The relationship between returns and earnings forecasts generates our first hypothesis. Intuitively, the ξ_t error terms in equation (24) are being referenced. Indeed, if returns and earnings are perfectly

[22] These biases may be investigated when individual analysts are interpreted as issuing private return forecasts. Thus, analysts that herd are less confident, while those issuing bold forecasts are more confident.

correlated, then $Var[\xi_t]$ is zero for this transformation.

Hypothesis 1. *Investors focus more on earnings when this state variable has a stronger relationship with earnings. In particular, the return forecast derived from earnings is assigned a higher information portfolio weight when the return implications of earnings are more certain.*

Our second hypothesis concerns earnings uncertainty and refers to the ζ_t error terms in equation (27). For example, if earnings evolve according to a known deterministic process agreed upon by all analysts, then $Var[\zeta_t]$ is zero.

Hypothesis 2. *Assets with greater earnings uncertainty exhibit stronger biases. In particular, return forecasts are more disparate when earnings dispersion is higher.*

The first hypothesis is critical to verifying the predictions of information portfolio theory since it is unique to our framework, while the second hypothesis is consistent with several behavioral interpretations (Zhang (2005)). According to our first hypothesis, even during periods of high return uncertainty, information portfolio theory predicts that investors attempt to find the "best" available information sources.

Analyst earnings forecasts serve as public information in our empirical tests. We consider all domestic primary stocks listed on the NYSE, AMEX and NASDAQ with analyst coverage. The monthly stock return and market capitalization data are obtained from CRSP while analyst forecasts are from the I/B/E/S Summary History dataset. The intersection of the CRSP and I/B/E/S datasets over the January, 1976 to December, 2004 sample period is utilized. The start date is determined by the beginning of the I/B/E/S Summary History dataset. Forecast revisions are scaled by stock prices retrieved from I/B/E/S to account for adjustments such as stock dividends and stock splits. Finally, we obtain book to market ratios (B/M) from Compustat.

We construct an uncertainty measure to proxy for the return dispersion in equation (5) as well as a sensitivity measure to gauge the relative informativeness of earnings versus everything else when forecasting returns.

5.3 Sensitivity of Returns to Forecast Revisions

Each month, we estimate stock price sensitivities to earnings information by computing the correlation coefficient between stock returns and forecast revisions over the previous twelve months.[23] These correlations proxy for the return implications of analyst forecasts. In particlar, stocks with higher correlations are more influenced by earnings since our sensitivity measure parallels the transformation from an earnings state variable into a return signal.

I/B/E/S contains summary statistics on analyst forecasts for the third Thursday of each month (referred to as the I/B/E/S compilation date hereafter). We define the forecast revision for firm i in month t as

$$rev_{i,t} = \frac{FY1_{i,t} - FY1_{i,t-1}}{P_{i,t}}, \qquad (42)$$

where $FY1_{i,t}$ and $FY1_{i,t-1}$ are the mean analyst forecast for fiscal year 1 in month t and $t-1$ respectively, while $P_{i,t}$ is the stock price provided by I/B/E/S on the compilation date in month t.[24] For each $rev_{i,t}$, we compute the contemporaneous stock return $ret_{i,t}$ defined as the return of stock i between two I/B/E/S compilation dates in month $t-1$ and month t. Once again, the stock prices on the I/B/E/S compilation dates are extracted from I/B/E/S.

Using the monthly forecast revisions and stock returns, we then find the return-forecast sensitivity of stock i in month t by computing the correlation coefficient between rev_i and ret_i over the past 12 months. Based on this sensitivity measure, the stocks are sorted into three groups every month consisting of the bottom 30%, middle 40% and top 30% respectively. For ease of illustration, these three groups are labeled low sensitivity (S1), medium sensitivity (S2) and high sensitivity (S3) stocks.

[23] We also estimate the correlation coefficient using observations from the previous 6 and 24 months. Our results are robust to these alternative estimates of the correlation coefficient.

[24] Additional adjustments on $rev_{i,t}$ are performed in the month when a firm announces its fiscal year earnings since analyst forecasts switch to subsequent fiscal years after the announcement. Thus, the $FY1$ estimates in two consecutive months could be forecasts for two different fiscal years. For example, suppose a firm announces its fiscal year earnings in month t. If the announcement date is before the I/B/E/S compilation date in that month, $rev_{i,t}$ is defined as its mean $FY1$ estimate in month t minus its mean $FY2$ estimate in month $t-1$. Conversely, if the announcement occurs after the I/B/E/S compilation date in that month, then $rev_{i,t}$ remains defined as the difference in the mean $FY1$ estimates between month t and $t-1$. However, $rev_{i,t+1}$ is defined as the mean $FY1$ estimate in month $t+1$ minus the mean $FY2$ estimate in month t.

5.4 Earnings Uncertainty

Our theory also asserts that biases are more severe when state variables are more uncertain. The uncertainty of earnings information is measured using the standard deviation of analyst forecasts scaled by stock price[25]

$$stdev_{i,t} = \frac{\sigma_{i,t}}{P_{i,t}}. \tag{43}$$

Along with the sensitivity classifications, we divide the stocks into three uncertainly groups each month according to equation (43) which are comprised of the bottom 30%, middle 40% and top 30%. These three groups are referred to as low uncertainty (U1), medium uncertainty (U2) and high uncertainty (U3) stocks.

Table 1 provides an overview of the sensitivity and uncertainty portfolios. Furthermore, we investigate whether there are significant differences among the portfolios in terms of value/growth and large/small characteristics as well as analyst coverage. The Spearman rank correlation coefficients among the sensitivity measure, the uncertainty measure, B/M, size and the number of analysts are computed each month, with their time series average reported in Panel A. Each month we also compute the average rankings of B/M, size and number of analysts for the stocks in the sensitivity and uncertainty portfolios. The ranking is normalized to [0, 1]. Thus, a ranking of 0.5 is the median and mean observation. Their time series averages are recorded in Panel B.

The statistics indicate low correlation between the sensitivity measure and the uncertainty measure (0.062), B/M ratio (0.013) and size (0.016). The uncertainty measures correlation with size is also very low (-0.018). On the other hand, the uncertainty measure has a positive correlation with B/M (0.265). In other words, higher dispersion stocks tend to be high B/M or value stocks which is consistent with the findings in Doukas, Kim and Pantzalis (2004). The correlation between the uncertainty measure and B/M is confirmed in Panel B as the average ranking of B/M for the stocks in the low uncertainty portfolio (U1) is 0.40, while the average ranking for the medium (U2) and high (U3) uncertainty portfolios are 0.51 and 0.59 respectively. The pattern is also consistent in the double-sorted portfolios (e.g. S1U1 is the portfolio of the stocks belonging to both S1 and U1). Besides this relationship,

[25] As a robustness test, the mean analyst forecast is also used to normalize $\sigma_{i,t}$ instead of the stock price. The results from this alternative are nearly identical to those using equation (43). Consequently, for brevity, they are unreported but available upon request.

the sensitivity and uncertainty portfolios are unrelated to B/M, size and analyst coverage factors. The average rankings of the three variables (B/M, size and number of analysts) for the stocks in each sensitivity and uncertainty portfolio are all close to 0.5 (with the exception of B/M and the uncertainty portfolios). Therefore, the portfolios have similar B/M, size and analyst coverage characteristics, and are well represented by an average stock.

5.5 Earnings Momentum Strategies

When the first two hypotheses are combined, the result is the following prediction for the profitability of earnings momentum strategies. This third hypothesis states that these cross-sectional returns are largest when conditioning on a credible earnings forecast during periods of high return uncertainty.

Hypothesis 3. *Earnings momentum profits are largest for stocks with high (previous) uncertainty and sensitivity measures.*

Earnings momentum is implemented as in Jegadeesh and Titman (1993), but with forecast revisions over the past 6 months instead of stock returns. The forecast revision for firm i in month t is defined as

$$REV6_{i,t} = \sum_{j=0}^{5} rev_{i,t-j}\,, \tag{44}$$

where $rev_{i,t}$ is defined in equation (42). We rank the stocks according to equation (44) and assign them to one of five quintile portfolios each month. The bottom quintile portfolio contains stocks with the most unfavorable earnings forecast revision, while the top quintile contains those with the most favorable revision. Overlapping portfolios are then constructed to compute equally-weighted returns each month. For instance, the portfolio having the most favorable revision (E5) consists of six overlapping portfolios from the previous six ranking months. The return for this portfolio is the simple average return of the six portfolios formed over the past six months. If a stock's return is missing during the holding period, it is replaced with the corresponding value-weighted market return. The earnings momentum portfolio is the zero-investment portfolio that buys the most favorable revision portfolio and sells the least favorable revision portfolio, E5-E1, each month.

Our earnings momentum strategy differs slightly from the standard price momentum strategy in another respect. After ranking stocks according to their past returns, Jegadeesh and Titman (1993) skip one month before buying stocks to avoid bid-ask spread and short-term stock price reversal. This one month gap is not inserted into our strategies for two reasons. First, we rank stocks based on their earnings which, unlike past returns, is not subject to the bid-ask spread problem. Second, almost all earnings consensus estimates are available between the 10^{th} and the 20^{th} day of the month. Consequently, about half a month has already been omitted before we start holding positions at the beginning of next month.

5.6 Earnings Momentum Conditioned on Sensitivity and Uncertainty

Chan, Jegadeesh and Lakonishok (1996) document strong earnings momentum profits and suggest that earnings momentum is caused by the slow response of market participants to earnings information. If earnings momentum is caused by market under-reaction to earnings information, our theory would predict that earnings momentum is stronger for stocks whose earnings information is more credible, and those with more uncertain earnings. Thus, we hypothesize that earnings momentum strategies are more profitable for stocks in the high sensitivity and high uncertainty portfolios.

Table 2 reports earnings momentum profits and illustrates the importance of return sensitivity to earnings and earnings uncertainty. When the earnings momentum strategy is implemented using the full sample, the strategy generates an average return of 0.69% per month with a t-statistic of 4.38.

Next, we implement the strategy separately for the three sensitivity groups (S1, S2 and S3). The momentum profit remains significant in each of the three groups. More interestingly, the profit increases monotonically from the low sensitivity group (S1) to the high sensitivity group (S3), with the profit of the latter being about 50% higher than the former (0.79% vs. 0.52%). To clarify, the grouping of S1, S2 and S3 is determined before the stocks are assigned to the earnings momentum portfolios (E1 to E5), and thus before the buying or selling of stocks.

The momentum profit pattern is identical in the three uncertainty groups, increasing monotonically from U1 to U3, the profit of U3 being approximately 70% higher than U1 (0.74% vs. 0.44%). When the earnings momentum strategy is applied to double-sorted portfolios on sensitivity and uncertainty, the monotonic increasing pattern of the momentum profits continues. Within each sensitivity group, the profit increases monotonically from U1 to U3 (e.g. within the medium sensitivity group, the profit is 0.49%, 0.64% and 0.79% for S2U1, S2U2 and S2U3 respectively). In addition, within each uncertainty group, the profit increases monotonically from S1 to S3 (e.g. within the medium uncertainty group, the profit is 0.45%, 0.64% and 0.73% for S1U2, S2U2 and S3U2 respectively).

There is existing evidence that momentum profits are affected by factors such as the B/M ratio, documented in Daniel and Titman (1999), along with size and analyst coverage, as reported in Hong, Lim and Stein (2000). Our descriptive statistics in Table 1 indicate that our sensitivity and uncertainty results are not manifestations of these factors.

In particular, our uncertainty measure is positively correlated with B/M, implying low uncertainty stocks tend to be growth stocks. Daniel and Titman (1999) find stronger momentum among growth stocks, and attribute this finding to investor overconfidence. If uncertainty is irrelevant, the positive correlation between uncertainty and B/M would indicate higher momentum profit amongst low rather than high uncertainty stocks. Therefore, our ability to find increasing momentum profits from U1 to U3 attests to the importance of conditioning on uncertainty.

The sensitivity and uncertainty measures are also weakly positively correlated with analyst coverage, although this feature is not found in Panel B of Table 1. Hong, Lim and Stein (2000) report higher momentum profits for stocks with less analyst coverage, consistent with the slow diffusion of information. Their findings also predict less momentum profits for the high sensitivity and high uncertainty stocks, while we find increasing momentum profits from S1 to S3 and U1 to U3. Consequently, the sensitivity and uncertainty measures both contain important conditional information that is not captured by the existing literature.

Overall, we can reasonably conclude that our earnings momentum results, which are derived from sensitivity and uncertainty measures for the return implications of earnings and variability in earnings respectively, are not driven by book-to-market, size and analyst coverage effects documented in the existing literature.

6 Conclusions

We introduce an information portfolio which minimizes the aggregate uncertainty of multiple return forecasts. This portfolio summarizes investor beliefs, and yields an expected return which exhibits biases that are similar to overconfidence, biased self-attribution, representativeness and conservatism as well as limited attention. Therefore, despite being optimal, expected returns display biases that have previously been attributed to investor psychology.

Furthermore, biases and return predictability are strongest in periods of high return uncertainty. As a consequence, information quality can (at least partially) explain the biases in expected returns. Moreover, testable implications of information portfolio theory distinct from psychology are available. In contrast to Bayesian frameworks, these implications are independent of any assumed prior distribution.

By examining the profits of earnings momentum strategies, we document the importance of return sensitivity to earnings as well as earnings uncertainty. The two pillars of information theory are verified since momentum profits increase monotonically from low to high sensitivity stocks, and from low to high uncertainty stocks. More importantly, the sensitivity results continue after controlling for the effects of information uncertainty. Thus, investors condition their beliefs in accordance with information portfolio theory since more accurate return forecasts are assigned greater influence on the expected return. The significance of our sensitivity and uncertainty measures is not driven by factors such as book-to-market, size and analyst coverage.

References

[1] Baker, M., and J. Wurgler, 2005. Investor sentiment and the cross-section of stock returns. Forthcoming in Journal of Finance.

[2] Barber, B. M., T. Odean, and N. Zhu, 2003. Systematic noise. Working Paper, University of California, Davis.

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

[4] Berger, J. O., 1985. Statistical Decision Theory and Bayesian Analysis (2nd Edition). Springer-Verlag, New York.

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

[6] Bernardo, A. E., and O. Ledoit, 2000. Gain, loss, and asset pricing. Journal of the Political Economy 108, 144-172.

[7] Bondarenko, O., and P. Bossaerts, 2000. Expectations and learning in Iowa. Journal of Banking and Finance 24, 1535-1555.

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

[9] Carlson, M., A. Fisher, and R. Giammarino, 2004. Corporate investment and asset pricing dynamics: Implications for the cross-section of returns. Journal of Finance 59, 2577-2603.

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

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

[12] Daniel, K., and S. Titman, 2005. Market reactions to tangible and intangible information. Forthcoming in Journal of Finance.

[13] Daniel, K., and S. Titman, 1999. Market efficiency in an irrational world. Financial Analysts Journal 55, 28-40.

[14] Doukas, J. A., C. Kim, and C. Pantzalis, 2004. Divergent opinions and the performance of value stocks. Financial Analysts Journal 60, 55-64.

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

[16] Gomes, J., L. Kogan, and L. Zhang, 2003. Equilibrium cross-section of returns. Journal of Political Economy 111, 693-732.

[17] 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.

[18] Hong, H., J. Scheinkman, and W. Xiong, 2005. Asset float and speculative bubbles. Forthcoming in Journal of Finance.

[19] Jackson, A., and T. Johnson, 2006. Unifying underreaction anomalies. Journal of Business 79, forthcoming.

[20] Jagannathan, R., and Y. Wang, 2005. Lazy Investors, Discretionary Consumption, and the Cross Section of Stock Returns. Working Paper, Northwestern University.

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

[22] Kumar, A., 2005. When do investors exhibit stronger behavioral biases? Working Paper, Notre Dame University.

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

[24] Peng, L., and W. Xiong, 2004. Limited attention and asset prices. Forthcoming in Journal of Financial Economics.

[25] Ross, S., 1976. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341-360.

[26] Vassalou, M., and K. Apedjinou, 2004. Corporate innovation, price momentum, and equity returns. Working Paper, Columbia University.

[27] Zhang, X. F., 2005. Information uncertainty and stock returns. Forthcoming in Journal of Finance.

Appendices

A Proof of Proposition 1

Denote the Lagrangian of equation (7) as

$$L(W, \lambda) \;=\; \frac{1}{2} W^T \Theta W + \lambda (W^T \mathbf{1} - 1)\,, \tag{45}$$

which generates two equations

$$\frac{\partial L(W,\lambda)}{\partial W} = \Theta W + \lambda \mathbf{1} = 0 \tag{46}$$

$$\frac{\partial L(W,\lambda)}{\partial \lambda} = W^T\mathbf{1} - 1 = 0 \tag{47}$$

involving two unknowns; W and the Lagrangian multiplier λ. Equation (46) is equivalent to

$$W = -\lambda\Theta^{-1}\mathbf{1}\,. \tag{48}$$

Multiplying the transpose of equation (48) by the $\mathbf{1}$ vector yields

$$W^T\mathbf{1} = -\lambda\mathbf{1}^T\Theta^{-1}\mathbf{1} \tag{49}$$

which implies

$$1 = -\lambda\mathbf{1}^T\Theta^{-1}\mathbf{1}\,, \tag{50}$$

due to the $W^T\mathbf{1} = 1$ constraint. Therefore, the λ parameter is solved as

$$-\lambda = \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\,. \tag{51}$$

Substituting equation (51) into equation (48) produces the final result

$$W = \left(\frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\right)\Theta^{-1}\mathbf{1}\,, \tag{52}$$

which satisfies the constraint

$$W^T\mathbf{1} = \left(\frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\right)\mathbf{1}^T\Theta^{-1}\mathbf{1} = 1\,. \tag{53}$$

B Covariances and the Perceived Return

The partial derivative of the investor's perceived return in equation (21) with respect to σ_{12} equals

$$\begin{aligned}\frac{\partial \text{Perceived Return}}{\partial \sigma_{12}} &= \frac{-(\mu_1+\mu_2)\left[\sigma_1^2+\sigma_2^2-2\sigma_{12}\right]+2\left[\sigma_2^2\,\mu_1+\sigma_1^2\,\mu_2-\sigma_{12}\,(\mu_1+\mu_2)\right]}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2} \\ &= \frac{(\mu_2-\mu_1)\,\left(\sigma_1^2-\sigma_2^2\right)}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}\,,\end{aligned} \tag{54}$$

according to the quotient rule of calculus. In general, the sign of this derivative may be either positive or negative. According to the numerator of equation (54), when either the return forecasts or their historical accuracies are identical, the investor's perceived return is invariant to the covariance. Figure 1 finds the perceived return is insensitive to σ_{12} over a range of values.

The partial derivative of the perceived return's aggregate uncertainty in equation (22) with respect to σ_{12} equals

$$\begin{aligned}\frac{\partial \text{Aggregate Uncertainty of Perceived Return}}{\partial \sigma_{12}} &= \frac{-2\sigma_{12}\left[\sigma_1^2+\sigma_2^2-2\sigma_{12}\right]+2\left[\sigma_1^2\sigma_2^2-(\sigma_{12})^2\right]}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2} \\ &= \frac{2\sigma_{12}\left[\sigma_{12}-(\sigma_1^2+\sigma_2^2)\right]+2\sigma_1^2\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}\,. \qquad (55)\end{aligned}$$

As confirmed by Figure 1, the perceived return's uncertainty is sensitive to the sign of the sample covariance. In particular, when σ_{12} is negative, equation (55) is large and positive. Thus, return uncertainty decreases as the covariance becomes more negative.

C Proof of Proposition 3

Recall from equation (16) that the perceived return is distributed $\mathcal{N}\left(W^T\mu, \nu + W^T\Theta W\right)$ under the distributional assumption of equation (9). To prove Proposition 3, the following utility maximization problem is solved

$$\begin{aligned}&\max_f E\left\{U\left[M\,\left((1-f)\,(1+r_f)+f\,\left(1+W^T\mu\right)\right)\right]\right\} \\ =\;& \max_f -E\left[\exp\left\{-a\,M\left(1+r_f+f\,\left(W^T\mu-r_f\right)\right)\right\}\right] \\ =\;& \max_f -\exp\left\{-a\,M\,f\,W^T\mu+a\,M\,f\,r_f+\frac{a^2\,M^2\,f^2}{2}\,\left[\nu+W^T\Theta W\right]\right\}, \qquad (56)\end{aligned}$$

where the last equality results from the moment generating function of a normal distribution. The necessary maximization involves setting the following partial derivative of equation (56) with respect to f

$$\left(-a\,M\,W^T\mu+a\,M\,r_f+a^2\,M^2\,f\,\left[\nu+W^T\Theta W\right]\right)\left(-e^{\left\{-a\,M\,f\,W^T\mu+a\,M\,f\,r_f+\frac{a^2\,M^2\,f^2}{2}\,\left[\nu+W^T\Theta W\right]\right\}}\right) \qquad (57)$$

to zero. This requires the first term in the above product to be zero

$$-a\,M\,W^T\mu+a\,M\,r_f+a^2\,M^2\,f\,\left[\nu+W^T\Theta W\right] \;=\; 0\,. \qquad (58)$$

Therefore, the optimal investment in the risky asset equals

$$f = \frac{W^T \mu - r_f}{a\, M\, \left[\nu + W^T \Theta W\right]}\,, \tag{59}$$

which becomes

$$\begin{aligned} f &= \frac{1}{a\, M} \frac{\left(\frac{\mathbf{1}^T \Theta^{-1} \mu}{\mathbf{1}^T \Theta^{-1} \mathbf{1}} - r_f\right)}{\nu + \frac{1}{\mathbf{1}^T \Theta^{-1} \mathbf{1}}} \\ &= \frac{\mathbf{1}^T \Theta^{-1} \left(\mu - r_f \mathbf{1}\right)}{a\, M\, \left[1 + \nu \mathbf{1}^T \Theta^{-1} \mathbf{1}\right]}\,, \end{aligned} \tag{60}$$

after substituting in the results of Proposition 2.

Observe that the optimal portfolio weights from Proposition 1 transform equation (59) into equation (60). When there is no uncertainty regarding the asset's true expected return, equation (59) implies $f = \frac{\eta - r_f}{a\, M\, \nu}$ since $W^T \mu = \eta$ and $W^T \Theta W = 0$.

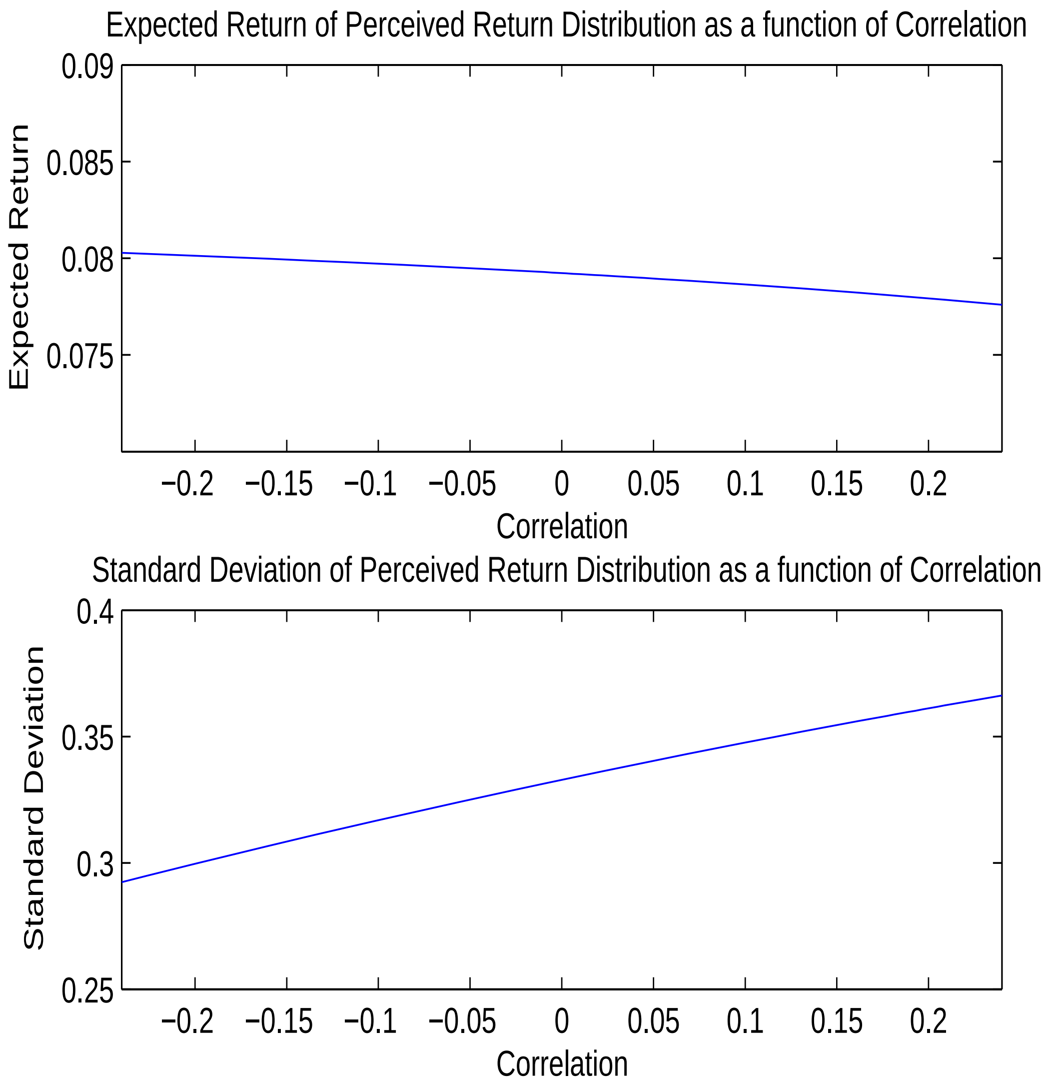

Figure 1: Impact of correlation between two return forecasts on the perceived return and its uncertainty according to equations (21) and (22) respectively. The above figure is derived from the following parameter values; $\mu_1 = 0.07$, $\mu_2 = 0.10$, $\sigma_1 = 0.40$ and $\sigma_2 = 0.60$. The σ_j parameter denotes the square root historical accuracy computed in equation (2) for the j^{th} information source, and serves as the uncertainty of μ_j for $j = 1, 2$.

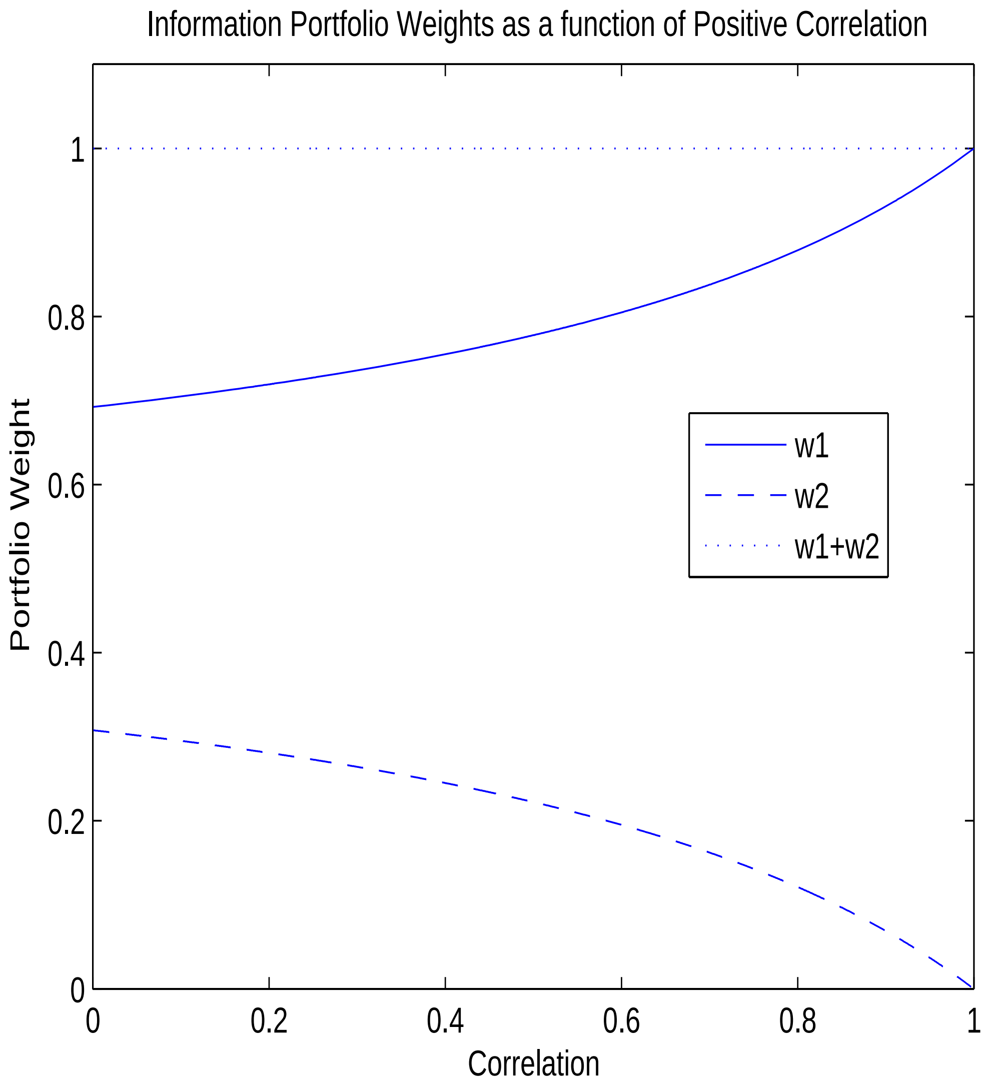

Figure 2: Impact of correlation between return forecasts on information portfolio weights. The above figure is derived from the following parameter values; $\mu_1 = 0.07$, $\mu_2 = 0.10$, $\sigma_1 = 0.40$ and $\sigma_2 = 0.60$ (as in Figure 1). Observe that higher positive correlation reduces the portfolio weight of the less accurate information source.

Table 1: Descriptive Statistics

This table describes our sensitivity and uncertainty measures as well as the characteristics of our dataset pertaining to B/M, size and number of analysts. The sensitivity measure is estimated monthly for each stock by computing the correlation coefficient between returns and price-scaled analyst forecast revisions over the previous 12 months. The uncertainty measure represents the price-scaled standard deviation of analyst forecasts for every stock each month. The sensitivity measure, uncertainty measure and number of analysts are derived from the I/B/E/S Summary History dataset, while B/M is the book-to-market ratio using the most recent quarterly data from Compustat. Size denotes the stock's market capitalization as reported in CRSP. The Spearman rank correlation coefficients among the five variables are computed each month from January 1976 to December 2004. Panel A reports the time series average of the Spearman correlation coefficients. Panel B reports growth/value, big/small and analyst coverage characteristics for the sensitivity and uncertainty portfolios. The sensitivity (uncertainty) portfolios denoted S1, S2 and S3 (U1, U2 and U3) represent the bottom 30%, middle 40% and top 30% of stocks ranked according to their sensitivity (uncertainty) measures. Double-sorted portfolios are also formed (e.g., S1U1 consists of stocks that belong to both S1 and U1). Each month, stocks are also ranked by B/M, size and number of analysts. This ranking is then normalized to the [0,1] interval. The average ranking for B/M, size and number of analysts in each sensitivity and uncertainty portfolio is computed monthly. The numbers in Panel B are the time series average for these monthly rankings in each sensitivity and uncertainty portfolio.

Panel A: Spearman Rank Correlation Coefficients

	Sensitivity	Uncertainty	B/M	Size	# of Analysts
Sensitivity	1	0.062	0.013	0.016	0.111
Uncertainty		1	**0.265**	-0.018	0.129
B/M			1	-0.275	-0.091
Size				1	0.833
# of Analyst					1

Panel B: Characteristics of Sensitivity and Uncertainty Portfolios

	B/M	Size	# of Analysts
S1	0.49	0.51	0.50
S2	0.50	0.50	0.50
S3	0.51	0.50	0.51
U1	**0.40**	0.50	0.48
U2	**0.51**	0.52	0.52
U3	**0.59**	0.48	0.50
S1U1	**0.39**	0.51	0.49
S1U2	**0.51**	0.52	0.51
S1U3	**0.58**	0.49	0.50
S2U1	**0.40**	0.50	0.48
S2U2	**0.51**	0.51	0.51
S2U3	**0.59**	0.48	0.49
S3U1	**0.40**	0.49	0.47
S3U2	**0.50**	0.51	0.52
S3U3	**0.60**	0.48	0.51

Table 2: Earnings Momentum Strategies

This table describes the profitability of earnings momentum strategies applied to stocks with varying levels of earnings uncertainty and return sensitivity to earnings. At the end of each month from July 1977 to December 2004, stocks from the intersection of the CRSP and I/B/E/S datasets are ranked on the basis of changes in consensus analyst earnings forecasts, measured by cumulative price-deflated revisions in the past six months. Stocks are assigned to five quintile portfolios, and equally weighted returns are computed for each portfolio. The bottom 20% is assigned to the E1 portfolio and the top 20% denotes the E5 portfolio. The trading strategy 6-0-6 in Jegadeesh and Titman (1993) is then implemented. Each month, the portfolio containing the most favorable (unfavorable) past revisions is an overlapping portfolio consisting of the E5 (E1) portfolios during the previous six months. Returns for the favorable (unfavorable) overlapping portfolios are the average returns over the six E5 (E1) portfolios. If a stock's return is missing during the holding period, it is replaced with the corresponding value-weighted market return. The earnings momentum portfolio (E5-E1) is the zero-cost portfolio that buys the most favorable revision portfolio and sells the least favorable revision portfolio (E5-E1) every month. Panel A reports the results for the strategy using the full sample. Panel B reports the results for stocks sorted on their sensitivity to analyst forecast revisions (S1, S2 and S3). Stocks are assigned to these groups before the earnings momentum portfolios are formed. Panel C reports the results when stocks are grouped according to their price-scaled standard deviation of analyst forecasts (U1, U2 and U3). These uncertainty groups are also constructed prior to the formation of the earnings momentum portfolios. Panel D reports our results after double-sorting by the sensitivity and uncertainty measures (e.g. S1U1 represents the group of stocks belonging to S1 and U1).

Panel A: Strategy using full sample

	E1	E2	E3	E4	E5	**E5-E1**	***t-stat***
All	1.09	1.22	1.25	1.48	1.79	**0.69**	***4.38***

Panel B: Strategy conditional on sensitivity of stock price to earnings information

	E1	E2	E3	E4	E5	**E5-E1**	***t-stat***
S1	1.24	1.28	1.23	1.38	1.76	**0.52**	***3.36***
S2	1.12	1.24	1.25	1.46	1.81	**0.69**	***3.76***
S3	1.11	1.26	1.34	1.54	1.90	**0.79**	***4.20***

Panel C: Strategy conditional on uncertainty of earnings information

	E1	E2	E3	E4	E5	**E5-E1**	***t-stat***
U1	1.21	1.14	1.17	1.40	1.65	**0.44**	***2.33***
U2	0.96	1.10	1.16	1.32	1.59	**0.64**	***4.34***
U3	0.81	1.05	1.25	1.29	1.55	**0.74**	***5.04***

Panel D: Strategy conditional on both sensitivity and uncertainty

	E1	E2	E3	E4	E5	**E5-E1**	***t-stat***
S1U1	1.61	1.38	1.25	1.37	1.72	**0.11**	***0.40***
S1U2	1.18	1.22	1.23	1.27	1.62	**0.45**	***2.50***
S1U3	0.95	1.30	1.28	1.46	1.66	**0.71**	***4.42***
S2U1	1.43	1.31	1.25	1.42	1.93	**0.49**	***2.24***
S2U2	1.08	1.25	1.28	1.45	1.72	**0.64**	***3.66***
S2U3	0.87	1.20	1.35	1.46	1.66	**0.79**	***4.53***
S3U1	1.41	1.35	1.27	1.62	1.92	**0.52**	***2.28***
S3U2	1.15	1.25	1.44	1.50	1.88	**0.73**	***4.53***
S3U3	0.82	1.11	1.40	1.44	1.68	**0.86**	***3.94***