Return Uncertainty and the Appearance of Biases in Expected Returns*

Dong Hong[†] and Mitch Warachka[‡]

April 2006

Abstract

We study the relationship between return uncertainty and behavioral finance by introducing an information portfolio which combines multiple return forecasts for a single asset into an estimate of its unknown expected return. Our optimal information portfolio minimizes the aggregate forecast error of an asset's estimated expected return when combining the return forecasts. The expected return from this minimization exhibits momentum as well as the *appearance* of overconfidence, biased self-attribution, representativeness, conservatism and limited attention. Although these characteristics coincide with expected return uncertainty, they are induced by the optimal information portfolio weights assigned to return forecasts rather than behavioral biases. Empirically, our optimal information portfolio yields testable implications distinct from psychology which we verify using analyst earnings forecasts and their revisions.

JEL Classification: G12, G14

Keywords: Behavioral Biases, Momentum, Information Portfolio

*We thank Bhagwan Chowdhry, Bruce Grundy, Steve Hogan, David Hirshleifer, Ravi Jagannathan, Avanidhar Subrahmanyam, Wei Mun Wang, Guojun Wu and Jun Yu for their helpful comments and suggestions.

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

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

1 Introduction

When testing market efficiency using historical return data, the empirical asset pricing literature usually ignores the range of return forecasts for an asset that were available to investors. For example, multifactor asset pricing models estimate a *single* expected return after conditioning on realized factor returns. However, a consensus regarding the correct multifactor formulation continues to be elusive since the number of required factors and their composition are controversial. More importantly, even if market participants agreed on a multifactor model, factor returns and the corresponding factor loadings for an individual asset are unknown ex-ante. Consequently, these unknown inputs are sources of expected return uncertainty. Furthermore, price targets and intrinsic value measures, such as the residual income valuation model in Lee, Myers and Swaminathan (1999), can exacerbate expected return uncertainty by providing alternative return forecasts.

Motivated by this uncertainty, we examine *multiple* return forecasts for an *individual* asset whose expected return is unknown. Return forecasts are issued by information sources after interpreting state variables such as the firm's projected earnings or the prospects for its industry. However, the true accuracy of each information source is unknown. Instead, the accuracy of an information source is estimated according to its time series of prior forecast errors. This estimated accuracy depends on the return implications and dynamics of the information source's underlying state variable. Covariances between the forecast errors of different information sources are also estimated.

The information portfolio combines the return forecasts for an individual asset into an estimated 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 examines multiple return forecasts for a single asset whose expected return is unknown.[1] The optimal information portfolio minimizes the aggregate forecast error of an asset's expected return estimate by assigning higher portfolio weights to more accurate information sources.[2] The estimated

[1] To illustrate our notion of expected return uncertainty, a BusinessWeek survey reported annual return forecasts for the S&P 500 ranging between -29.5% and 31.0% with a standard deviation of 7.61%. This uncertainty is magnified for individual stocks whose expected returns are determined, at least in part, by the market's expected return. Furthermore, the average forecast of 7.87% is not the optimal estimate for the S&P 500's expected return unless the 76 forecasts are equally accurate.

[2] 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, we refrain from

expected return of an asset implied by the optimal information portfolio is labeled the investor's *perceived* return.

With regards to behavioral finance, the perceived return exhibits the appearance of overconfidence and biased self-attribution as well as representativeness and conservatism. These two pairs of psychological biases have previously been incorporated into the behavioral finance literature by Daniel, Hirshleifer and Subrahmanyam (1998) and Barberis, Shleifer and Vishny (1998) respectively. For example, the optimal information portfolio emphasizes accurate sources of private information, while downplaying the investor's less accurate private return forecasts. Furthermore, state variables with trends in their dynamics are assigned larger information portfolio weights whenever this predictability improves forecast accuracy. A property which mimics limited attention is also instilled into the perceived return since return forecasts receive smaller information portfolio weights if they are positively correlated with forecasts from more accurate information sources. All of these perceived return characteristics are induced by our optimal information portfolio weights rather than psychology. Momentum and subsequent reversals in the perceived return also result from the dynamic updating of the information portfolio weights. These optimal portfolio fluctuations reflect changes in the estimated accuracy of each information source as additional forecast errors become available.

Momentum as well as the appearance of psychological biases are most salient during periods of high expected return uncertainty when fewer forecast errors are available. Intuitively, events that alter a firm's capital structure or investment strategy as well as technological innovations reduce the number of relevant previous forecast errors and increase expected return uncertainty. Therefore, information portfolio theory does not assume the return implications of such idiosyncratic events are immediately understood and agreed upon by all information sources. Indeed, every return forecast would be without error and identical under this extreme assumption. Thus, our framework examines the *limits to available information* when determining an investor's perceived return. Although compatible with multifactor asset pricing models, information portfolio theory allows information sources to disagree on ex-ante factor returns as well as the factor loadings of an individual asset.[3] Consequently, the market's expected return is not assumed to be known, nor are the firm's future earnings when implementing

referring to the information portfolio weights as linear regression coefficients since their optimality is independent of any distributional assumption and does not require unbiased return forecasts.

[3] In the context of information portfolio theory, ex-ante factor returns serve as state variables while factor loadings represent their return implications.

intrinsic value measures.

Unlike Bayesian models in behavioral finance which incorporate psychological biases by imposing assumptions on the investor's prior distribution, we examine the optimal aggregation of multiple return forecasts. Therefore, characteristics of the perceived return which mimic behavioral biases are outputs from information portfolio theory rather than inputs. This important distinction yields unique testable implications of information portfolio theory. In contrast, Brav and Heaton (2002) demonstrate the difficulty of distinguishing between behavioral and rational explanations for return anomalies using Bayesian techniques.[4]

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) report that behavioral biases appear stronger during periods of high uncertainty. Besides event and time dependence, Baker and Wurgler (2005) report that firm characteristics such as size and age explain a firm's sensitivity to investor sentiment, while Vassalou and Apedjinou (2004) report that momentum strategies are most profitable for firms with high levels of corporate innovation. These empirical regularities appear to be consistent with information portfolio theory as well as psychological biases. However, a relative ranking of the information sources by their estimated accuracies is equivalent to the existence of an information portfolio. Thus, for a given level of uncertainty, information portfolio theory posits that accurate sources of information have the greatest influence on the investor's perceived return. In contrast, this optimal weighting is not predicted by psychology.

As a consequence, after controlling for state variable uncertainty, information portfolio theory asserts that investors focus their attention on state variables which have experienced the highest correlation with realized returns. Empirically, we verify the main testable implications of information portfolio theory by examining earnings momentum. For a given level of earnings uncertainty, psychology predicts stronger momentum when earnings are less informative, while information portfolio theory predicts the opposite. Therefore, the first aspect of our empirical study measures the *sensitivity* of returns to earnings revisions by computing firm-specific correlations between these variables. We find momentum profits increase monotonically from low to high sensitivity stocks by 50%. This evidence is consistent with investors focusing on earnings when this state variable has been informative. The sec-

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

ond 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. Most importantly, portfolios derived from double sorts on the sensitivity and uncertainty measures continue to display both relationships. Consequently, after controlling for earnings uncertainty, firms whose earnings are more informative experience greater momentum. This finding is consistent with the central prediction of information portfolio theory. Several robustness checks verify that our results are not driven by book-to-market, size and analyst coverage.

However, if knowledge of investor psychology improves forecast accuracy, then information portfolio theory and psychology are compatible. Therefore, the exact decomposition of the perceived return into the effects of psychology versus information portfolio theory is ultimately an empirical question. Nonetheless, our empirical implementation demonstrates that the contribution of information portfolio theory is crucial.

Information portfolio theory also enhances applications of utility maximization by providing a general formulation to estimate an investor's perceived return and its aggregate forecast error. For example, an investor with exponential utility reduces their exposure to a risky asset when the aggregate forecast error is high.

The remainder of this paper begins with the introduction of the optimal information portfolio in Section 2. Section 3 illustrates the impact having a limited number of forecast errors when estimating the accuracy of an information source, and examines the ability of time-varying optimal information portfolio weights to induce momentum (and reversals) in the perceived return. Section 4 links the optimal information portfolio with return characteristics which have previously been 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 industry and macroeconomic conditions. Each state variable forecast is interpreted by an information source who expresses its estimated return implications for a particular asset.[5] In practice, an individual analyst can issue earnings forecasts and long term growth rate projections 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 sources of information can originate from an individual analyst, the firm or the investor.

To simplify the exposition of our framework, but without loss of generality, each return forecast is generated by a single state variable.[6] From an academic perspective, this structure enables our framework to address issues related to which sources of information influence expected returns. For example, Brav and Lehavy (2003) examine the marginal importance of analyst price targets to the price formation process in the presence of earnings forecast revisions and stock recommendations. Furthermore, this structure allows the information portfolio to aggregate over the widest possible array of return forecasts. Although the economic intuition underlying our framework is identical if information sources interpret multiple state variables before issuing their return forecasts, this modification reduces the amount of aggregation performed by the information portfolio.

The return forecast issued by an information source can possess private as well as public characteristics. For example, state variables such as earnings forecasts, while publically available when issued by sell-side analysts, require additional interpretation by the investor to become return forecasts. Conversely, the conversion of analyst price targets into return forecasts is immediate, implying these sources of information are entirely public. The prior returns of an asset also constitute a source of public information whose corresponding return forecasts are similar to buy/sell signals in technical analysis. For emphasis, information sources are only assumed to issue return forecasts. The mechanism for estimating their accuracy is addressed in the next subsection.

In summary, we consider $J > 1$ return forecasts for a single asset originating from J unique

[5] Although sales are usually reported in millions of dollars and earnings stated on a per share basis, information portfolio theory abstracts from these scale complications by aggregating across their return implications.

[6] The next section demonstrates that information portfolio theory is able to replicate the expected return estimates from multifactor asset pricing models.

information sources who evaluate the return implications of $K \geq 1$ state variable forecasts.[7] The inequality $J \geq K$ enables information sources to disagree on the return implications of a forecasted state variable.

2.1 Estimating Information Source Accuracy

The accuracy of each information source is critical to the information portfolio's solution. This property is estimated 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

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

over the previous n periods. 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 accuracy of the j^{th} information source is estimated as

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

according to their previous forecast errors over the last n periods. By replacing the t subscript in $\sigma_{j,t}^2$ with an asterix, we denote the true but unknown accuracy of the j^{th} information source as $\sigma_{j,*}^2$ which proxies for their true skill at forecasting the asset's expected return.[8] Therefore, $\sigma_{j,t}^2$ in equation (2) is the investor's estimate of $\sigma_{j,*}^2$ based on the information source's prior n forecast errors. Intuitively, this estimate represents the *credibility* of the $\mu_{j,t}$ forecast issued at $t-1$. From a statistical perspective, equation (2) calculates the mean-squared error (MSE) of equation (1).[9]

[7]When state variable dynamics are random, K represents the number of state variable forecasts rather than the number of actual state variables. This structure allows each state variable forecast to generate a distinct return forecast.

[8]The true accuracy $\sigma_{j,*}^2$ of the j^{th} information source may be time-varying and state-dependent. However, this parameter is written as a constant for notational simplicity since only the estimated accuracies from equation (2) are involved in the solution for the optimal information portfolio.

[9]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 $E[\epsilon]$. Information sources may employ Bayesian methods when generating their return forecasts with the usual tradeoff between variance and

Throughout the remainder of our paper, the *accuracy* of an information source refers to its estimate in equation (2). However, with the true accuracy of each information source being unknown, *accuracy* and *estimated accuracy* are used interchangeably. Observe that the estimated accuracy of an information source is derived from its time series of prior return forecasts, with state variable forecasts serving an intermediate role.

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

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

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, although the *estimates* in equations (2) and (3) should be denoted as $\hat{\sigma}^2_{j,t}$ and $\hat{\sigma}_{j,k,t}$ respectively, the hats are omitted for notational simplicity.

In our current exposition, the value of n in equations (1), (2) and (3) is specific to an individual asset.[10] Intuitively, established firms in stable industries have ample forecast errors to estimate the accuracy of each information source. Conversely, initial public offerings and companies undergoing a significant restructuring, undertaking a large investment, or experiencing major technological innovations have fewer *relevant* forecast errors. This notion of relevance has n being reduced after significant corporate events, and parallels Brav and Heaton (2002)'s concept of a random *change point* in the economy. However, the arrival of a change point is not necessarily unknown in our framework. Instead, the change point's impact on expected returns is uncertain.[11] Thus, we examine whether expected

bias arising from an informative prior. By computing the mean-squared error of prior forecast errors, the potential for optimism to bias analyst earnings forecasts, price targets and stock recommendations is addressed. Nonetheless, the investor can adjust the return forecasts to account for known biases, although decomposing forecast errors into their bias and variance components is not required.

[10]If n is specific to an individual information source, then n_j would denote the number of relevant forecast errors for the j^{th} information source. For example, n_j could proxy for the experience of an information source. Chen, Liu and Qian (2005) document the importance of experience to the credibility of buy-side analyst forecasts, while Nicolosi, Peng and Zhu (2004) report that experienced individual investors earn higher returns. However, for ease of exposition, all J information sources are evaluated using n previous forecast errors since our initial focus is on a firm-specific information environment.

[11]From an empirical perspective, a reduction in n after corporate events is motivated by the estimation of time-

return uncertainty is responsible for causing momentum and the appearance of return characteristics that mimic behavioral biases. Empirical evidence linking corporate events and return predictability is documented by Jackson and Johnson (2006) as well as Daniel and Titman (2005). Vassalou and Apedjinou (2004) also report that corporate innovation increases return predictability.

Overall, the vector of return forecasts available 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 $\mu_{t-1}, \ldots, \mu_{t-n}$ vectors over the last n periods yields a Θ_t matrix summarizing the estimated accuracy for the J information sources as well as their estimated covariances which are described by equations (2) and (3) respectively. Therefore, the Θ_t matrix is an estimate of the true but unknown variance-covariance matrix for the J return forecasts in equation (4). To simplify our notation, we suppress the t subscripts on μ and Θ for the remainder of this paper.

2.2 Optimal Information Portfolio

When combining the J return forecasts, the investor minimizes the aggregate forecast error of the asset's estimated expected return by solving the following optimization problem

$$\begin{aligned} \min_{W} \quad & \frac{1}{2} W^T \Theta W \\ \text{subject to:} \quad & W^T \mathbf{1} = 1 , \end{aligned} \tag{5}$$

where $\mathbf{1}$ denotes a J-dimensional vector of ones. The solution for W is referred to as the *optimal* information portfolio. The objective function in equation (5) is related to Peng and Xiong (2004)'s minimization for the variance of beliefs regarding subsequent dividends. Hong, Scheinkman and Xiong (2005) also minimize the variance of different information sources.

varying parameters in the empirical asset pricing literature. For example, a company's market beta is typically not estimated using all returns since the firm's inception. Instead, to account for changes in the company and its operating environment, firm-specific factor sensitivities are calibrated over different subperiods.

As proven in the next subsection, after imposing a common distributional assumption on every return forecast, the objective function in equation (5) is equivalent to finding the best linear unbiased estimator (BLUE) of the asset's expected return given available forecasts. Therefore, equation (5) 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 in Appendix A.

Proposition 1. *The solution for the optimal information portfolio W in equation (5) equals*

$$W = \frac{\Theta^{-1}\mathbf{1}}{\mathbf{1}^T\Theta^{-1}\mathbf{1}} . \tag{6}$$

When private information sources are evaluated by the investor, this optimal information portfolio is investor-specific in addition to being firm-specific.

An information portfolio which equally-weights each return forecast, until there is statistically significant evidence that $\sigma^2_{j,*}$ differs across the J information sources, and the optimal information portfolio in equation (6) converge as the number of forecast errors in equation (1) increases. The equally-weighted portfolio has the investor utilizing $W = \frac{1}{J}\mathbf{1}$ unless the $\sigma^2_{j,t}$ estimates from equation (2) enable them to reject the null hypothesis that the true accuracy of each information source is identical. Besides earlier research by Peng and Xiong (2004) as well as Hong, Scheinkman and Xiong (2005), support for the objective function in equation (5) is provided below in equation (15) for *all* values of n.

2.3 Regression Interpretation of Optimal Information Portfolio

Denote the asset's true return distribution as $\mathcal{N}(\eta, \nu)$ with η being its unknown expected return. Corporate or macroeconomic events that generate expected return uncertainty may also cause η to vary over time but this parameter is written as a constant for notational simplicity.

The main result of this subsection is that after imposing a common distributional assumption on every return forecast

$$\mu \overset{d}{\sim} \mathcal{N}(\eta\mathbf{1}, \Theta) , \tag{7}$$

the objective function in equation (5) is equivalent to finding the best linear unbiased estimate of η. Specifically, from a linear regression perspective, the true model for the asset's return is described by

$$y = \eta + e , \tag{8}$$

where the error terms e are i.i.d. random variables from a $\mathcal{N}(0,\nu)$ distribution. Therefore, the asset's realized return y is emitted by the true $\mathcal{N}(\eta,\nu)$ distribution. Appendix B considers a special case of equation (8) which has η generated by a N-factor model

$$y \;=\; \left[\beta_0 + \sum_{j=1}^{N} \beta_j\, f_j\right] + e\,. \tag{9}$$

However, regardless of η's specification, its corresponding linear estimator $\hat{y}$ equals

$$\hat{y} \;=\; W^T \mu\,. \tag{10}$$

A linear regression procedure minimizes the mean-squared error of the $y - \hat{y}$ deviations

$$y - \hat{y} \;=\; \eta - W^T \mu + e\,, \tag{11}$$

by choosing the optimal *coefficients* W given a set of *independent variables* which are the return forecasts μ in our framework. The coefficients are required to produce an unbiased estimator which implies

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

The $W^T\mathbf{1} = 1$ constraint is an immediate consequence of equation (12) which follows from the distributional assumption in equation (7).[12] With $W^T\mu$ being an unbiased estimate of η, minimizing the mean-squared error in equation (11) is equivalent to minimizing

$$\begin{aligned} Var\left[y - \hat{y}\right] &= Var\left[\eta - W^T \mu + e\right] \\ &= Var\left[W^T \mathcal{N}\left(\eta \mathbf{1}, \Theta\right) + e\right] \\ &= W^T \Theta W + \nu\,, \end{aligned} \tag{13}$$

since the $\mathcal{N}(0,\nu)$ distribution for the error terms is independent of the normal distribution in equation (7) while η is not random. Equation (13) implies the investor minimizes $W^T\Theta W$ since the asset's true

[12] Equation (7) implies $W^T\mu$ is an unbiased estimator of η but this property does not imply that $W^T\mu$ equals η. Indeed, confidence intervals and hypothesis tests evaluate the point estimates from linear regression models.

return variance ν is not a function of the information portfolio. In summary, the best linear unbiased estimate of the asset's expected return minimizes $W^T\Theta W$ subject to the $W^T\mathbf{1} = 1$ constraint.[13] Consequently, statistical justification underlying linear regression models also applies to our objective function in equation (5). To clarify, $W^T\Theta W$ is not an estimate of ν. Indeed, even if $W^T\Theta W$ equals zero or η is known (as in classical portfolio theory), the asset is not riskless provided ν is non-zero.

To determine the asset's ex-ante return distribution, consider the *prediction* interval for next period's return

$$\tilde{y}_p = W^T\mu + e \,, \tag{14}$$

which is conditioned on W and a vector of return forecasts. For emphasis, equation (14) is not intended to calibrate the W coefficients since $\tilde{y}_p$ is the asset's unobserved (random) ex-ante return as signified by the tilde. Instead, conditional on W, the asset's ex-ante return distribution equals

$$\tilde{y}_p \overset{d}{\sim} \mathcal{N}\left(W^T\mu, W^T\Theta W + \nu\right) \,, \tag{15}$$

according to equation (13). Thus, return uncertainty reflects the asset's true variability denoted ν as well as the aggregate forecast error $W^T\Theta W$ of the J forecasts. Consequently, equation (15) provides further justification for the objective function in equation (5). Observe that equation (15) is valid under a weaker assumption than equation (7) which simply requires

$$W^T\mu \overset{d}{\sim} \mathcal{N}\left(\eta, W^T\Theta W\right) \,. \tag{16}$$

Equation (7) immediately implies equation (16) but the converse is not true. Intuitively, equation (7) assumes all J return forecasts are unbiased estimates of η, while equation (16) only assumes the aggregate forecast $W^T\mu$ is an unbiased estimate. From an practical perspective, equation (16) does not require individual information sources to issue unbiased forecasts. Instead, the investor's information portfolio is assumed to combine the return forecasts into an unbiased estimate of the asset's true expected return. The distribution of $\tilde{y}_p$ in equation (15) is invoked when the relationship between our optimal information portfolio and momentum is examined.

For emphasis, the objective function in equation (5) is independent of the distributional assumptions in equation (7) as well as equation (16). In particular, the optimal information portfolio in

[13] Minimizing $W^T\Theta W$ in equation (13) is equivalent to minimizing $\frac{1}{2}W^T\Theta W$ in equation (5).

Proposition 1 does not require the J return forecasts or their optimal combination to be unbiased estimates for η since equations (2) and (3) evaluate the mean-squared forecast error of an information source. More importantly, fluctuations in the optimal information portfolio weights over time are crucial to our interpretation of the investor's perceived return in the next two sections. Therefore, we refrain from referring to our information portfolio weights as linear regression coefficients.

2.4 Perceived Return and Return Uncertainty

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

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

$$W^T \mu = \frac{\mathbf{1}^T \Theta^{-1} \mu}{\mathbf{1}^T \Theta^{-1} \mathbf{1}}, \tag{17}$$

while

$$W^T \Theta W = \frac{1}{\mathbf{1}^T \Theta^{-1} \mathbf{1}}, \tag{18}$$

is the aggregate forecast error of the perceived return in equation (17).

Proof: The perceived return follows immediately from equation (6) while the aggregate forecast error is computed as[14]

$$\begin{aligned} W^T \Theta W &= \frac{1}{\mathbf{1}^T \Theta^{-1} \mathbf{1}} \mathbf{1}^T \Theta^{-1} \, \Theta \, \Theta^{-1} \mathbf{1} \frac{1}{\mathbf{1}^T \Theta^{-1} \mathbf{1}} \\ &= \frac{1}{\mathbf{1}^T \Theta^{-1} \mathbf{1}} . \end{aligned} \tag{19}$$

Ex-ante, the investor is unaware of the asset's true expected return denoted η. As a consequence, the investor is compelled to combine the J return forecasts and rely on the perceived return in equation

[14] A negative portfolio weight implies the investor reverses the sign of this information source's return forecast when computing their perceived return.

(17) which has the lowest aggregate forecast error amongst all other estimates for the asset's expected return.

The cross-sectional dispersion across the J forecasts of the μ vector in equation (4) equals

$$\sigma_\mu^2 = \frac{1}{J-1} \sum_{j=1}^{J} (\mu_j - \bar{\mu})^2 , \tag{20}$$

where $\bar{\mu}$ is defined as the average return forecast

$$\bar{\mu} = \frac{1}{J} \sum_{j=1}^{J} \mu_j , \tag{21}$$

does not have an explicit role in our solution for the information portfolio. Nonetheless, equation (20) offers an economically intuitive definition for expected return uncertainty. For example, dispersion in the return forecasts inferred from price targets immediately generates a proxy for σ_μ^2 although sources of information besides price targets are also available to the investor. Note that $\bar{\mu}$ in equation (21) is not the optimal estimate for an asset's expected return unless every information source is equally skilled with an identical true but unknown accuracy.

To clarify, $W^T \Theta W$ and σ_μ^2 are not equivalent. Indeed, if all information sources issue identical return forecasts, then σ_μ^2 in equation (20) is zero and the information portfolio is irrelevant since any combination of return forecasts yields the same perceived return under the $W^T \mathbf{1} = 1$ constraint. However, provided η remains unknown, the aggregate forecast error $W^T \Theta W$ is positive.

Intuitively, the optimal information portfolio cannot prevent information sources from disagreeing on an asset's expected return and generating a large σ_μ^2 value. Instead, by minimizing their aggregate forecast error $W^T \Theta W$, our optimal information portfolio finds the most "accurate" perceived return.

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

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\sigma_2^2 + \sigma_1^2} \begin{bmatrix} \sigma_2^2 \\ \sigma_1^2 \end{bmatrix} . \tag{22}$$

Therefore, the investor's perceived return equals

$$\frac{\sigma_2^2\,\mu_1 + \sigma_1^2\,\mu_2}{\sigma_1^2 + \sigma_2^2} , \tag{23}$$

while

$$\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} , \tag{24}$$

is the aggregate forecast error of the perceived return in equation (23).

According to equation (23), the return forecast issued by a more accurate information source has a larger portfolio weight and greater influence on the investor's perceived return. Intuitively, accuracy enhances the credibility of an information source. The utility maximization approach in Cheng, Liu and Qian (2005) produces a pair of weights similar to equation (22) for signals issued by sell-side versus buy-side analysts.

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

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} \begin{bmatrix} \sigma_2^2 - \sigma_{12} \\ \sigma_1^2 - \sigma_{12} \end{bmatrix} . \tag{25}$$

The perceived return for the asset equals

$$\frac{\sigma_2^2\,\mu_1 + \sigma_1^2\,\mu_2 - \sigma_{12}\,(\mu_1 + \mu_2)}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} , \tag{26}$$

while

$$\frac{\sigma_1^2\sigma_2^2 - (\sigma_{12})^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} , \tag{27}$$

is the aggregate forecast error of the perceived return in equation (26).

A negative covariance, $\sigma_{12} < 0$, between two information sources represents "offsetting" forecast errors. Appendix C proves that a negative covariance reduces the aggregate forecast error in equation (27). We utilize this property in Section 4 to demonstrate that our optimal information portfolio weights generate perceived returns which exhibit the appearance of several behavioral biases.

3 Information Source Accuracy and Momentum

In this section, we begin by investigating the relationship between the number of relevant forecast errors in equation (1) and the estimated accuracy of a single information source. We then study two information sources whose time-varying information portfolio weights instill momentum into the perceived return, especially when the asset's expected return uncertainty is high and few relevant forecast errors are available to estimate the accuracy of each information source.

Intuitively, a small n undermines the investor's ability to estimate the true accuracy of each information source. More formally, the next two subsections demonstrate that the estimates $\sigma_{j,t}^2$ generated by equation (2) are likely to increase, along with the aggregate forecast error $W^T \Theta W$ in equation (18) as a consequence, when the number of relevant forecast errors is limited. As additional forecast errors become available, the optimal information portfolio weights are updated and generate momentum along with subsequent reversals in the perceived return.

The uncertainty surrounding an individual asset's expected return increases when information sources examine distinct state variables, have unique forecasting techniques for their state variable, and interpret its return implications differently. Moreover, a public information source does not necessarily disclose these components of their return forecast. Instead, at each point in time, the investor observes a collection of return forecasts and relies on the estimated accuracy of each information source to define the Θ matrix. However, the tendency for $W^T \Theta W$ to increase due to larger $\sigma_{j,t}^2$ estimates after corporate events which reduce n does not require the investor to understand the origin of every return forecast.

3.1 Return Implications of a State Variable

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

$$\mu_{j,t} \quad = \quad \hat{\alpha} + \hat{\beta} V_t \, . \tag{28}$$

The hats signify the unknown coefficients of the transformation, while the state variable V_t in equation (28) is not random. As mentioned above, other information sources may employ a transformation different than equation (28) or interpret another state variable when issuing their return forecast.

According to equation (29) below, the j^{th} information source calibrates the α and β coefficients in equation (28) using realized returns and state variables

$$y_{t-i} \quad = \quad \alpha + \beta V_{t-i} + \xi_{t-i} \, , \tag{29}$$

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 obtaining the estimates $\hat{\alpha}$ and $\hat{\beta}$ from the regression model in equation (29), the information source invokes equation (28) to convert V_t into $\mu_{j,t}$. This $\mu_{j,t}$ return forecast simply equals the predicted value $\tilde{y}_t$ from the linear regression in equation (29) at $t-1$.

At time t, an additional forecast error $\epsilon_t^j = y_t - \mu_{j,t}$ is appended to the time series of forecast errors in equation (1). In particular, the realization of $\left(\epsilon_t^j\right)^2$ at t augments equation (2) when estimating the j^{th} information source's accuracy. To illustrate the importance of n, the expectation of the squared forecast error ϵ_t^j at $t-1$ is evaluated as[15]

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

[15]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 the linear regression in equation (29) providing unbiased coefficient estimates. Thus, equations (28) and (29) imply the j^{th} information source issues unbiased return forecasts. This simplification is without loss of generality since equation (2) evaluates the mean-squared error of information sources and therefore does not require them to issue unbiased return forecasts.

For large n, the $\hat{\alpha}$ and $\hat{\beta}$ estimates converge to α and β respectively, implying equation (30) reduces to $Var\left[\xi_t\right]$. Therefore, equation (30) converges to σ_ξ^2 which equals the unknown true accuracy $\sigma_{j,*}^2$ of the j^{th} information source. Recall that the investor cannot compute the decomposition in equation (30) if the transformation in equation (28) is not disclosed by the information source.

However, when n is small, estimation error in $\hat{\alpha}$ and $\hat{\beta}$ is severe. Lewellen and Shanken (2002) examine the asset pricing implications of parameter uncertainty and demonstrate that return predictability cannot necessarily be exploited by investors. In our framework, a small n can undermine the credibility of a knowledgeable information source or an important state variable. For example, Jagannathan and Wang (2005) find that consumption explains cross-sectional returns but is dominated by the SMB and HML factors of Fama-French (1993) in empirical applications due to the limitations of consumption data.

Equation (30) also illustrates the importance of predictability in the return implications of a state variable. If the conversion of V_t into $\mu_{j,t}$ is deterministic, implying the ξ_{t-i} error terms in equation (29) are identically zero, then the α and β coefficients are known.[16] Conversely, when the relationship between a stock's expected return and a state variable is unreliable, the information source's estimated accuracy in equation (2) is likely to be poor. Nonetheless, the investor can overestimate an information source's true accuracy since equation (30) is the expectation of next period's squared forecast error. The likelihood that an unskilled information source, by chance, has a low mean-squared forecast error is higher when n is small. For example, the relationship between y_{t-i} and V_{t-i} could eventually be discredited once additional return and state variable realizations are available, but nonetheless improve an information's source's estimated accuracy at an earlier point in time. Thus, the return implications of a state variable revealed "ex-post" (after obtaining additional forecast errors) to be spurious can influence the investor's "ex-ante" optimal information portfolio. Overall, the investor is more likely to underestimate or overestimate the true accuracy of information sources when fewer relevant forecast errors are available.

Finally, even the idealized environment in equation (30) has two important complications. First, the α and β parameters may be time-varying, which complicates their estimation even as n increases. Second, as discussed in the next subsection, V_t could represent a forecast for the state variable.

[16] Transforming price targets into return forecasts involves a deterministic (perfectly predictable) function, although not the linear relationship in equation (28).

3.2 State Variable Dynamics

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

$$\tilde{V}_t \;=\; \hat{a} + \hat{b}V_{t-1} \,. \tag{31}$$

The dynamics of V_t are estimated by the information source at $t-1$ as

$$V_{t-i} \;=\; a + bV_{t-i-1} + \zeta_{t-i} \,, \tag{32}$$

using data 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 (32) is utilized to estimate the a and b coefficients, while the ζ_{t-i} error terms signify the random evolution of the state variable. The $\tilde{V}_t$ notation contains a tilde to emphasize that the information source is forecasting this state variable, in contrast to equation (28) where V_t is known.

When equation (28) with known α and β parameters is combined with equation (31), the following return forecast is generated by the j^{th} information source

$$\begin{aligned} \mu_{j,t} &= \alpha + \beta\tilde{V}_t \\ &= \alpha + \beta\left[\hat{a} + \hat{b}V_{t-1}\right] . \end{aligned} \tag{33}$$

For clarification, the conversion of the state variable into its return forecast continues to be specified by equation (28). 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 estimated accuracy. An economy in which the α, β, a and b coefficients all require calibration would produce a complicated estimation error in equation (35) below involving cross-products. In addition, j superscripts implicitly index the α, β, a and b coefficients as well as the ξ and ζ error terms but they are omitted for notational simplicity.

Inserting the true dynamics of the state variable in equation (32) into equation (29) implies

$$y_{t-i} \;=\; \alpha + \beta\left[a + bV_{t-i-1} + \zeta_{t-i}\right] + \xi_{t-i} \,. \tag{34}$$

When combined, equations (33) and (34) imply the following expectation at time $t-1$ for next period's squared forecast error[17]

$$
\begin{aligned}
E\left[\epsilon_t^j\right]^2 &= Var\left[y_t - \mu_{j,t}\right] \\
&= Var\left[\xi_t\right] + \beta^2 Var\left[\zeta_t\right] \\
&+ \left\{\beta^2 Var\left[\hat{a}\right] + \beta^2\left(V_{t-1}\right)^2 Var\left[\hat{b}\right] + 2\beta^2 V_{t-1} Cov\left[a-\hat{a}, b-\hat{b}\right]\right\} \qquad (35) \\
&= \text{Transformation Uncertainty} + \text{State Variable Uncertainty} \\
&+ \text{Estimation Error in State Variable Dynamics}\,.
\end{aligned}
$$

To clarify, $Var\left[\zeta_t\right]$ 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 the second line of equation (35) tends toward zero as $n \to \infty$, implying equation (2) converges to $\sigma_\xi^2 + \beta^2\sigma_\zeta^2$ which equals the j^{th} information source's true accuracy $\sigma_{j,*}^2$.

Equation (35) implies that after controlling for $Var\left[\xi_t\right]$, information sources who condition their return forecasts on predictable state variables are likely to have superior estimated accuracies in comparison to those conditioning on unpredictable state variables. Indeed, if V_t evolves deterministically, then $Var\left[\zeta_t\right]$ equals zero while the a and b coefficients are known, implying equation (35) reduces to $Var\left[\xi_t\right]$.

In general, state variable forecasts and their return implications are not required to arise from linear relationships as in our previous illustrations. Nonetheless, the importance of n as well as predictability in state variable dynamics and their return implications extends beyond these linear specifications when estimating an information source's accuracy. Specifically, the estimation errors arising from a small n would likely increase the $\sigma_{j,t}^2$ estimates from equation (2) and consequently the aggregate forecast error $W^T\Theta W$ in equation (18). In addition, state variables whose dynamics and return implications exhibit predictability improve the estimated accuracy of their corresponding information source. This improvement occurs even if the predictability is eventually found to be spurious. More importantly,

[17] Equations (31), (33) and (34) imply $\mu_{j,t}$ is an unbiased return forecast. Specifically, $E\left[\epsilon_t^j\right]$ equals zero since the linear regression in equation (32) ensures $E\left[\hat{a}\right] = a$ and $E\left[\hat{b}\right] = b$ under the assumptions imposed on ζ_{t-i} and ξ_{t-i}. However, this property is not a requirement of information portfolio theory since equation (2) minimizes mean-squared error and does not require unbiased return forecasts.

as additional forecast errors become available, fluctuations in the $\sigma^2_{j,t}$ estimates cause the optimal information portfolio weights to be updated. The ability of time-varying optimal information portfolio weights to induce momentum and reversals is analyzed in the next subsection.

3.3 Momentum in Perceived Return

To examine momentum (and reversals) in the investor's perceived return, we consider an optimistic and pessimistic information sources for a firm that has recently initiated a large investment. The profitability (earnings / cashflow) of this investment represents the relevant state variable in our subsequent analysis. More importantly, differences of opinion between the information sources are not assumed to result from psychological biases.

At the initial timepoint t_1, the high return forecast from the optimistic information source is denoted $\mu_{H,1}$ while its low return counterpart is denoted $\mu_{L,1}$. For simplicity but without loss of generality, assume these return forecasts are independent with $\sigma^2_{H,1}$ equaling $\sigma^2_{L,1}$ at t_1. According to equation (23) in Corollary 1, the investor's perceived return over the $(t_1, t_2]$ horizon is the average of the two forecasts.

During the $(t_1, t_2]$ interval, information regarding the success of the investment is revealed, with the firm's realized return denoted $r_{1,2}$ over this horizon. This realized return is identical to y_2 in equation (1) but notated more descriptively as $r_{1,2}$. In particular, there are two scenarios at t_2, the first indicating success and the second failure. The ex-ante probability attached to these scenarios is irrelevant when the firm's realized return sequence is studied at t_3 or a later timepoint. Furthermore, over the $(t_2, t_3]$ interval, assume $\mu_{H,2}$ continues to exceed $\mu_{L,2}$ with the disparity between these forecasts depending on the uncertainty prevailing at time t_2 surrounding the firm's expected return. To illustrate return extrapolation in our framework, consider the following two scenarios defined by $r_{1,2}$.

Investment appears to be successful over $(t_1, t_2]$ interval:

$$\left\{ \begin{array}{l} \text{Realized return } r_{1,2} \text{ is high} \\ \sigma^2_{H,2} < \sigma^2_{L,2}\text{, estimated accuracy of high return information source improves since } r_{1,2} \text{ is high} \\ \text{Perceived return over } (t_2, t_3] \text{ horizon is closer to } \mu_{H,2} \\ \hline \text{Perceived return appears to extrapolate from high realized return} \end{array} \right.$$

Investment appears to be failing over $(t_1, t_2]$ interval:

$$\begin{cases} \text{Realized return } r_{1,2} \text{ is low} \\ \sigma^2_{L,2} < \sigma^2_{H,2}, \text{ estimated accuracy of low return information source improves since } r_{1,2} \text{ is low} \\ \text{Perceived return over } (t_2, t_3] \text{ horizon is closer to } \mu_{L,2} \\ \hline \text{Perceived return appears to extrapolate from low realized return} \end{cases}$$

The disparity between $\sigma^2_{H,2}$ and $\sigma^2_{L,2}$ at t_2 is similar to the expected return uncertainty in equation (20). If n is small, either $\mu_{H,2}$ or $\mu_{L,2}$ exerts a strong influence on the investor's perceived return as the optimal information portfolio is updated to reflect changes in the estimated accuracy of each information source once $r_{1,2}$ is revealed at t_2. In contrast, a large number of relevant forecast errors enables the investor to better estimate the true accuracy of both information sources, implying their information portfolio weights would be relatively stable.

More formally, let t_1 denote the initiation of a large investment which is assumed to generate expected return uncertainty. This uncertainty persists until t_2, when the winner and loser portfolios are constructed. Thus, formation and holding periods of the momentum strategy correspond to $[t_1, t_2)$ and $[t_2, t_3)$ respectively.[18] Information portfolio theory cautions that momentum profits can arise from idiosyncratic events and firm characteristics which reduce the number of relevant forecast errors, thereby undermining the estimation of each information source's accuracy, as illustrated in equations (30) and (35). In the aftermath of these events, the optimal information portfolio weights fluctuate as additional forecast errors become available. Empirically, events capable of reducing n are studied by Jackson and Johnson (2006).[19]

In summary, equations (30) and (35) imply a tendency towards larger $\sigma^2_{j,t}$ estimates in equation (2) after significant corporate events. These larger estimates increase an asset's aggregate forecast error $W^T\Theta W$ which equations (15) and (16) translate into higher risk from the investor's perspective. In particular, the asset's ex-ante return variance equals $W^T\Theta W + \nu$, with the aggregate forecast error

[18] Several intermediate return forecasts and their associated forecast errors are possible within the $[t_1, t_2)$ horizon. Thus, n is not necessarily equal to one at t_2.

[19] 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. This persistence suggests their impact on a firm's earnings dynamics are not immediately understood.

augmenting the asset's true return variance. Consequently, momentum is strongest when the investor confronts higher *forecast* risk. For emphasis, the optimal information portfolio has no control over the firm's expected return uncertainty in equation (20). Instead, information portfolio theory's explanation for momentum assumes σ_{μ}^2 is non-zero. Zhang (2005) finds that momentum profits are strongest for firms experiencing high levels of expected return uncertainty. Whether these momentum profits arise from our optimal information portfolio weights or investor psychology is addressed empirically in Section 5.

One may argue that the empirical evidence concerning long term reversals motivates a mean-reverting prior distribution when issuing return forecasts. However, if the optimistic (pessimistic) information source at t_1 decreases (increases) its return forecast at t_2, then expected return uncertainty is reduced. Indeed, if $\mu_{H,1}$ and $\mu_{L,1}$ converge to a common return forecast at t_2, then the uncertainty created by the investment is resolved during the $(t_1, t_2]$ horizon and σ_{μ}^2 is zero. Hence, return extrapolation attributable to the optimal information portfolio continues as long as there is uncertainty regarding the firm's expected return. Our explanation only requires persistence in the relative optimism of the information sources. Specifically, we assume the same information source issues *relatively* high (low) forecasts until the expected return uncertainty dissipates.

Intuitively, returns exhibit momentum due to events whose return implications are not immediately understood and agreed upon by all information sources, although the return forecasts issued by information sources are allowed to be updated at intermediate timepoints.[20] Overall, the estimated accuracy of an information source who issues high (low) return forecasts improves following high (low) realized returns. This improvement increases their optimal information portfolio weight. Therefore, information sources are assigned larger information portfolio weights when their past, hence current, return forecasts are similar to realized returns. As a consequence, the investor's perceived return appears to *extrapolate* from past returns.[21]

Reversals in the perceived return can also be attributed to the optimal information portfolio. The

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

[21] Thus, multiple return sequences exist ex-ante with the investment's success determining a realized return sequence and the corresponding series of optimal information portfolio weights. 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.

existing literature usually obtains momentum and reversals from the misinterpretation of a single state variable. In Barberis, Shleifer and Vishny (1998), a firm's true earnings evolve as a random walk while the investor believes this state variable is either in a trending or mean-reverting regime. Instead of two regimes, our framework considers multiple state variables, hence multiple sources of information whose time-varying optimal portfolio weights induce reversals in the investor's perceived return. As an example, prior returns could temporarily constitute the most accurate source of information for an IPO when its earnings are difficult to forecast. However, the influence of prior returns would diminish once the IPO's earnings dynamics are better understood.[22]

More formally, consider the earlier environment where one information source is relatively more optimistic than its counterpart. The optimal information portfolio weights assigned to $\mu_{H,3}$ and $\mu_{L,3}$ depend on the estimated accuracy of the optimistic and pessimistic information sources at t_3, hence the realized returns over the $(t_1, t_2]$ and $(t_2, t_3]$ horizons. Even if the asset's true expected return η is constant, reversals in the perceived return occur when realized returns alter the relative ranking of the information sources by their estimated accuracy. In particular, the optimistic (pessimistic) information source can be more accurate at t_2 than its pessimistic (optimistic) counterpart but less accurate at t_3 if $r_{1,2}$ is high (low) while $r_{2,3}$ is low (high). This situation corresponds to the intuition behind the above IPO example involving past returns and earnings. The potential for reversals to result from realized returns that benefit one information source at the expense of another increases with the asset's true return variance. Specifically, for a given n, a higher value of ν implies more variable realized returns, and consequently more dramatic fluctuations in the information portfolio weights.

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 zero mean. However, information portfolio theory allows different information sources to utilize distinct statistical methodologies when forecasting state variables or ascertaining their return implications. Disparate return forecasts can also originate from information sources analyzing different state variables. Consequently, the law of iterated expectations

[22] Therefore, another implication of information portfolio theory is that conflicts of interest which compromise the accuracy of affiliated analyst forecasts have a greater impact on an IPO's perceived return before non-affiliated analysts can establish their credibility.

does not ensure homogenous return forecasts across the J information sources. Instead, disagreement over future state variables and their impact on the asset's return justifies the existence of multiple information sources. From a practical perspective, we assume the asset's valuation is sufficiently complex to prevent information sources from obtaining identical estimates for its expected return.

Our framework's structure allows an asset's return forecasts to be determined by multifactor asset pricing models when factors such as the market return are interpreted as state variables. An individual asset's true expected return is unknown for several reasons in these formulations; randomness in the dynamics of the factors, estimation error in the factor loadings for individual assets, and uncertainty regarding the number of required factors as well as their composition. Appendix B discusses these issues in more detail. Alternative return forecasts could be generated by price targets and intrinsic value measures which are studied in Brav and Lehavy (2003) and Lee, Myers and Swaminathan (1999) respectively.

However, the standard econometric approach when testing market efficiency estimates a single expected return, which fails to account for uncertainty surrounding the interpretation of available information. Instead, realized factor returns replace ex-ante forecasts, while their corresponding beta coefficients are also assumed to be known. Moreover, these firm-specific beta coefficients are usually estimated using time series data and fixed over a given horizon. In contrast, our information portfolio weights are time-varying and apply to a cross-section of return forecasts.

One concern regarding information portfolio theory may be the appearance of *systematic* expected return biases that appear to indicate the presence of psychological biases. In the next section, we demonstrate that our optimal information portfolio weights induce return characteristics that mimic psychological biases in the behavioral finance literature.

4 Properties of the Perceived Return

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

Return characteristics induced by the optimal information portfolio that mimic psychological biases coincide with expected return uncertainty. Indeed, if all information sources issue identical return forecasts for an asset, then the information portfolio is irrelevant since any combination of these forecasts yields the same perceived return. Furthermore, the appearance of overconfidence, biased self-attribution, representativeness, conservatism and limited attention is more pronounced when n is small due to larger fluctuations in the information portfolio weights and greater difficulty in estimating the true accuracy of each information source.

4.1 Appearance of Overconfidence and Biased Self-Attribution

To analyze the appearance of overconfidence in the perceived return, we examine two information sources. This first information source is private and the second public, with their return forecasts and estimated accuracies denoted by pr and pb subscripts respectively. The investor's perceived return appears to exhibit overconfidence *whenever* the information portfolio weight w_{pr} for a private information source exceeds the information portfolio weight w_{pb} of a public information source. Later in this subsection, we demonstrate that the appearance of overconfidence can occur even when the true accuracy of the private and public information sources are identical.

Interpretation 1. *Appearance of Overconfidence*
Corollary 1 implies the following information portfolio weights for private and public information

$$\begin{bmatrix} w_{pb} \\ w_{pr} \end{bmatrix} = \frac{1}{\sigma_{pr}^2 + \sigma_{pb}^2} \begin{bmatrix} \sigma_{pr}^2 \\ \sigma_{pb}^2 \end{bmatrix} . \tag{36}$$

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

$$\frac{1}{\sigma_{pr}^2 + \sigma_{pb}^2} \left[\sigma_{pr}^2 \, \mu_{pb} + \sigma_{pb}^2 \, \mu_{pr} \right] , \tag{37}$$

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

According to equation (37), whenever a private information source is more accurate than its public counterpart, the investor's perceived return mimics overconfidence. Recall that multiple private information sources may exist because forecasts for state variables such as earnings require further interpretation by the investor to become return forecasts. In contrast, price targets yield explicit return forecasts which constitute public information sources. Overall, let the return forecast μ_{pr} in equation (37) be associated with the investor's most accurate source of private information. Provided this private information source has been more accurate than a public information source over the last n periods, the investor appears to be overconfident. We formalize this property after introducing a perceived return characteristic which mimics biased self-attribution.

In the context of information portfolio theory, the investor exhibits the appearance of biased self-attribution when the estimated accuracy for one of their private information sources is superior to a public information source, while another private information source is less accurate.[23] Consider two private information sources along with the original public information source. The private information source which is more accurate than the public information source is denoted with a c subscript, while the less accurate private information source has a d subscript.

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*

$$[w_c,\, w_d,\, w_{pb}] \;=\; \frac{1}{D}\left[\sigma_d^2\,\sigma_{pb}^2,\, \sigma_c^2\,\sigma_{pb}^2,\, \sigma_c^2\,\sigma_d^2\right] , \tag{38}$$

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*

$$\frac{\left[\sigma_c^2\,\mu_d + \sigma_d^2\,\mu_c\right]\sigma_{pb}^2 + \sigma_c^2\,\sigma_d^2\,\mu_{pb}}{D} , \tag{39}$$

which is influenced more by μ_c *than* μ_d.

The $\sigma_d^2 > \sigma_c^2$ property ensures the information portfolio weight w_c for μ_c exceeds the information portfolio weight w_d assigned to the μ_d return forecast. Thus, the investor's perceived return gravitates

[23] Intuitively, to connect estimated accuracy with terminology in the psychology literature, *confirming* private information sources are more accurate than a public information source according to equation (2), while *disconfirming* private information sources have been less accurate than all public information sources over the last n periods.

towards their 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 weights.

Interestingly, the investor's perceived return may exhibit the appearance of overconfidence even if their private information sources are inaccurate on average since less accurate private information sources receive smaller information portfolio weights. For example, if the investor successfully predicts the return implications of industry characteristics, but cannot reliably interpret a firm's earnings, then the importance of industry data is accentuated by the information portfolio at the expense of earnings. By implication, the investor pursues trading strategies derived from private sources of information which have provided them with *individual* success, regardless of the technique's generality.

Consider two private information sources whose estimated accuracies are denoted $\sigma^2_{pr,1}$ and $\sigma^2_{pr,2}$ respectively. Let p_j represent the probability that the true accuracy of the j^{th} private information source is less than its public counterpart which summarizes their *relative* skill. These probabilities equal one-half when their return forecasts originate from a common distribution, which implies the true accuracy of each information source is identical.[24] The following four scenarios summarize the appearance of overconfidence and biased self-attribution after n periods.

[24]Having the three return forecasts emanate from the true return distribution is for ease of illustration but without loss of generality. If the true accuracy of each information source is distinct, then the optimal information portfolio simply converges to a different set of weights as n increases although the economic intuition is identical.

Scenario	Estimated Accuracies	Probability	Investor Appears to Exhibit
A	$\sigma^2_{pr,1}\,,\,\sigma^2_{pr,2} \;<\; \sigma^2_{pb}$	$(1-p_1)\,(1-p_2)$	Overconfidence from both private sources
B	$\sigma^2_{pr,1} \;<\; \sigma^2_{pb} \;<\; \sigma^2_{pr,2}$	$(1-p_1)\,p_2$	Overconfidence from 1^{st} private source and biased self-attribution
C	$\sigma^2_{pr,2} \;<\; \sigma^2_{pb} \;<\; \sigma^2_{pr,1}$	$p_1\,(1-p_2)$	Overconfidence from 2^{nd} private source and biased self-attribution
D	$\sigma^2_{pb} \;<\; \sigma^2_{pr,1}\,,\,\sigma^2_{pr,2}$	$p_1\,p_2$	No Overconfidence

Observe that the perceived return exhibits the appearance of overconfidence in scenarios A, B and C, with a cumulative probability of $1 - p_1\,p_2$. Therefore, when p_1 and p_2 both equal one-half, the investor's perceived return appears to exhibit overconfidence in 75% of the scenarios. Furthermore, in scenarios B and C, an 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%. Nonetheless, the appearance of overconfidence and biased self-attribution occurs despite the two private and public information sources possessing equal skill at forecasting returns. The appearance of overconfidence and biased self-attribution is further justified when a perceived return characteristic which mimics limited attention is examined later in this section. For emphasis, the appearance of these biases occurs despite the three information sources having identical true accuracies.

Finally, our results continue to apply when there are more public than private information sources.[25] When two public information sources and one private information source are considered, the following

[25] 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.

four scenarios are relevant after n periods.

Scenario	Estimated Accuracies					Probability	Investor Appears to Exhibit
A	$\sigma^2_{pb,1}\,,\,\sigma^2_{pb,2}$	$<$	σ^2_{pr}			$p_1\,p_2$	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 and biased self-attribution
C	$\sigma^2_{pb,2}$	$<$	σ^2_{pr}	$<$	$\sigma^2_{pb,1}$	$p_2\,(1-p_1)$	Limited overconfidence from 1^{st} public source and biased self-attribution
D			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 two public information sources. Indeed, the investor's private information sources may perform poorly on average. 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 an information source's estimated accuracy improves when predictability in a state variable's dynamics as well as its return implications reduces its mean-squared forecast error. These properties imply that *trends* can increase an information source's portfolio weight since trends imply predictability.[26] Even predictability in state variable dynamics and their relationship with realized returns that is subsequently found to be spurious when n is larger can influence the optimal information portfolio at earlier timepoints. This property distinguishes between

[26] As an example, a strong trend in a binomial sequence occurs when its realizations consist predominately of either up or down movements, implying the estimated binomial probability over the last n observations is near zero or one.

ex-ante estimates of an information source's accuracy, computed when a limited number of relevant forecast errors were available, and ex-post estimates which are only known later in the sample period.

Consider two information sources, labeled consistent and inconsistent, with the former arising from predictability in the dynamics of a state variable or its return implications. The consistent and inconsistent information sources are denoted by c and d subscripts respectively, with the property $\sigma_I^2 > \sigma_C^2$ induced by predictability underlying the consistent source of information.

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*

$$\begin{bmatrix} w_C \\ w_I \end{bmatrix} = \frac{1}{\sigma_I^2 + \sigma_C^2} \begin{bmatrix} \sigma_I^2 \\ \sigma_C^2 \end{bmatrix}, \tag{40}$$

which implies the perceived return

$$\frac{1}{\sigma_C^2 + \sigma_I^2} \left[\sigma_C^2 \, \mu_I + \sigma_I^2 \, \mu_C \right], \tag{41}$$

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 consistency without an information source possessing superior knowledge regarding the true dynamics of a state variable or its relationship with future returns. Indeed, when n is small, the consistency of an information source could be temporary. For example, prior returns or industry trends may generate consistent sources of information for an IPO until its earnings dynamics can be reliably estimated.

More formally, assume the return forecasts from two information sources both emanate from the true return distribution. Therefore, the true accuracy of these two information sources is identical, implying any trend that causes either of the two information sources to have a higher estimated accuracy is spurious. Nonetheless, for a finite n, equation (2) implies one of the information sources is more accurate as described by the following two scenarios.[27]

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

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

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 accuracies estimated using equation (2), this distance decreases as n increases since the return forecasts from both information sources arise from the same distribution.

Furthermore, one cannot analyze the return implications of a single state variable in isolation unless the estimated accuracy of its corresponding information source is far superior to other sources of information. This caveat limits a researcher's ability to detect conservatism by examining a single information source.

Interpretation 4. *Appearance of Conservatism*
According to Corollary 1, the investor's perceived return

$$\frac{\sigma_C^2\,\mu_I + \sigma_I^2\,\mu_C}{\sigma_C^2 + \sigma_I^2}\,, \tag{42}$$

is an aggregate estimate for the asset's expected return. Therefore, conservatism cannot be established by evaluating a single source of information.

Moreover, suppose the return implications of two state variables, such as earnings and sales, are negatively correlated after a large investment or period of price discounting. Equation (42) would be replaced by

$$\frac{\sigma_C^2\,\mu_I + \sigma_I^2\,\mu_C - \sigma_{C,I}\,\left(\mu_C + \mu_I\right)}{\sigma_C^2 + \sigma_I^2 - 2\sigma_{C,I}} \tag{43}$$

according to equation (26). As demonstrated in the next subsection, negatively correlated return forecasts receive larger information portfolio weights and have more influence on the investor's perceived return relative to the estimated accuracy of their information sources. Consequently, examining negatively correlated information sources in isolation, rather than ascertaining their aggregate influence, can create the impression that the investor's perceived return exhibits conservatism.

4.3 Appearance of Limited Attention

The $\sigma_{1,2}$ covariance term in Corollary 2 incorporates the appearance of limited attention into the perceived return. Barber, Odean and Zhu (2003) present empirical evidence of this bias for individual investors. In the context of information portfolio theory, the investor assigns lower portfolio weights to information sources whose return forecasts are positively correlated with a more accurate source of information. Intuitively, 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.[28] In contrast, if their return forecasts are negatively correlated and offer alternative perspectives on the asset's expected return, then the investor benefits from analyzing both information sources. More formally, consider the portfolio weights in Corollary 2

$$\begin{aligned} w_1 &= \frac{\sigma_2^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} \\ w_2 &= \frac{\sigma_1^2 - \sigma_{12}}{\sigma_2^2 + \sigma_1^2 - 2\sigma_{12}} . \end{aligned} \tag{44}$$

If the two forecasts are independent, then σ_{12} equals zero and both portfolio weights are positive. However, when σ_{12} equals σ_1^2, the Cauchy-Schwartz inequality implies $\sigma_1^2 \leq \sigma_2^2$ with the first information

[28] *Herding* by information sources induces positive correlation between a subset of return forecasts. This positive correlation further reduces the optimal portfolio weights assigned to less accurate information sources and mitigates the impact of herding on the investor's perceived return.

source being more accurate than the second, while the portfolio weights in equation (44) become[29]

$$\begin{aligned} w_1 &= \frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2 - \sigma_1^2} = 1 \\ w_2 &= \frac{\sigma_1^2 - \sigma_1^2}{\sigma_2^2 - \sigma_1^2} = 0\,. \end{aligned} \tag{45}$$

Consequently, a large positive covariance between the two information sources eliminates the second return forecast from the perceived return as illustrated in Figure 1. In contrast, negatively correlated forecasts have a greater influence over the investor's perceived return relative to their estimated accuracy. Therefore, when attempting to detect conservatism, it is essential to evaluate the aggregate impact of contradictory information instead of focusing on the return implications of a single state variable.

Limited attention also magnifies the appearance of overconfidence since private information sources receive larger portfolio weights than their estimated accuracies alone justify if the previous return forecasts they issued have been negatively correlated with public information sources. In particular, *contradictory* private sources of information can be overweighted relative to more accurate but positively correlated public information sources.

In summary, the portfolio weights assigned to information sources are dependent on the estimated covariances between their prior forecast errors in equation (3). For example, the investor may rely on broadly defined sector information rather than firm-specific characteristics if the latter are positively correlated within an industry.

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 are not prevented from incorporating psychology. Indeed, 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.

[29]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 implies $\sigma_1^2 \leq \sigma_2^2$.

For example, suppose all J return forecasts are identical and equal to $\mu_\circ$, with this common expectation further *assumed* to be the result of at least one psychological bias. The investor's perceived return equals $\mu_\circ$ regardless of the information portfolio and reflects investor psychology. Conversely, the use of psychology could increase the expected return uncertainty in equation (20) due to disagreements over the relative importance of specific behavioral biases. As a result, the relevance of information portfolio theory is enhanced by differences of opinion regarding investor psychology. The exact decomposition of the perceived return into the effects of psychology versus the optimal information portfolio is ultimately an empirical question.

By estimating an investor's expected return, information portfolio theory enhances rather than contradicts utility maximization. Indeed, the investor's perceived return and its aggregate mean-squared forecast error are critical inputs in further asset pricing applications. Proposition 3 below, whose proof is in Appendix D, provides a utility maximizing application of information portfolio theory.

Proposition 3. *Assume the investor has a negative exponential utility function, $U(M) = 1 - e^{-\gamma M}$, with initial wealth M. Under the return distribution in equation (15), the optimal fraction of wealth f invested in the risky asset equals*

$$f \;=\; \frac{\mathbf{1}^T \Theta^{-1} \left(\mu - r_f \mathbf{1}\right)}{\gamma\, M\, \left[1 + \nu \mathbf{1}^T \Theta^{-1} \mathbf{1}\right]}\,, \tag{46}$$

where r_f represents the riskfree interest rate.

As an explicit illustration, consider two identical return forecasts issued by information sources with identical accuracies ($\mu_1 = \mu_2 = \mu_\circ$, $\sigma_1^2 = \sigma_2^2 = \sigma_\circ^2$) with σ_{12} describing their off-diagonal covariance element as in Corollary 2. Under these specifications, the solution for f in equation (46) reduces to

$$f \;=\; \frac{2\,\left(\mu_\circ - r_f\right)}{\gamma\, M} \left(\frac{1}{2\nu + \sigma_\circ^2 + \sigma_{12}}\right). \tag{47}$$

Observe that when information sources forecast higher returns or are more accurate, the investor increases their exposure to the risky asset. When the return forecasts are negatively correlated, the investor also purchases more of the risky asset. According to equation (47), accurate return forecasts offset the investor's risk aversion.

As a special case of equation (47), the fraction of wealth allocated to the risky asset equals $\frac{\eta - r_f}{\gamma\, M\, \nu}$ when the asset's true expected return is known since $\sigma_\circ^2$ and σ_{12} are zero while $\mu_\circ = \eta$.

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 variability. Alternatively, the attribution bias in Gervais and Odean (2001) utilizes improper Bayesian updating to create overconfidence.

In contrast, the perceived return characteristics induced by the time-varying optimal information portfolio weights arise from aggregating across multiple return forecasts and provide testable implications of information portfolio theory which are independent of any prior distribution. 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.

5 Empirical Implementation

Information portfolio theory does not assume that investor beliefs are influenced by psychological biases. Instead, uncertainty surrounding an asset's expected return enables our time-varying optimal information portfolio weights to induce these return characteristics. 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. First, the return characteristics induced by information portfolio theory 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, the appearance of return characteristics that mimic overconfidence, biased self-attribution, representativeness, conservatism and limited attention in the perceived return is more prevalent, along with return momentum. Corporate restructurings, significant investments as well as technological innovations all reduce the relevance of previous observations and increase the uncertainty surrounding a firm's expected return.

Second, predictability in the dynamics of a state variable or its relationship with an asset's return influences the investor's perceived return provided this predictability improves an information source's estimated accuracy. The possibility that statistically insignificant predictability in state variable dynamics or spurious return implications influence the optimal information portfolio is most likely in periods of high expected return uncertainty when there are only a limited number of relevant forecast errors available to estimate the accuracy of each information source.

Third, negatively correlated return forecasts reduce the investor's aggregate forecast error. Consequently, contradictory sources of information have greater influence over the investor's perceived return. This property increases the magnitude of return characteristics which mimic overconfidence and conservatism.

Two testable hypotheses involving private return forecasts are also available. First, investors overweight their accurate private information sources (successes) at the expense of their less accurate private information sources (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 fewer forecast errors to estimate their true accuracy. This implication arises from n in equation (1) is specific to an information source rather than being common across all information sources.

To differentiate between the implications of psychological theories versus information portfolio theory, the estimated accuracy of each information source is crucial. In particular, for a given level of state variable uncertainty, information portfolio theory posits that momentum is highest when conditioning on an informative state variable (highly correlated with returns). In contrast, psychology predicts greater momentum when state variables are uninformative.

However, psychological biases and information portfolio theory are not necessarily incompatible. The extent to which they both influence the perceived return is ultimately an empirical question. For example, if accurate information sources incorporate investor psychology into their return forecasts, then psychology undeniably impacts the investor's perceived return.

5.2 Hypotheses and Data

In our empirical implementation, there are implicitly two information sources; the first issues return forecasts after interpreting earnings while the second considers a state variable representing everything except earnings. Testing information portfolio theory requires us to examine the transformation of earnings into return forecasts as well as the underlying variability in earnings. We examine earnings momentum strategies based on analyst forecast revisions and forecast dispersion to verify the predictions of information portfolio theory.

The relationship between returns and earnings forecasts generates our first hypothesis. Intuitively, $Var[\xi_t]$ in equation (35) is being referenced. The first hypothesis is derived from the information portfolio weights assigned to return forecasts that arise from earnings. Information portfolio theory predicts that investors focus their attention on a firm's earnings when this state variable has experienced a stronger relationship with its realized returns.

Hypothesis 1. *Earnings momentum is stronger for stocks when the return implications of their earnings are more certain.*

Our second hypothesis concerns earnings uncertainty and refers intuitively to $Var[\zeta_t]$ in equation (35). Higher earnings uncertainty translates into greater expected return uncertainty and momentum.

Hypothesis 2. *Earnings momentum is stronger for stocks with higher earnings uncertainty.*

The first hypothesis is critical to verifying information portfolio theory, while the second hypothesis also has a behavioral interpretation (Zhang (2005)). Behavioral theory (e.g. Hirshleifer (2001)) posits that psychological biases are strongest in environments with high uncertainty as well as poor information. Consequently, behavioral theory and our framework are both consistent with our second hypothesis. However, for a given level of earnings uncertainty, behavioral theory predicts stronger earnings momentum when earnings are less informative (low sensitivity of returns to earnings), while information portfolio predicts the opposite. Therefore, in contrast to psychology, the optimal information portfolio predicts investors attempt to find the "best" available sources of information, even during periods of high expected return uncertainty. Hence, the first hypothesis is crucial to distinguishing between our framework and psychological explanations for momentum.

Our empirical tests 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 (20) 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.[30] 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 earnings state variables into return forecasts.

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

$$rev_{i,t} \quad = \quad \frac{FY1_{i,t} - FY1_{i,t-1}}{P_{i,t}}\,, \tag{48}$$

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.[31]

[30]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.

[31]Additional adjustments on $rev_{i,t}$ are performed in the month when a firm announces its fiscal year earnings since analyst forecasts switch to the subsequent fiscal year 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.

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.

5.4 Earnings Uncertainty

Our theory also asserts that momentum is more pronounced when state variables are more uncertain. The uncertainty of earnings information is measured using the standard deviation of analyst forecasts scaled by stock price[32]

$$stdev_{i,t} = \frac{\sigma_{i,t}}{P_{i,t}}. \tag{49}$$

Along with the sensitivity classifications, we divide the stocks into three uncertainly groups each month according to equation (49) 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.

[32] As a robustness test, the mean analyst forecast is also used to normalize $\sigma_{i,t}$ instead of the stock price. The results under this alternative normalization are nearly identical to those using equation (49). Consequently, for brevity, they are unreported but available upon request.

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, 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.

Hypothesis 3. *Earnings momentum is strongest 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

$$REV6_{i,t} \;=\; \sum_{j=0}^{5} rev_{i,t-j}\,, \tag{50}$$

where $rev_{i,t}$ is defined in equation (48). We rank the stocks according to equation (50) 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 these profits are 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 earnings 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 for the sensitivity and uncertainty measures are not attributable to book-to-market, size and analyst coverage effects documented in the existing literature.

6 Conclusions

We estimate an individual asset's true but unknown expected return using an optimal information portfolio which minimizes the aggregate forecast error associated with a combination of multiple return forecasts. Therefore, unlike existing estimates for an asset's expected return, information portfolio theory does not assume the return implications of events such as changes in a firm's capital structure or investment strategy are immediately understood and agreed upon by market participants. Instead, motivated by the uncertainty surrounding an asset's expected return, multiple return forecasts are aggregated into an estimate for the asset's expected return.

Each return forecast is issued by an information source after interpreting a state variable. Examples of state variables include ex-ante factor returns or a firm's future earnings. However, the true accuracy of each information source is unknown and requires estimation. Therefore, prior forecast errors for an information source are utilized to estimate its accuracy. The optimal information portfolio then assigns larger information portfolio weights to the return forecasts of more accurate information sources.

The expected return implied by our optimal information portfolio exhibits the appearance of overconfidence, biased self-attribution, representativeness and conservatism as well as limited attention. However, these characteristics arise from the time-varying optimal information portfolio weights rather than psychology. Their appearance, along with momentum, are strongest during periods of high expected return uncertainty when only a limited number of relevant forecast errors are available for estimating the accuracy of information sources.

Testable implications of information portfolio theory distinct from psychology are also provided. In contrast to Bayesian frameworks, these implications are independent of any 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 information sources have a greater influence on expected returns. The importance of our sensitivity and uncertainty measures is not attributable to book-to-market, size and analyst coverage.

Several testable implications of information portfolio theory have been left for future research and are the subject of on-going research. Furthermore, applications of information portfolio theory could study return volatility and trade volume arising from fluctuations in the information portfolio weights. Extending our framework to incorporate multiple assets would also enable its cross-sectional return implications to be examined.

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

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

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

[7] Brav, A., and R. Lehavy, 2003. An empirical analysis of analysts' target prices: Short term informativeness and long-term dynamics. Journal of Finance 58, 1933-1967.

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

[9] Cheng, Y., M. H. Liu, and J. Qian, 2005. Buy-side analysts, sell-side analysts, and investment decisions of money managers. Forthcoming in Journal of Financial and Quantitative Analysis.

[10] Clement, M. B., and S. Y. Tse, 2005. Financial analyst characteristics and herding behavior in forecasting. Journal of Finance 60, 307-341.

[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] Gervais, S., and T. Odean, 2001. Learning to be overconfident. Review of Financial Studies 14, 1-27.

[17] Hirshleifer, D., 2001. Investor psychology and asset pricing. Journal of Finance 56, 1533-1596.

[18] 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.

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

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

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

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

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

[24] Lee, C., J. Myers, and B. Swaminathan, 1999. What is the intrinsic value of the Dow? Journal of Finance 54, 1693-1741.

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

[26] Nicolosi, G., L. Peng, and N. Zhu, 2004. Do individual investors learn from their trading experience? Working Paper, University of Cincinnati.

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

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

[29] 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 (5) as

$$L(W,\lambda) = \frac{1}{2}W^T\Theta W + \lambda(W^T\mathbf{1} - 1)\,, \tag{51}$$

which generates two equations

$$\frac{\partial L(W,\lambda)}{\partial W} = \Theta W + \lambda\mathbf{1} = 0 \tag{52}$$

$$\frac{\partial L(W,\lambda)}{\partial \lambda} = W^T\mathbf{1} - 1 = 0 \tag{53}$$

involving two unknowns; W and the Lagrangian multiplier λ. Equation (52) is equivalent to

$$W = -\lambda\Theta^{-1}\mathbf{1}\,. \tag{54}$$

Multiplying the transpose of equation (54) by the $\mathbf{1}$ vector yields

$$W^T\mathbf{1} \;=\; -\lambda\mathbf{1}^T\Theta^{-1}\mathbf{1} \tag{55}$$

which implies

$$1 \;=\; -\lambda\mathbf{1}^T\Theta^{-1}\mathbf{1}\,, \tag{56}$$

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

$$-\lambda \;=\; \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\,. \tag{57}$$

Substituting equation (57) into equation (54) produces the final result

$$W \;=\; \left(\frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\right)\Theta^{-1}\mathbf{1}\,, \tag{58}$$

which satisfies the constraint

$$W^T\mathbf{1} \;=\; \left(\frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}\right)\mathbf{1}^T\Theta^{-1}\mathbf{1} = 1\,. \tag{59}$$

B Multifactor Models and the Information Portfolio

A three factor version of equation (9) describes the asset's true expected return as

$$\eta \;=\; \beta_0 + \beta_1\,f_1 + \beta_2\,f_2 + \beta_3\,f_3\,, \tag{60}$$

from which a vector μ of three return forecasts may be formed to represent the return implications of each individual factor

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \;=\; \begin{bmatrix} \alpha_{0,1} + \alpha_{1,1}\,f_1 \\ \alpha_{0,2} + \alpha_{1,2}\,f_2 \\ \alpha_{0,3} + \alpha_{1,3}\,f_3 \end{bmatrix}\,. \tag{61}$$

The elements of μ in equation (61) arise from single factor versions of the linear regression in equation (9)

$$y \;=\; [\alpha_{0,j} + \alpha_{1,j}\,f_j] + \epsilon^j \quad \text{for } j = 1,\, 2,\, 3 \tag{62}$$

$$\;=\; \mu_j + \epsilon^j\,, \tag{63}$$

where ϵ^j are mean zero error terms whose variances and covariances define the Θ matrix. The $\alpha_{0,j}$ intercepts ensure equation (62) provides three unbiased return estimates for η. The three intercept terms are[33]

$$\alpha_{0,j} \;=\; \beta_0 + \sum_{k=1}^{3} \beta_k \, f_k + \left[\beta_j - \alpha_{1,j}\right] \, f_j \quad \text{for } k \neq j \, . \tag{64}$$

With the return forecasts in equation (62) conforming to equation (7), equation (12) implies $W^T \mu$ offers an *unbiased* estimate of the asset's true expected return η in equation (60). Therefore, regardless of N, multifactor models for an asset's expected return are incorporated into information portfolio theory when the individual factor returns represent distinct state variables. The return forecasts μ_j associated with each factor in equation (63) replace the single expected return estimate arising from the multifactor model in equation (9).

Observe that the $\alpha_{0,j}$ and $\alpha_{1,j}$ coefficients are calibrated using a time series regression, along with the β_0 and β_j coefficients for $j = 1, \ldots, N$. However, the information portfolio weights differ from the α and β coefficients since the W vector sums to one and is derived from the variances and covariances of the ϵ^j errors in equation (63) which are computed using equations (2) and (3) respectively.

In practice, the number of return forecasts J would exceed the number of factors N since their ex-ante returns are random and an individual asset's beta coefficients also require estimation. These sources of uncertainty illustrate the generality of information portfolio theory which is not restricted to a single return forecast for each asset. For example, if the Fama-French (1993) model estimates expected returns, then information sources can disagree on the return prospects for small versus large market capitalization stocks, value versus growth stocks as well as the overall market. Thus, the expected return uncertainty σ^2_μ in equation (20) is larger than the dispersion across the three elements of equation (61) which assumes the random factor returns next period are known along with the asset's respective factor loadings. This property applies to any specification for η and allows the state variables to represent industry and macroeconomic trends as well as firm-specific earnings forecasts.

[33]When the factors are orthogonal, estimates for the β_j coefficients in equation (9) equal the $\alpha_{1,j}$ regression estimates in equation (62), which reduces equation (64) to $\alpha_{0,j} = \beta_0 + \sum_{k=1}^{3} \beta_k \, f_k$ for $k \neq j$ by eliminating the $\left[\beta_j - \alpha_{1,j}\right] \, f_j$ term.

C Covariances and the Perceived Return

The partial derivative of the investor's perceived return in equation (26) with respect to σ_{12} equals

$$\begin{aligned}\frac{\partial \text{Perceived Return}}{\partial \sigma_{12}} &= \frac{-\left(\mu_1+\mu_2\right)\left[\sigma_1^2+\sigma_2^2-2\sigma_{12}\right]+2\left[\sigma_2^2\,\mu_1+\sigma_1^2\,\mu_2-\sigma_{12}\,\left(\mu_1+\mu_2\right)\right]}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2} \\ &= \frac{\left(\mu_2-\mu_1\right)\,\left(\sigma_1^2-\sigma_2^2\right)}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}\,. \end{aligned} \tag{65}$$

The sign of this derivative may be either positive or negative. According to the numerator of equation (65), when either the return forecasts or their estimated accuracies are identical, the investor's perceived return is invariant to σ_{12}.

The partial derivative of the perceived return's aggregate mean-squared forecast error in equation (27) with respect to σ_{12} equals

$$\begin{aligned}\frac{\partial \text{Aggregate Forecast Error of Perceived Return}}{\partial \sigma_{12}} &= \frac{-2\sigma_{12}\left[\sigma_1^2+\sigma_2^2-2\sigma_{12}\right]+2\left[\sigma_1^2\sigma_2^2-(\sigma_{12})^2\right]}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2} \\ &= \frac{2\sigma_{12}\left[\sigma_{12}-(\sigma_1^2+\sigma_2^2)\right]+2\sigma_1^2\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2-2\sigma_{12}\right)^2}\,. \end{aligned} \tag{66}$$

When σ_{12} is negative, equation (66) is large and positive which implies the aggregate forecast error surrounding an asset's perceived return is reduced as σ_{12} becomes more negative.

D Proof of Proposition 3

Recall from equation (15) that the asset's ex-ante return is distributed $\mathcal{N}\left(W^T\mu, W^T\Theta W+\nu\right)$ under the distributional assumption in equation (16). To prove Proposition 3, the following utility maximization problem is solved

$$\begin{aligned} & \max_f E\left\{U\left[M\,\left((1-f)\,(1+r_f)+f\,\left(1+W^T\mu\right)\right)\right]\right\} \\ = & \max_f -E\left[\exp\left\{-\gamma\,M\left(1+r_f+f\,\left(W^T\mu-r_f\right)\right)\right\}\right] \\ = & \max_f -\exp\left\{-\gamma\,M\,f\,W^T\mu+\gamma\,M\,f\,r_f+\frac{\gamma^2\,M^2\,f^2}{2}\,\left[\nu+W^T\Theta W\right]\right\}\,, \end{aligned} \tag{67}$$

where the last equality results from the moment generating function of a normal distribution. This maximization involves setting the partial derivative of equation (67) with respect to f

$$\left(-\gamma\,M\,W^T\mu+\gamma\,M\,r_f+\gamma^2\,M^2\,f\,\left[\nu+W^T\Theta W\right]\right)\left(-e^{\left\{-\gamma\,M\,f\,W^T\mu+\gamma\,M\,f\,r_f+\frac{\gamma^2\,M^2\,f^2}{2}\,\left[\nu+W^T\Theta W\right]\right\}}\right) \tag{68}$$

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

$$-\gamma\, M\, W^T\mu + \gamma\, M\, r_f + \gamma^2\, M^2\, f\, \left[\nu + W^T\Theta W\right] \;=\; 0\,. \tag{69}$$

Therefore, the optimal investment in the risky asset equals

$$f \;=\; \frac{W^T\mu - r_f}{\gamma\, M\, \left[\nu + W^T\Theta W\right]}\,, \tag{70}$$

which becomes

$$\begin{aligned} f \;&=\; \frac{1}{\gamma\, M}\frac{\left(\frac{\mathbf{1}^T\Theta^{-1}\mu}{\mathbf{1}^T\Theta^{-1}\mathbf{1}} - r_f\right)}{\nu + \frac{1}{\mathbf{1}^T\Theta^{-1}\mathbf{1}}} \\ &=\; \frac{\mathbf{1}^T\Theta^{-1}\left(\mu - r_f\mathbf{1}\right)}{\gamma\, M\, \left[1 + \nu\mathbf{1}^T\Theta^{-1}\mathbf{1}\right]}\,, \end{aligned} \tag{71}$$

after substituting in the results of Proposition 2.

Observe that the optimal portfolio weights from Proposition 1 transform equation (70) into equation (71). When the asset's true expected return is known, equation (70) implies $f = \frac{\eta - r_f}{\gamma\, M\, \nu}$ since $W^T\mu = \eta$ and $W^T\Theta W = 0$.

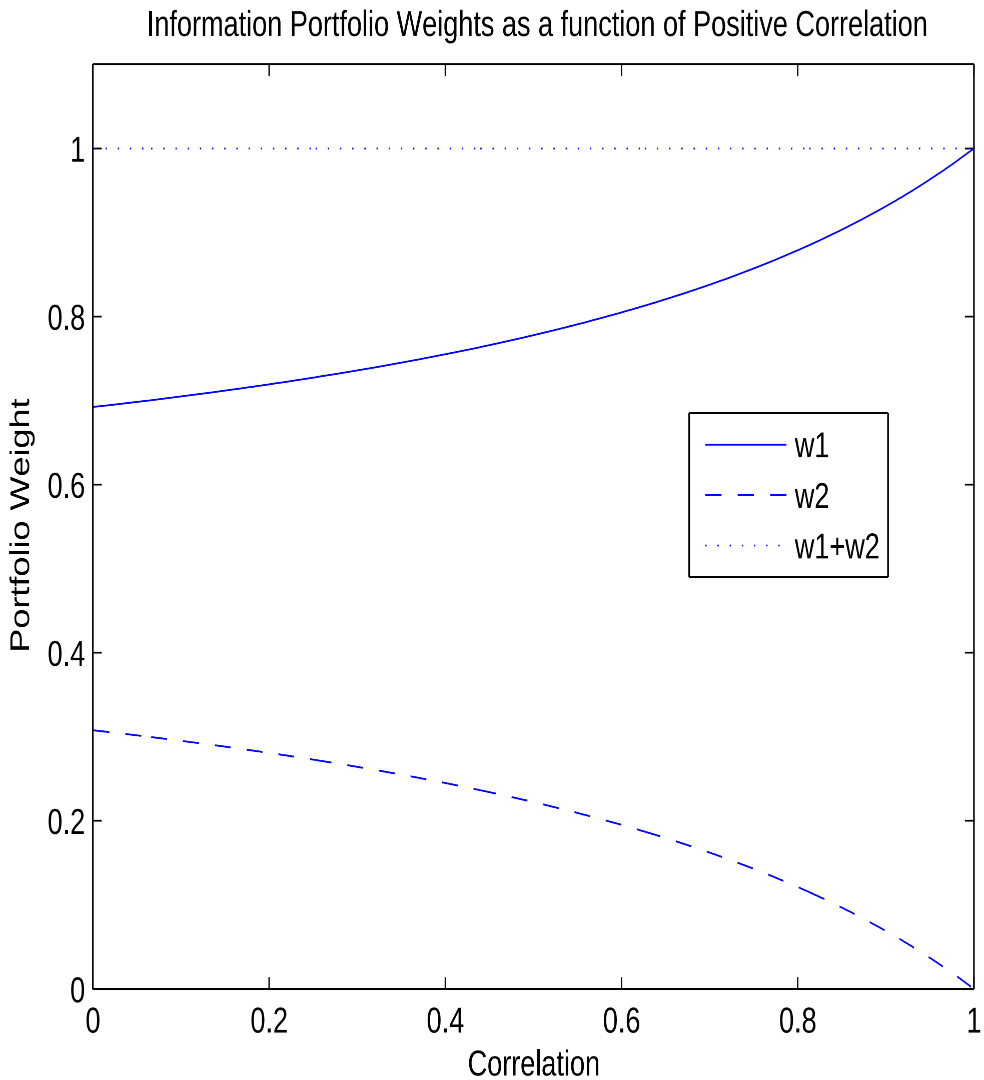

Figure 1: Impact of correlation between the return forecasts of two information sources on their corresponding information portfolio weights in equation (44). The above plot utilizes the following parameter values; $\mu_1 = 0.07$, $\mu_2 = 0.10$, $\sigma_1 = 0.40$ and $\sigma_2 = 0.60$. Therefore, the first information source is more accurate than the second.

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***