Dynamic Forecasting Behavior by Analysts: Theory and Evidence

Jonathan Clarke

Ajay Subramanian

DuPree College of Management
Georgia Institute of Technology
755, Ferst Drive,
Atlanta, GA 30332

March 2003

Dynamic Forecasting Behavior by Analysts: Theory and Evidence

Abstract

We examine the dynamic forecasting behavior of investment analysts in response to their prior performance relative to their peers within a continuous time/multi-period framework. Our model predicts a U-shaped relationship between the *boldness* of an analyst's forecast, that is, the deviation of her forecast from the consensus and her prior relative performance. In other words, analysts who significantly out perform or under perform their peers issue bolder forecasts than intermediate performers. We then test these predictions of our model on observed analyst forecast data. Consistent with our theoretical predictions, we document an approximately U-shaped relationship between analysts' prior relative performance and the deviation of their forecasts from the consensus. Our theory examines the impact of both *explicit incentives* in the form of compensation structures and *implicit incentives* in the form of career concerns, on the dynamic forecasting behavior of analysts. Consistent with existing empirical evidence, our results imply that analysts who face greater employment risk (that is, the risk of being fired for poor performance) have greater incentives to herd, that is, issue forecasts that deviate less from the consensus. Our multi-period model allows us to examine the *dynamic* forecasting behavior of analysts in contrast with the extant two-period models that are *static* in nature. Moreover, the model also differs significantly from existing theoretical models in that it does not rely on any specific assumptions regarding the existence of asymmetric information and/or differential analyst abilities.

Key Words: Analysts, Career Concerns, Dynamic Forecasting, Herding

1. Introduction

We examine the dynamic forecasting behavior of investment analysts in response to their prior performance relative to their peers within a continuous time/multi-period framework. Our model predicts a U-shaped relationship between the boldness of an analyst's forecast, that is, the deviation of her forecast from the consensus, and her prior relative performance. In other words, analysts who significantly out-perform or under-perform their peers issue bolder forecasts than intermediate performers. We then test these predictions of our model on observed analyst forecast data. Consistent with our theoretical predictions, we document an approximately U-shaped relationship between deviation of analysts' forecasts from the consensus and their prior relative performance.

Our model relies on two assumptions: an analyst's compensation is convex in her performance relative to her peers and faces significant negative career concerns, that is, the risk of losing her job for poor relative performance. These assumptions are consistent with existing empirical and anecdotal evidence. Although data on analysts' compensation is not available, Wise (2000) notes that there are large pay discrepancies between analysts being named to *Institutional Investor's* All-America research team and those that do not make the team. Given that Leone and Wu (2002) find that *Institutional Investor* all-stars have superior performance to non-star analysts, this would imply convexity in the compensation structures for analysts. Mikhail, Walther, and Willis (1999) find that an analyst is more likely to be fired if her forecast accuracy declines relative to her peers. However, they find no association between an analyst's probability of turnover and absolute forecast error. In other words, this suggests that it is relative, rather than absolute performance, that matters.

The model we propose and investigate can be briefly described as follows. At each forecasting date, an analyst faces the choice between a *bold* strategy and a *conservative* strategy.

The analyst's forecast deviates to a greater extent from the consensus forecast under the bold strategy than the conservative strategy. The bold and conservative strategies may result from private signals received by the analyst or from publicly available information about the firm being covered. Under either strategy, the analyst has a nonzero probability of either outperforming or under performing the median analyst. The analyst is compensated at the end of each forecasting period and her compensation is a convex function of her prior relative performance over some time horizon. There is a nonzero probability that the analyst may be fired if her relative performance is below an exogenous level. The analyst has inter-temporal preferences for the periodic cash flows that comprise her compensation and dynamically chooses the bold or conservative strategy at each date to maximize her expected utility. We assume that the analyst has linear preferences purely for simplicity. Our analysis can easily be generalized to incorporate risk aversion.

We explicitly solve the analyst's stochastic dynamic optimization problem and show that the optimal policy for the analyst can be described as follows: there exist two thresholds of prior relative performance such that if the analyst either outperforms the higher threshold or under performs the lower threshold, she chooses the bold strategy and if her prior relative performance lies between the thresholds, she chooses the conservative strategy. Therefore, the analyst makes a bold forecast if she either significantly out-performs or under-performs the median analyst and makes a conservative forecast if she is an intermediate performer.

The intuition for these results is the following. When the analyst significantly outperforms the median analyst, the convexity of her compensation structure induces her to take on the increased risk of issuing a bold forecast. On the other hand, if she significantly under performs the median, she faces a substantial risk of being fired for poor relative performance. Therefore, she takes on the increased risk of the bold strategy to increase the probability that her performance will rise above

the level where she may be fired. At intermediate levels of prior relative performance, the analyst trades off the higher expected compensation from choosing the bold strategy for the lower probability that her performance may decline below the level where she may be fired from choosing the conservative strategy. In general, there exists a nonempty intermediate region of prior relative performance where she prefers the conservative strategy.

We also show that if the analyst faces little or no risk of being fired, she will always choose the bold strategy. This result follows easily from the convexity of her compensation structure. Our results have immediate implications for *herding* by analysts. In particular, they show that intermediate performers tend to deviate less from the consensus, that is, herd more than significant out-performers and under-performers. Moreover, our result that an analyst always chooses the bold strategy if she faces little or no employment risk indicates that the *incentive to herd* increases with employment risk. In particular, this implies that more experienced analysts who face lower employment risk herd less than less experienced analysts. This prediction is consistent with the empirical results of Hong, Kubik, and Solomon (2000) who examine a sample of stock analysts' earnings forecasts and show that younger analysts tend to herd more than older, more experienced analysts who face lower employment risk.[1]

Our results therefore highlight the impact of both *explicit incentives* in the form of compensation structures and *implicit incentives* in the form of career concerns on the forecasting behavior of investment analysts. Several recent papers have emphasized the importance of career concerns in various contexts. Scharfstein and Stein (1990), Prendergast and Stole (1996) and Morris (1997) argue that career concerns may induce corporate and/or fund managers to ignore private information and follow the herd or avoid following it when their actions are observable. In Scharfstein and Stein (1990), "smart" managers receive correlated information, while "dumb"

---

[1] Chevalier and Ellison (1999) obtain similar results in their empirical investigation of mutual fund managers.

managers receive uncorrelated noise. Thus, a manager who learns that his private information differs from that of another manager believes that he is more likely to be "dumb", ignores his information and "herds".[2] In Prendergast and Stole (1996), managers have private information about the precision of their information. A bolder action signals that a young manager knows his information to be good, and hence young managers have an incentive to take excessively bold actions. Older managers have an incentive to become jaded and do not change their actions a great deal from period to period. Zwiebel (1995) proposes a model of the behavior of corporate managers where taking an unobserved unconventional action increases the variance of the market's ex post assessment of a manager's ability. In a result that is reminiscent of the results that we obtain, Zwiebel (1995) shows that average managers prefer the conventional action because it reduces the risk of their being fired, while high or low ability managers may prefer unconventional actions.

Our paper differs significantly from the above papers in that we focus on investigating the *dynamic* forecasting behavior of analysts in a multi-period/continuous time framework, rather than the two-period models examined in the above papers. Welch (2000), in particular, notes that these existing *static* theories are "designed to explain a steady state in which all analysts herd perfectly, not to explain an ever varying *time-series* of recommendations or a residual difference in opinions across analysts." (pg. 370).   Our dynamic model attempts to address some of these concerns. Moreover, in contrast with the above papers, we do not make any specific assumptions about whether analysts possess different abilities and whether they receive private signals. Therefore, our results do not rely on the existence of asymmetric information and/or differential abilities, but depend only the convexity of compensation in relative performance and the existence of significant career concerns.

---

[2] Graham (1999) considers a model patterned after Scharfstein and Stein (1990) to investigate the herding behavior of investment newsletters. Trueman (1994) also proposes a model that demonstrates herding by analysts.

Using data from the IBES history tapes on individual analysts' forecasts between 1988 and 2000, we test the predictions of our model. Specifically, each quarter we rank analysts into deciles based on past forecasting performance calculated over varying time horizons that range from 1 year to 3 years. For each analyst, we calculate boldness as the normalized deviation from the consensus as in Hong, Kubik, and Solomon (2000). We then compute median *future* boldness for each decile of *past* relative performance. Consistent with our predictions, we find evidence of a U-shaped relation between future boldness and past performance. It is possible that this U-shaped relation is a result of analysts persisting in choosing bold or conservative strategies and not due to dynamic alteration of strategies by analysts. In other words, an analyst who is always bold is more likely to take on extreme positions in relative performance. We control for this possibility in our empirical tests, and find that the U-shaped pattern still holds. As an alternative methodology, we test for the U-shaped relation using the Fama-MacBeth (1973) regression methodology. These tests lead to qualitatively similar conclusions.

Our empirical results contribute to the existing literature in that we investigate the forecasting behavior of all analysts in response to their prior performance and document, for the first time, a *non-monotonic U-shaped relationship* between the boldness of an analyst's forecast and her prior relative performance. Hilary and Menzly (2002) also empirically investigate the relationship between boldness and past performance. Consistent with our findings, they report that analysts who have performed well in the past tend to issue bolder forecasts. However, they do not document that analysts who have performed poorly in the past also issue bolder forecasts. They propose a behavioral model based on overconfidence to explain their findings. Our model offers a rational explanation for bold forecasting behavior by out-performers and under-performers. Hong et al (2002) empirically examine the effect of career concerns on boldness and find that more

experienced analysts are more likely to deviate from the consensus. From an empirical standpoint,

we complement their findings by examining and documenting the effect of prior performance on

analysts' incentives to deviate from the consensus.[3]

The remainder of the paper proceeds as follow. In **Section 2** we present the model. In

Section 3, we state our main results regarding the optimal forecasting behavior of analysts. In

Section 4, we present the results of empirical tests of our predictions. **Section 5** concludes the

paper. All detailed proofs are relegated to the **Appendix.**

2. The Model

The primary focus of this paper is the dynamic forecasting behavior of an analyst in response to

his prior performance relative to his peers. We consider an infinite horizon, multi-period

framework. The set of forecasting dates is characterized by the set

$\Gamma \equiv \{....,-3\Delta,-2\Delta,-\Delta,0,\Delta,2\Delta,3\Delta,...\}$. The set of forecasting dates is a doubly infinite set to

emphasize the point that there is no "initial date" in our framework. The fundamental economic

variable we model is the cumulative forecasting performance, that is, the sum of the forecasting

errors, for an analyst over a time horizon T . T is the exogenously specified horizon over which the

analyst's performance is evaluated. The actual value of T does not play an important role in our

analysis and it may also differ across analysts.[4] If $e(t)$ denotes the cumulative forecasting error of

an analyst at date $t \in \Gamma$, and $e_m(t)$ denotes the median cumulative forecasting error for all analysts

over the same time horizon t, then

(1) $q(t) = e(t) - e_m(t)$

---

denotes the relative forecasting performance of the analyst at date t. $q(.)$ is a stochastic process

with support in $(-\infty, \infty)$. The median forecasting error is used as a benchmark purely for

concreteness. Our theoretical results do not depend on the specific benchmark used. At any date t,

we assume that an analyst has the choice between adopting a *bold* strategy or a *conservative*

strategy. If he adopts the bold (conservative) strategy, then the change in his relative forecasting

performance over the next period Δ is a normally distributed random variable with mean μ and

standard deviation $\sigma_1(\sigma_2)$ with $\sigma_1 > \sigma_2$. Therefore,

(2)
$$q(t+\Delta) - q(t) = \mu\Delta + \sigma_1 N$$
$$q(t+\Delta) - q(t) = \mu\Delta + \sigma_2 N$$

under the bold and conservative strategies 1 and 2 respectively. In the above, N is a standard

normal random variable.[5]

It is important to emphasize here that the bold and conservative strategies may represent

either private or publicly observable signals. In other words, our theoretical framework does *not*

rely on any specific assumptions regarding the observability of the analysts' signals on the basis of

which they make their forecasts. Therefore, *asymmetric information* is *not* a crucial ingredient of

our model although our model is certainly consistent with it. Further, we do not make any specific

assumptions regarding the abilities of analysts, that is, they may all possess the same ability or have

different abilities. The fact that neither asymmetric information nor differential analyst abilities

plays a crucial role in our theoretical analysis makes it significantly different from earlier theoretical

frameworks that have been applied to examine analyst behavior.[6] In all these frameworks,

asymmetric information and/or differential analyst abilities play important roles. In summary, the

nature of the information that analysts possess and the analysts' forecasting abilities do not affect

---

[5] Since the analyst can only choose either the bold or the conservative strategy at any date, we use the same notation for the standard normal random variable without loss of generality.
[6] See, for example, Scharfstein and Stein (1991), Zwiebel (1995), Prendergast and Stole (1996), and Graham (1999)

our model as well as our theoretical results. Ours is a simple rational model of forecasting strategy choices by analysts.

There is an exogenously specified level of relative performance q_b such that if $q(.)$ exceeds q_b, the analyst is fired with some probability $\alpha\Delta \in (0,1)$.[7] We also assume that the analyst bears significant personal costs from being fired. These proportional costs are described by the parameter $\delta \in [0,1]$, that is, the analyst loses a proportion δ of her expected future compensation is he is fired.

The analyst is assumed to be risk-neutral and his compensation is assumed to be convex in his relative performance. The risk-neutrality of the analyst is assumed purely for simplicity of exposition and does not affect our results qualitatively. For analytical tractability, we assume that the compensation of the analyst at date t is given by

(3) $$C(t) = \Delta[g\exp(-cq(t))] \text{ where } g > 0, c > 0.$$

For subsequent notational simplicity, we normalize g to 1. From (3), we note that as the analyst's performance varies from being very good to very bad, $q(t)$ varies from being significantly negative to becoming significantly positive. The goal of the risk-neutral analyst is to choose his forecasting policy ξ representing his choice between a bold and conservative strategy at every forecasting date, in order to maximize his discounted expected compensation. At any date t, his optimal future forecasting policy should therefore solve

(4)
$$v(q(t)) = \sup_\xi E[(1-\delta)v(q(\tau_b)) + \sum_{s=t}^{\tau_b} e^{-\beta s} C_\xi(s)] =$$
$$\sup_\xi E[(1-\delta)v(q(\tau_b)) + \sum_{s=t}^{\tau_b} e^{-\beta s} \Delta(\exp(-cq_\xi(s)))]$$

---

[7] The assumption that the analyst is not fired with certainty is consistent with the empirical findings of Hong, Kubik, and Solomon (2000). They find that the probability of an analyst getting fired increases with underperformance , but it is significantly different from one.

In the above, τ_b denotes the random date at which the analyst is fired. Notice that the analyst's payoff when he is fired is a proportional of his optimal value function $v(q(\tau_b))$ that is his maximum expected future compensation. We have incorporated the fact that the analyst's optimal policies are clearly stationary, that is, his decision at any date depends only on his relative performance at that date. The subscripts on q_ξ, C_ξ denote the dependence of the analyst's relative performance and compensation on the forecasting policy he chooses. β is the analyst's constant discount rate for future cash flows. We can use standard dynamic programming techniques to obtain the following Hamilton-Jacobi-Bellman equation for the analyst's optimal value function:

$$v(q) = \sup_{i=1,2} e^{-\beta\Delta} E[v(q^i(\Delta)) + \Delta\exp(-cq^i(\Delta)) - 1_{q^i(\Delta)>q_b} 1_{fired=1} \delta v(q^i(\Delta))]$$

In the above, $q^i(\Delta)$ represents the analyst's relative performance at the end of the next period if he follows strategy i and his relative performance at the beginning of the period is q. The first term in the expectation is the optimal value function at the end of the period, the second term is the analyst's compensation for the period, and the third term represents the proportional personal costs he pays if he is fired due to his relative performance being above the threshold level q_b. The subscript $fired = 1$ represents the event that the analyst is fired. Incorporating the exogenous probability $\alpha\Delta$ that the analyst is fired when his relative performance is above the threshold, we obtain

$$v(q) = \sup_{i=1,2} e^{-\beta\Delta} E[v(q^i(\Delta)) + \Delta\exp(-cq^i(\Delta)) - 1_{q^i(\Delta)>q_b} \alpha\delta\Delta v(q^i(\Delta))]$$

The Continuous Time Framework

For analytical tractability, we now make the standard continuous time approximation that is valid when the time period between successive dates Δ is small relative to the time period over which the analyst's forecasting behavior is being investigated. In this case, equation (2) for the evolution of the analyst's relative performance is replaced by

(5) $dq(t) = \mu dt + \sigma_i dB(t)$ under strategy $i \in \{1,2\}$ where $B(.)$ is a Brownian motion.

The analyst's objective in (4) is replaced by

(6)
$$v(q(t)) = \sup_\xi E[(1-\delta)v(q(\tau_b)) + \int_t^{\tau_b} ds e^{-\beta s} C_\xi(s)] =$$
$$\sup_\xi E[(1-\delta)v(q(\tau_b)) + \int_t^{\tau_b} ds e^{-\beta s} (\exp(-cq_\xi(s)))]$$

For subsequent analytical and notational convenience, we re-define the analyst's optimization problem in terms of the process $p(t) = \exp(-q(t))$. Using Ito's lemma, the evolution of the process $p(.)$ under strategy $i \in \{1,2\}$ is given by

(7) $dp(t) = p(t)[(\frac{\sigma_i^2}{2} - \mu)dt + \sigma_i dB(t)]$

and the analyst's objective is to choose his forecasting policy to maximize

(8) $u(p(t)) = \sup_\xi E[(1-\delta)u(p(\tau_b)) + \int_t^{\tau_b} ds e^{-\beta s} (p_\xi(s)^c)]$

where $u(.)$ is the analyst's optimal value function as a function of $p = e^{-q}$. Notice now that low (high) values of $p(.)$ represent good (bad) relative performances by the analyst. In order to ensure that the analyst's value function is defined, we assume that $\beta > \mu_i$, $i = 1,2$.

If $u(.)$ is the optimal value function of the dynamic optimization problem (8), then, as mentioned earlier, the analyst's optimal policy is stationary, that is, his choice of strategy at any

date is not explicitly dependent on time and depends only on the current measure of his relative performance p. We denote the analyst's optimal value function by $u(p)$. We may use traditional dynamic programming arguments analogous to those used in deriving the discrete-time Hamilton-Jacobi-Bellman equation to write down the following formal Hamilton-Jacobi-Bellman equation for u:

$$0 = \sup_{i=1,2} [1_{p>p_b}(-\beta u) + 1_{p \le p_b}(-(\beta + \alpha\delta)u) + p^c + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp}]$$

where $\mu_i = \frac{1}{2}\sigma_i^2 - \mu$. In the dynamic programming framework, the variable p above represents the value of the state variable $P(.)$ so that the term p^c is the instantaneous rate of compensation of the analyst.

Hence, in regions where strategy i is optimal, the value function $u(p)$ must satisfy the system of ordinary differential equations:

$$-\beta u + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp} + p^c = 0; \ p > p_b$$

$$-(\beta + \alpha\delta)u + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp} + p^c = 0; \ p < p_b$$

It can be shown that the general solution to the ordinary differential equations has the form:

(9)
$$u(p) = A p^{\eta_i^+} + B p^{\eta_i^-} - \frac{p^c}{\frac{1}{2}\sigma_i^2 c^2 + (\mu_i - \frac{1}{2}\sigma_i^2)c - \beta}; \ p > p_b$$

$$= C p^{\rho_i^+} + D p^{\rho_i^-} - \frac{p^c}{\frac{1}{2}\sigma_i^2 c^2 + (\mu_i - \frac{1}{2}\sigma_i^2)c - (\beta + \alpha\delta)}; \ p < p_b$$

where η_i^+, η_i^- and ρ_i^+, ρ_i^- are the positive and negative roots respectively of the quadratic equations:

(10)
$$\frac{1}{2}\sigma_i^2 x^2 + (\mu_i - \frac{1}{2}\sigma_i^2)x - \beta = 0$$

$$\frac{1}{2}\sigma_i^2 x^2 + (\mu_i - \frac{1}{2}\sigma_i^2)x - (\beta + \alpha\delta) = 0$$

In order to ensure that the analyst's value function exists, we assume that her discount rate $\beta > \mu_i$

and is high enough so that

(11) $c < \eta_i^+, c < \rho_i^+$.

The following lemma collects properties of the roots $\eta_i^+, \eta_i^-, \rho_i^+, \rho_i^-$ that will be used frequently.

Lemma 1

a) $\eta_1^+ < \eta_2^+, \rho_1^+ < \rho_2^+, \eta_1^+ < \rho_1^+, \eta_2^+ < \rho_2^+$

Proof. In the Appendix.

We now state (without proof) the following well-known verification theorem for the

analyst's optimal value function.

Proposition 1: *Suppose u is a function that is differentiable on $(0, \infty)$ and twice differentiable on*

$(0, \infty) \setminus \{p_b\}$ *satisfying the HJB equation*

$$0 = \sup_{i=1,2}[1_{p > p_b}(-\beta u) + 1_{p \le p_b}(-(\beta + \alpha\delta)u) + p^c + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp}]$$

and $\lim_{p \to \infty} u(p)/p^c < \infty$ (no bubbles condition). Then u is the analyst's optimal value function.

Proof. See Karatzas and Shreve [1998].

This completes the formulation of the model and the mathematical preliminaries.

3. The Analyst's Optimal Forecasting Policy

In this section, we explicitly derive the optimal forecasting policies for the analyst for all

possible pairs of bold and conservative strategies 1 and 2 characterized by the volatility parameters

(σ_1, σ_2) with $\sigma_1 > \sigma_2$ and his discount rate β is high enough to ensure that condition (11) is satisfied. We show that there exist two levels p_l, p_h with $p_l \leq p_h$ of prior relative performance as measured by the process $p(.)$ such that it is optimal for the analyst to choose the conservative strategy when his prior performance lies between the thresholds and switch to the bold strategy above the threshold p_h and below the threshold p_l. Therefore, the analyst chooses the bold strategy when he is either a significant out performer or under performer and the conservative strategy when he is an intermediate performer. We may have $p_l = p_h$ in which case the analyst always chooses the bold strategy and the switching of strategies is sub-optimal. We provide a necessary and sufficient condition for the switching of strategies to be optimal for the analyst.

The intuition underlying these results is the following. If the analyst is a significant out performer, the risk of his getting fired is very low. Therefore, he chooses the bold strategy since it increases his expected compensation. On the other hand, if the analyst significantly under performs the threshold p_b where he may be fired, he faces significant risk of being fired. He therefore, chooses the bold strategy to maximize the probability that his performance may increase above the level below p_b. At intermediate levels of performance, the analyst trades off the higher expected compensation from choosing the bold strategy for the lower employment risk from choosing the conservative strategy. In general, there exists an intermediate region of relative performance where the analyst prefers the conservative strategy. We now proceed to formalize this intuition.

Consider the class of policies defined by the trigger r with $p_b \leq r$ where the analyst always chooses strategy 1 for $p \leq p_b, p \geq r$ and strategy 2 for $p_b < p < r$. It can be shown that the value function u_r of such a policy has the following functional form:

$$(12) \quad \begin{aligned} u_r &= A_r p^{\rho_1^+} - \frac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - (\beta + \alpha\delta)}; p \le p_b \\[2ex] &= B_r p^{\eta_2^+} + C_r p^{\eta_2^-} - \frac{p^c}{\frac{1}{2}\sigma_2^2 c^2 + (\mu_2 - \frac{1}{2}\sigma_2^2)c - \beta}; p_b < p \le r \\[2ex] &= D_r p^{\eta_1^-} - \frac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - \beta}; r < p \end{aligned}$$

where the coefficients are determined by continuity and differentiability conditions at the points

p_b, r and their dependence on the trigger r is explicitly indicated. Define the differential operators

$L_i, L_i'; i = 1,2$ as follows:

$$L_i u = -\beta u + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp}; \quad L_i' u = -(\beta + \alpha\delta)u + \mu_i p u_p + \frac{1}{2}\sigma_i^2 p^2 u_{pp}$$

The following propositions completely characterize the optimal policies for the analyst.

Proposition 2: *Suppose*

$$(13) \quad L_2(u_{p_b}) + p^c \, |_{p=p_b+} > 0$$

There exist a threshold level of relative performance $p_* > p_b$ such that if

a) $p \le p_b, p \ge p_*$, it is optimal for the analyst to choose the bold strategy 1 and if

b) $p_b < p < p_*$, it is optimal for the analyst to choose the conservative strategy 2.

Proof. In the Appendix.

Condition (13) of the proposition is therefore a sufficient condition for the analyst to choose the

conservative strategy 2 in some region of prior relative performance. Intuitively, the condition

expresses the fact that the employment risk of the analyst and the difference between the risks of the

bold and conservative strategies are high enough to ensure that it is optimal for the analyst to choose

the conservative strategy when his prior performance is "close" to the threshold p_b. The following proposition shows that condition (13) is also necessary for the optimality of switching to the conservative strategy in some region of prior performance.

Proposition 3. *Suppose*

(14) $L_2(u_{p_b}) + p^c \mid_{p=p_b+} \leq 0$

Then the optimal policy for the analyst is to always choose the bold strategy.

Proof. In the Appendix.

If condition (14) is satisfied, the employment risk of the analyst and/or the difference in the risks of the two strategies is low enough that it is sub-optimal for the analyst to deviate from the bold strategy.

The results of the above propositions imply that it is either always optimal for the analyst to choose strategy 1 or there exists a non-empty intermediate region $[p_b, p^*]$ of prior performance where the analyst optimally chooses strategy 2. If the analyst chooses strategy 2, he deviates less from the median analyst than if he chooses strategy 1. Suppose now that strategy 1 represents a private noisy signal for the analyst and strategy 2 represents the *herding* strategy. In this setting, our results imply that when the analyst is a significant out performer or under performer, he trusts his private signal whereas if he is an intermediate performer, he chooses to *herd*. We would like to emphasize again that our model does not make any specific assumptions about whether the analyst receives private signals. Hence, our result that significant out performers and under performers deviate more significantly in their forecasts from the median than intermediate performers does *not* depend on whether there is asymmetric information and/or the analysts possess different forecasting abilities. Our model and results depend on two main assumptions: an analyst's compensation is convex in her relative performance and faces significant career concerns.

4. Empirical Results

Data Description:

In order to test the above propositions on past forecast accuracy and boldness, we collect forecasts from the IBES Detailed History database over the period 1988 to 2000. The Detailed History database tracks the identity of the analyst issuing the forecast, her employer, the date of the forecast, and the actual value of her forecast. This dataset also allows us to identify and track each analyst across time, even if they switch investment.

In order to construct our measures of forecast accuracy and forecast boldness, we use the ranking methodology introduced in Hong, Kubik, and Solomon (2000). The procedure is as follows. We use the I/B/E/S data to construct a quarterly performance measure based on an analyst's forecast accuracy. We define $F_{i,j,t}$ as the most recent earning-per-share forecast of quarterly earnings issued by analyst i on stock j in quarter t. Our measure of analyst i's accuracy for firm j in year t is the absolute difference between her forecast and the realized earnings-per-share of the firm, $A_{j,t}$:

$$forecast \quad error = \left| F_{i,j,t} - A_{j,t} \right|$$

We then sort the analysts who cover a firm in a quarter based on their forecast errors given above. We then assign a ranking based on this sorting: the best analyst receives a rank of one, the second best analyst receives a rank of two, and so on. In the case of ties, we assign each analyst the midpoint value of the ranks that they take up. Since the maximum rank an analyst can receive for a firm depends on the number of analysts who cover the firm, we scale an analyst's rank by the number of analysts who cover the firm. The formula for this score measure is:

$$accuracy \quad score_{i,j,t} = 100 - \left[\frac{rank_{i,j,t} - 1}{number \quad analysts_{j,t} - 1} \right] \times 100,$$

where number of analysts$_{j,t}$ is the number of analysts who cover the firm in a given quarter.[8] We

then calculate the average score for each analyst over the previous four, eight, and 12 quarters.

Higher overall scores correspond to better analyst performance.

We use a similar procedure to construct a measure of an analyst's forecast boldness. Let

$$\overline{F}_{-i,j,t} = \frac{1}{n}\sum_{m\in -i} F_{m,j,t} \text{ , where } -i \text{ is the set of all analysts other than analyst } i \text{ who produce an}$$

earnings estimate for stock j in year t, and n is the number of analysts in $-i$. Hence, $\overline{F}_{-i,j,t}$ is a

measure of the consensus forecast made by all other analysts except analyst i following stock j in

quarter t.

$$boldness_{i,j,t} = \left| F_{i,j,t} - \overline{F}_{-i,j,t} \right|$$

We then replicate the previous ranking methodology for constructing the analyst accuracy score as
in the previous subsection.

Empirical Findings

Table 1 presents summary statistics for our sample of analysts' forecasts. Our sample contains a

large number of analysts from a number of different investment firms. The average number of

unique investment firms each quarter is 366.37, while the average number of unique analysts

issuing forecasts each quarter is 1,763.67. The average analyst in our sample issues quarterly

earnings-per-share estimates for 7.32 stocks.[9] The average stock in a sample has 5.56 different

analysts providing coverage. Note that we require at least two analysts to be covering the stock in

order to calculate our boldness and performance rankings. The average analyst in our sample has

---

[8] For example, the lowest rated analyst for each firm would receive a score of zero, while the highest rated analyst
would receive a score of 100.
[9] The maximum number of stocks covered by an analyst is 94. This could be attributed to team of analysts rather than
an individual.

4.46 years of experience. Finally, average analyst boldness and average analyst accuracy average

50.38 and 50.46. By construction, the median values of both of these variables is 50.00.

Table 2 examines whether analysts actively move between conservative and bold

forecasting strategies. The table shows a transition matrix relating average past boldness to future

boldness. The table is constructed as follows. For each quarter in our sample, we calculate the

average boldness score for each analyst over the previous four quarters and then divide analysts into

deciles based on this score. We similarly rank analyst into boldness deciles based on their current

boldness score. The results suggest that analysts actively move between bold and conservative

strategies. For example, of the analysts ranked into the lowest past boldness decile, only 15.93%

pursue the least bold strategy in the subsequent period. Similarly, of the analyst ranked into the

highest boldness decile based on past performance, only 18.02% fall into the same decile in the

subsequent period. It's worth noting that the p-value from a χ^2 test indicates that we can reject the

null hypothesis of equal proportions within each past boldness decile.

Table 3 presents our results on the relation between past performance and future boldness.

Each quarter, we rank analysts into deciles based on their average past performance. The average

past performance of each analyst is calculated using accuracy scores over the previous four, eight,

and twelve quarters. We then compute mean future boldness for each of these deciles. In order to

test for a U-shaped relation, we computer the average difference in boldness between past

performance deciles 5 and 6 and then test to see whether mean boldness in each decile is different

from this value. Panel A presents our findings for the case where past performance is calculated

over the previous four quarters. The results confirm the existence of a U-shaped relation. The

worst past performance decile has future boldness of 51.06, which is statistically significant.

Similarly the best past performance decile has future boldness of 50.61, which is statistically

significant. The results in Panel B and Panel C show that the results are not affected by the choice of the time horizon over which past forecast accuracy is calculate. Similar results obtain if we calculate average past performance using the previous eight quarters or the previous twelve quarters.

In Table 4, we control for the past boldness of the analyst. It is possible that the U-shaped relation documented in Table 3 is a result of analysts persisting in choosing bold or conservative strategies and not due to dynamic alteration of strategies by analysts. In other words, an analyst who is always bold is more likely to take on extreme positions in relative performance. The results presented in Table 4 panel are still consistent with a U-shaped relation between past performance and future boldness even after controlling for past boldness. In contrast to the results presented in Table 3, we find that the relation tends to be driven by the best past performance decile and the worst past performance decile. Panels B and C show that the choice of time horizon over which past forecast accuracy is calculated produces similar results.

Fama-MacBeth Regressions:

Our findings above are consistent with a U-shaped relation between boldness and past performance. In this section, we use Fama-MacBeth (1973) regressions to examine the robustness of our results. For every quarter in our sample period, we estimate a cross-sectional regression relating future boldness to past accuracy. Given the results in Table 4, we include in the regression model indicator variables if the analyst's past performance is either in the top performance decile or the bottom performance decile. Based on the existing literature, we also control for an analyst's experience and the number of firms covered by the analyst. *High experience* is an indicator variable taking the value of one if the analyst has more than four years of prior experience and zero

otherwise. *Number of firms* covered is the number of firms the analyst covers in a given quarter. We also control for the average boldness of the analyst over the previous four quarters.

In Table 5, we report the average coefficients from these regressions along with the associated p-value from a simple t-test for the statistical significance of the estimates. In specification (1), we find that analysts in the worst past performance decile and analysts in the best performance decile have significantly higher boldness than other analysts. Specification (2) shows that this result holds even after controlling for experience, past boldness, and the number of firms covered by the analyst. Interestingly, the coefficient on number of firms covered is negative and significant, indicating that analysts covering more stocks tend to issue more conservative forecasts. The final specification examines the interaction between past performance and experience and boldness. The results indicate that experienced analysts are more likely to deviate from the consensus following poor performance. However, experienced analysts with good past performance are not more likely to issuer bolder forecasts.

5. Conclusions

We examine the dynamic forecasting behavior of investment analysts in response to their prior performance relative to their peers within a continuous time/multi-period framework. The model differs significantly from existing theoretical models in its dynamic nature and the fact that it does not rely on any specific assumptions regarding the existence of asymmetric information and/or differential analyst abilities. The central prediction of the model is that there is a U-shaped relationship between the *boldness* of an analyst's forecast, that is, the deviation of her forecast from the consensus and her prior relative performance. In other words, analysts who significantly out perform or under perform their peers issue bolder forecasts than intermediate performers.

We then test the predictions of our model on observed analyst forecast data. Consistent with our theoretical predictions, we document an approximately U-shaped relationship between the deviation of analysts' forecasts from the consensus and their prior relative performance. This result is robust to different empirical methodologies. Consistent with prior empirical evidence, our results imply that analysts who face greater employment risk have greater incentives to herd. Our theoretical and empirical analyses therefore highlight the importance of both *explicit incentives* in the form of compensation structures and *implicit incentives* in the form of career concerns, on the dynamic forecasting behavior of analysts.

As noted by Welch (2000) one drawback of many existing models of herding is that they are static and designed only to explain a steady state in which all analysts herd perfectly. These models are unable to explain why analysts may deviate from the consensus some times and herd at other times. We address this issue by theoretically and empirically examining the dynamic forecasting behavior of investment analysts in response to their prior performance relative to their peers.

References

Fama, E. and J. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy* 81: 607-636.

Graham, John, 1999, "Herding among investment newsletters: Theory and evidence," *Journal of Finance* 54: 237-268.

Hilary, G. and L. Menzly, 2001, "Does past success lead analysts to become overconfident?" Working paper: University of Chicago.

Hong, H., J. Kubik, and A. Solomon, 2000, "Security analysts' career concerns and herding of earnings forecasts," *Rand Journal of Economics* 31: 121-144.

Hu, P., J. Kale, and A. Subramanian, 2003, "Relative risk choice by mutual fund managers," Working paper: Georgia Tech.

Jacob, J., T. Lys, and M. Neale, 1999,"Expertise in forecasting performance of security analysts," *Journal of Accounting and Economics* 28: 51-82.

Karatzas, I. And S. Shreve, 1998, Methods of Mathematical Finance, Springer-Verlag: New York.

Leone, Andrew and Joanna Wu, 2002, "What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts", working paper, University of Rochester.

Mikhail, M., B. Walther, and R. Willis, 1999, "Does forecast accuracy matter to security analysts?" *The Accounting Review* 74: 185-200.

Morris, S, 1997, "An instrumental theory of political correctness," Working paper: University of Pennsylvania.

Prendergast, C. and L. Stole, 1996, "Impetuous youngsters and jaded oldtimers: Acquiring a reputation for learning," *Journal of Political Economy* 1996: 1105-1134.

Rosen, S., 1981, "The economics of superstars," *The American Economic Review* 71: 845-858.

Scharfstein, D. and J. Stein, 1990, "Herd behavior and investment," *The American Economic Review* 80: 465-479.

Welch, I, 2000, "Herding among security analysts," *Journal of Financial Economics* 58: 369-396.

Trueman, B., 1994, "Analyst forecasts and herding behavior," *Review of Financial Studies* 71: 97-124

Zwiebel, J., 1995, "Corporate conservatism and relative compensation," *Journal of Political Economy* 103, 1-25:

Appendix

Proof of Lemma 1

We first note that

$$\frac{1}{2}\sigma_i^2(1)^2 + (\mu_i - \frac{1}{2}\sigma_i^2)(1) - \beta = \mu_i - \beta < 0$$

since $\mu_i < \beta$ by assumption. Therefore, the above implies that $\eta_i^- < 1 < \eta_i^+$ since η_i^+, η_i^- are the

roots of the quadratic equation $\frac{1}{2}\sigma_i^2 x^2 + (\mu_i - \frac{1}{2}\sigma_i^2)x - \beta = 0$.

We have

$$\frac{1}{2}\sigma_1^2\eta_2^{+2} + (\mu_1 - \frac{1}{2}\sigma_1^2)\eta_2^+ - \beta = \frac{1}{2}\sigma_1^2(\eta_2^{+2} - \eta_2^+) + \mu_1\eta_2^+ - \beta > \frac{1}{2}\sigma_2^2(\eta_2^{+2} - \eta_2^+) + \mu_2\eta_2^+ - \beta = 0$$

since $\mu_1 > \mu_2, \sigma_1 > \sigma_2, \eta_2^+ > 1$. Therefore,

$$\frac{1}{2}\sigma_1^2(\eta_2^+ - \eta_1^+)(\eta_2^+ - \eta_1^-) = \frac{1}{2}\sigma_1^2\eta_2^{+2} + (\mu_1 - \frac{1}{2}\sigma_1^2)\eta_2^+ - \beta > 0$$

It follows that η_2^+ must be greater than η_1^+, η_1^-, i.e. $\eta_1^+ < \eta_2^+$. We may similarly show that $\rho_1^+ < \rho_2^+$.

We now note that

$$\frac{1}{2}\sigma_i^2\eta_i^{+2} + (\mu_i - \frac{1}{2}\sigma_i^2)\eta_i^+ - (\beta + \alpha\delta) = (\frac{1}{2}\sigma_i^2\eta_i^{+2} + (\mu_i - \frac{1}{2}\sigma_i^2)\eta_i^+ - \beta) - \alpha\delta < 0$$

since the expression in the brackets on the right hand side above is zero by the definition of the root

η_i^+ and $\alpha\delta > 0$. From the above, it easily follows that we must have $\rho_i^- < \eta_i^+ < \rho_i^+$.

Proof of Proposition 2:

The proof proceeds by explicitly constructing a differentiable function u that satisfies the

hypotheses of Proposition 1, that is,

(A1) $\sup_{i=1,2}[1_{p>p_b} L_i u + 1_{p \leq p_b} L_i' u + p^c] = 0$ and $\lim_{p \to \infty} u(p)/p^c < \infty$

We first show that there exists p_* with $p_b < p_* < \infty$ such that the function u_{p_*} is twice differentiable everywhere except possibly at p_b and that

(A2) $L_i(u_{p_*}) + p^c \mid_{p=p_*} = 0; \ i = 1,2$

We begin by noting that the function $L_2(u_r) + p^c \mid_{p=r+}$ is a continuous function of r.

Condition (13) of the proposition therefore implies that there exists $p_b' > p_b$ such that

(A3) $L_2(u_{p_b'}) + p^c \mid_{p=p_b'+} > 0$

By the definition of the function $u_{p_b'}$ (setting $r = p_b'$ in (12)), we see that

(A4) $L_2(u_{p_b'}) + p^c \mid_{p=p_b'-} = 0$

and

(A5) $L_1(u_{p_b'}) + p^c \mid_{p=p_b'+} = 0$

Subtracting (A4) from (A3), we see that

$$\frac{1}{2}\sigma_2^2 (p_b')^2 [\frac{d^2}{dp^2}(u_{p_b'}) \mid_{p=p_b'+} - \frac{d^2}{dp^2}(u_{p_b'}) \mid_{p=p_b'-}] > 0$$

since $u_{p_b'}$ is differentiable everywhere by construction. This implies that

(A6) $\frac{1}{2}\sigma_1^2 (p_b')^2 [\frac{d^2}{dp^2}(u_{p_b'}) \mid_{p=p_b'+} - \frac{d^2}{dp^2}(u_{p_b'}) \mid_{p=p_b'-}] > 0$

(A5) and (A6) clearly imply that

(A7) $L^1(u_{p_b'}) + p^c \mid_{p=p_b'-} < 0$

We now show the existence of p_* such that

(A8) $L^1(u_{p_*}) + p^c \mid_{p=p_*-} = 0$

It is not difficult to show that this implies that u_{p_*} is twice differentiable at p_* and that

(A9) $L^1(u_{p_*}) + p^c \mid_{p=p_*} = L^2(u_{p_*}) + p^c \mid_{p=p_*} = 0$

We prove this by first showing that

(A10) $\lim_{r \to \infty} L^1(u_r) + p^c \mid_{p=r-} = \infty$

As $r \to \infty$, the value function u_r clearly approaches the value function u_∞ of the policy of choosing

strategy 2 for $p > p_b$. It is easy to see that the functional form of u_∞ for $p > p_b$ is

(A11) $u_\infty(p) = C_\infty p^{\eta_2^-} - \dfrac{p^c}{\dfrac{1}{2}\sigma_2^2 c^2 + (\mu_2 - \dfrac{1}{2}\sigma_2^2)c - \beta}$

We now note that

$\lim_{p \to \infty} L^1(u_\infty) + p^c =$

(A12)
$\lim_{p \to \infty}[C_\infty p^{\eta_2^-}(\dfrac{1}{2}\sigma_1^2(\eta_2^-)^2 + (\mu_2 - \dfrac{1}{2}\sigma_1^2)\eta_2^- - \beta) + \dfrac{[\dfrac{1}{2}(\sigma_2^2 - \sigma_1^2)(c^2 - c) + (\mu_2 - \mu_1)c]p^c}{\dfrac{1}{2}\sigma_2^2 c^2 + (\mu_2 - \dfrac{1}{2}\sigma_2^2)c - \beta}]$

Since $\eta_2^- < 0$, the limit of the first term above is zero. Since

$\dfrac{1}{2}(\sigma_2^2 - \sigma_1^2)(c^2 - c) + (\mu_2 - \mu_1)c = \dfrac{1}{2}(\sigma_2^2 - \sigma_1^2)c^2 < 0$ as $\mu_i = \dfrac{1}{2}\sigma_i^2 - \mu, \sigma_1 > \sigma_2$ and

$\dfrac{1}{2}\sigma_2^2 c^2 + (\mu_2 - \dfrac{1}{2}\sigma_2^2)c - \beta < 0$ as $\eta_2^- < c < \eta_2^+$, the second term is positive and goes to infinity as

$p \to \infty$. This implies that (A10) holds[10].

It now easily follows by continuity that there exists $p_* > p_b$ such that (A8) holds and

therefore (A9) holds. We now show that p_* is the required "optimal switching point" where u_{p_*} is

defined by setting $r = p_*$ in (12). By the result of Proposition 3, we need to show that

---

[10] Strictly this needs to be shown rigorously, but the arguments are quite straightforward and are available upon request.

$$L_2(u_{p_*}) + p^c \le 0; \text{ for } p \ge p_*$$

(A13) $\quad L_2'(u_{p_*}) + p^c \le 0; \text{ for } p \le p_b$

$$L_1(u_{p_*}) + p^c \le 0; \text{ for } p_b < p < p_*$$

For $p > p_*$,

(A14) $L^2(u_{p_*}) + p^c = D_{p_*} \dfrac{1}{2}\sigma_2{}^2 p^{\eta_1^-} (\eta_1^- - \eta_2^-)(\eta_1^- - \eta_2^+) + \dfrac{[\frac{1}{2}(\sigma_1{}^2 - \sigma_2{}^2)(c^2 - c) + (\mu_1 - \mu_2)c]p^c}{\frac{1}{2}\sigma_1{}^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1{}^2)c - \beta}$

Since $c < \eta_1^+ < \eta_2^+$ by hypothesis, $\dfrac{1}{2}\sigma_1{}^2 c^2 + (\mu_1 - \dfrac{1}{2}\sigma_1{}^2)c - \beta < 0$. Since $\mu_i = \dfrac{1}{2}\sigma_i{}^2 - \mu$ by

definition, $\dfrac{1}{2}(\sigma_1{}^2 - \sigma_2{}^2)(c^2 - c) + (\mu_1 - \mu_2)c = \dfrac{1}{2}(\sigma_1{}^2 - \sigma_2{}^2)c^2 > 0$ as $\sigma_1 > \sigma_2$. Therefore, the

second term on the right hand side above is always negative. Therefore, (A9) can hold only if

the first term on the right hand side of (A14) is positive and the second term is negative. Since

$\eta_1^- < 0 < c$, it follows that the expression

$$D_{p_*} \dfrac{1}{2}\sigma_2{}^2 p^{\eta_1^-} (\eta_1^- - \eta_2^-)(\eta_1^- - \eta_2^+) + \dfrac{[\frac{1}{2}(\sigma_1{}^2 - \sigma_2{}^2)(c^2 - c) + (\mu_1 - \mu_2)c]p^c}{\frac{1}{2}\sigma_1{}^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1{}^2)c - \beta}$$

is a *decreasing function* of p. Therefore, (A9) implies that

(A15) $\quad L^2(u_{p_*}) + p^c < 0 \text{ for } p > p_*$

Using the fact that u_{p_*} is twice differentiable at $p = p_*$, we can show (after some tedious algebra

that we omit for the sake of brevity) that (A15) implies that the coefficients B_{p_*}, C_{p_*} in the

definition (12) of u_{p_*} satisfy

(A16) $\quad B_{p_*} > 0, C_{p_*} < 0$

We now note that

$$L^1(u_{p_*}) + p^c = B_{p_*} \frac{1}{2}\sigma_1^2(\eta_2^+ - \eta_1^-)(\eta_2^+ - \eta_1^+)p^{\eta_2^+} + C_{p_*}\frac{1}{2}\sigma_1^2(\eta_2^- - \eta_1^-)(\eta_2^- - \eta_1^+)p^{\eta_2^-}$$

(A17)
$$+ \frac{[\frac{1}{2}(\sigma_2^2 - \sigma_1^2)(c^2 - c) + (\mu_2 - \mu_1)c]p^c}{\frac{1}{2}\sigma_2^2 c^2 + (\mu_2 - \frac{1}{2}\sigma_2^2)c - \beta}$$

Since $\eta_2^- < \eta_1^- < 0 < \eta_1^+ < \eta_2^+$, it follows from (A16) that the first and second terms on the right hand

side above are both increasing functions of p. Since

$$\frac{1}{2}(\sigma_2^2 - \sigma_1^2)(c^2 - c) + (\mu_2 - \mu_1)c = \frac{1}{2}(\sigma_2^2 - \sigma_1^2)c^2 < 0,$$ the third term above is positive and

therefore also an increasing function of p. Therefore, $L^1(u_{p_*}) + p^c$ is an *increasing function* of p

for $p_b < p < p_*$. (A8) and (A9) now clearly imply that

(A18) $L^1(u_{p_*}) + p^c < 0$ for $p_b < p < p_*$

It only remains to show that

(A19) $L_2'(u_{p_*}) + p^c < 0$ for $0 < p < p_b$

(A18) implies that the value function u_{p_*} of the analyst must be strictly greater than the value

function of choosing strategy 1 for all values of p. The value function of choosing strategy 1

always must be at least as great as $-\dfrac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - (\beta + \alpha\delta)}$ that is the value the analyst

obtains from choosing strategy 1 when he faces employment risk for all values of p, that is, there is

no finite threshold p_b beyond which the analyst does not face employment risk.

From the definition (12) of u_{p_*}, this easily implies that we must have

(A20) $A_{p_*} > 0$

We now note that

$$L_2(u_{p_*}) + p^c = (\frac{1}{2}\sigma_2^2(\rho_1^+)^2 + (\mu_2 - \frac{1}{2}\sigma_2^2)\rho_1^+ - (\beta + \alpha\delta))A_{p_*}p^{\rho_1^+} +$$

(A21) $\quad \dfrac{[\frac{1}{2}(\sigma_1^2 - \sigma_2^2)(c^2 - c) + (\mu_1 - \mu_2)c]p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - (\beta + \alpha\delta)}$ $\hspace{5cm}$ for $p < p_b$

Since $\rho_1^+ < \rho_2^+, c < \rho_1^+, \sigma_1 > \sigma_2$, (A20) implies that the first term on the right hand side above is

negative and the second term is also negative. This clearly implies (A19). Therefore, we have

shown that the value function u_{p_*} satisfies the hypotheses of Proposition 3 and is therefore the

optimal value function of the analyst. Hence, the policy of switching policies at p_* is optimal. This

completes the proof.

Proof of Proposition 3

Setting $r = p_b$ in (12), the value function u_{p_b} has the following functional form:

$$u_{p_b} = A_{p_b}p^{\rho_1^+} - \frac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - (\beta + \alpha\delta)}; p \le p_b$$

(A22) $\qquad = D_{p_b}p^{\eta_1^-} - \dfrac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - \beta}; p > p_b$

Since u_{p_b} must be at least as great as the value function of choosing strategy 1 when the analyst

faces employment risk for all values of p, we must have

(A23) $\quad A_{p_b} > 0$

However, the value function is strictly less than the value function $-\dfrac{p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - \beta}$ of

choosing strategy 1 when the analyst faces no employment risk. Therefore, we must have

(A24) $D_{p_b} < 0$

In order to establish the optimality of choosing strategy 1 for all values of p, we need to show that

(A25)
$$L_2(u_{p_b}) + p^c \leq 0, p > p_b$$
$$L_2'(u_{p_b}) + p^c \leq 0, p < p_b$$

We now note that

$$L_2'(u_{p_*}) + p^c = (\frac{1}{2}\sigma_2^2(\rho_1^+)^2 + (\mu_2 - \frac{1}{2}\sigma_2^2)\rho_1^+ - (\beta + \alpha\delta))A_{p_b} p^{\rho_1^+} +$$

(A26) $$\frac{[\frac{1}{2}(\sigma_1^2 - \sigma_2^2)(c^2 - c) + (\mu_1 - \mu_2)c]p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - (\beta + \alpha\delta)}$$ for $p < p_b$

Since $\rho_1^+ < \rho_2^+, c < \rho_1^+, \sigma_1 > \sigma_2$, (A23) implies that the first term on the right hand side above is

negative and the second term is also negative. Therefore, $L_2'(u_{p_*}) + p^c < 0, p < p_b$.

For $p > p_b$, we have

$$L_2(u_{p_b}) + p^c = (\frac{1}{2}\sigma_2^2(\eta_1^-)^2 + (\mu_2 - \frac{1}{2}\sigma_2^2)\eta_1^- - \beta)D_{p_b} p^{\eta_1^-} +$$

(A27) $$\frac{[\frac{1}{2}(\sigma_1^2 - \sigma_2^2)(c^2 - c) + (\mu_1 - \mu_2)c]p^c}{\frac{1}{2}\sigma_1^2 c^2 + (\mu_1 - \frac{1}{2}\sigma_1^2)c - \beta}$$

If $\eta_1^- < \eta_2^-$, then (A24) implies that the first term on the right hand side above is negative. Since

$1 \leq c < \eta_1^+$, the second term is also negative. On the other hand, if $\eta_1^- > \eta_2^-$, then the first term on the

right hand side of (A27) is positive and the second term is negative. This implies that $L_2(u_{p_b}) + p^c$

is a *decreasing function* of p for $p > p_b$. Hypothesis (14) of the proposition now implies that

$L_2(u_{p_b}) + p^c < 0, p > p_b$. Therefore, we have established (A25). Hence, the function u_{p_b} satisfies

the hypotheses of Proposition 3. Hence, the policy of always choosing strategy 1 is optimal. This completes the proof.

Table 1. Summary Statistics

This table shows various summary statistics for our sample of analysts' forecasts. The data consists of all quarterly forecasts of earnings per share between 1988 and 2000 contained in the IBES Detail History File. The number of investment firms each quarter and number of analysts issuing forecasts each quarter are the average number of investment firms submitting forecasts to IBES each quarter and the number of unique analysts submitting forecasts respectively. Experience is calculated for each analyst as the difference between the year of the forecast and the analyst's first year submitting forecasts to the IBES database. Analyst boldness and analyst accuracy are calculated using the ranking procedure of Hong, Kubik, and Solomon (2000).

	Mean	Median	Std. Deviation	Min	Max
Number of investment firms each quarter	366.37	386.00	90.99	194.00	510.00
Number of analysts issuing forecasts each quarter	1,763.67	1,607.00	626.50	768.00	2,861.00
Number of stocks covered per quarter	7.32	6.00	6.28	1.00	94.00
Number of analysts following a stock	5.56	4.00	4.35	2.00	38.00
Experience	4.46	3.00	3.96	0.00	20.00
Average analyst boldness	50.38	50.00	18.32	0.00	100.00
Average analyst accuracy	50.46	50.00	19.68	0.00	100.00

Table 2. Transition Matrix

This table shows a transition matrix relating past average boldness to future boldness.   The table is constructed as follows. For each quarter in our sample period, we rank analysts into deciles based on their average boldness score over the previous four quarters. We perform a similar ranking of the boldness scores for the current score. The p-value from a χ^2 test for equal proportions is reported for each past boldness decile.

	Boldness										
	Least Bold	2	3	4	5	6	7	8	9	Most Bold	p-value from χ^2 test for equal proportions
Least Bold	15.93%	11.09%	8.43%	7.62%	7.44%	10.46%	6.94%	7.58%	9.50%	15.02%	0.001
2	10.79%	11.90%	11.90%	10.80%	10.07%	9.87%	9.49%	8.85%	8.34%	8.00%	0.001
3	8.72%	11.53%	12.27%	11.59%	11.05%	9.74%	10.21%	9.45%	8.47%	6.97%	0.001
4	8.02%	10.79%	11.51%	12.18%	10.98%	10.12%	11.05%	10.06%	8.61%	6.68%	0.001
5	7.21%	9.96%	12.04%	11.76%	11.56%	9.77%	11.49%	10.29%	9.42%	6.51%	0.001
6	8.99%	9.45%	9.78%	10.31%	10.45%	11.46%	10.62%	10.91%	9.13%	8.89%	0.001
7	7.38%	9.48%	9.59%	10.91%	10.99%	9.97%	11.05%	11.76%	10.86%	8.01%	0.001
8	7.95%	8.80%	9.62%	10.17%	9.48%	10.17%	10.96%	11.26%	12.07%	9.52%	0.001
9	9.70%	8.82%	7.85%	8.29%	8.96%	9.76%	10.90%	11.19%	12.17%	12.37%	0.001
Most Bold	14.97%	8.26%	7.07%	6.55%	7.20%	10.54%	7.39%	8.67%	11.33%	18.02%	0.001

Average boldness over previous 4 quarters (row label)

Table 3. Past performance and future boldness
This table examines the relation between past performance and future boldness. Analysts are ranked each quarter on the basis of their average accuracy over the previous four, eight, and twelve quarters. The analysts are then sorted into deciles, with Decile 1 containing the worst performing analysts and Decile 10 containing the best performing analyst. We then compute mean future boldness for each decile. The p-value reported for each decile tests whether the mean boldness is statistically different from the average boldness of quartiles five and six.

Panel A: Performance rankings based on one year of past data

Past Performance quartile	Past Performance	Boldness	# of observations	P-value
Worst	26.38	51.72	51	0.00
2	40.31	51.27	51	0.00
3	44.46	50.53	51	0.00
4	47.25	50.19	51	0.01
5	49.54	49.89	51	0.02
6	51.61	49.28	51	0.02
7	54.00	49.20	51	0.07
8	56.97	49.60	51	0.96
9	61.54	49.89	51	0.26
Best	76.31	50.55	51	0.00

Panel B: Performance rankings based on two years of past data

Past Performance quartile	Past Performance	Boldness	# of observations	P-value
Worst	28.54	51.80	51	0.00
2	41.89	51.49	51	0.00
3	45.48	50.60	51	0.00
4	47.88	50.18	51	0.03
5	49.81	50.13	51	0.01
6	51.59	49.28	51	0.01
7	53.62	49.39	51	0.19
8	56.13	49.02	51	0.01
9	60.12	49.74	51	0.90
Best	74.24	50.96	51	0.00

Panel C: Performance rankings based on three years of past data

Past Performance quartile	Past Performance	Boldness	# of observations	P-value
Worst	29.24	51.75	51	0.00
2	42.42	51.30	51	0.00
3	45.89	50.76	51	0.00
4	48.13	50.48	51	0.00
5	49.93	49.99	51	0.03
6	51.63	49.41	51	0.03
7	53.53	49.11	51	0.01
8	55.88	49.32	51	0.17
9	59.72	49.68	51	0.84
Best	73.59	50.90	51	0.00

Table 4. Past Performance and future boldness controlling for past boldness
This table examines the relation between past performance and future boldness, controlling for past boldness. Past boldness is calculated as the average boldness score for each analyst over the previous four quarters. The p-value reported for each decile tests whether the mean boldness is statistically different from the average boldness of quartiles five and six.

Panel A: Performance rankings based on one year of past data.

Past Performance quartile	Boldness	# of observations	P-value
Worst	51.06	51	0.00
2	50.71	51	0.00
3	50.38	51	0.08
4	50.19	51	0.29
5	50.12	51	0.13
6	49.68	51	0.13
7	49.70	51	0.44
8	49.94	51	0.87
9	50.08	51	0.48
Best	50.61	51	0.00

Panel B: Performance rankings based on two years of past data

Past Performance quartile	Boldness	# of observations	P-value
Worst	51.09	51	0.00
2	50.74	51	0.00
3	50.24	51	0.18
4	50.62	51	0.00
5	49.94	51	0.65
6	49.81	51	0.65
7	50.06	51	0.57
8	49.60	51	0.19
9	49.78	51	0.75
Best	50.76	51	0.00

Panel C: Performance rankings based on three years of past data

Past Performance quartile	Boldness	# of observations	P-value
Worst	51.09	51	0.00
2	50.57	51	0.03
3	50.45	51	0.02
4	50.44	51	0.06
5	50.26	51	0.10
6	49.73	51	0.10
7	49.90	51	0.70
8	49.73	51	0.20
9	49.85	51	0.59
Best	50.65	51	0.02

Table 5. Fama-MacBeth Regressions

This table reports the results of Fama-MacBeth regressions. The dependent variable in each regression is the boldness score. Worst past performance decile is an indicator variable that takes the value of one if the analyst's average forecasting performance over the previous four quarters was in the bottom 10% and zero otherwise. Best past performance decile is an indicator variable that takes the value of one if the analyst's average forecasting performance over the previous four quarters was in the top 10% and zero otherwise. High experience is an indicator variable that takes the value of one if the analysts has four or more years of experience and zero otherwise. Number of firms covered is the number of firms covered by the analyst in the quarter. Past boldness is the average boldness score of the analyst over the previous four quarters. We include controls for industry effects. We use the IBES SIG code to define the industries. The average R^2 from the 51 quarterly regressions are reported. P-values are reported in parentheses.

	Boldness		
	(1)	(2)	(3)
Intercept	50.64	46.94	46.87
	(0.00)	(0.00)	(0.00)
Worst past performance decile	1.88	1.14	0.56
	(0.00)	(0.00)	(0.15)
Best past performance decile	0.80	0.50	0.67
	(0.00)	(0.05)	(0.04)
high experience		0.06	-0.04
		(0.71)	(0.80)
past boldness		0.08	0.08
		(0.00)	(0.00)
number of firms covered		-0.05	-0.05
		(0.00)	(0.00)
Worst past performance decile*high exp			1.80
			(0.01)
Best past performance decile*high exp			-0.94
			(0.22)
Industry Effects	YES	YES	YES
N	51	51	51
Average R^2	0.073	0.079	0.081