

## WORKING PAPER

Can Relative Performance Compensation  
Explain Analysts' Forecasts of Earnings?

Edward Kutsoati  
Department of Economics  
Tufts University

and

Dan Bernhardt  
Department of Economics  
University of Illinois

Discussion Paper 99-09  
Department of Economics  
Tufts University

Department of Economics  
Tufts University  
Medford, MA 02155  
(617) 627-3560

# Can Relative Performance Compensation Explain Analysts' Forecasts of Earnings?\*

Edward Kutsoati<sup>†</sup>      Dan Bernhardt<sup>‡</sup>

May 20, 1999

## Abstract

This paper derives and tests the implications of relative performance incentives for the forecasts of financial analysts. If, in addition to compensation for absolute forecast accuracy, analysts are compensated on the basis of how their forecasts compare with those of other analysts, then earlier forecasts affect later announcements. The theoretical predictions regarding the direction of bias are tested using data on individual analysts' forecasts of earnings per share (EPS). We find very strong evidence that the last analyst to report a forecast "over-emphasizes" his private information by reporting a contrarian forecast that overshoots the consensus (mean) forecast. We also find modest evidence that investors do not unravel biases in the forecasts of firms followed by fewer analysts.

---

\*We are grateful to the Institutional Brokers Estimate System (I/B/E/S), a service of I/B/E/S International Inc. for providing analysts' forecasts of earnings per share. We thank Lucy Ackert, Ted Juhl, Gregor Smith, Brett Trueman, Mike Waldman, Ján Zábajnik for very helpful suggestions, and Vaclav Polasek for help with programming. We are solely responsible for all errors.

<sup>†</sup>Department of Economics, Tufts University, Medford, MA, 02155

<sup>‡</sup>Department of Economics, 463 Commerce West, University of Illinois, Champaign, IL. 61820

# 1 Introduction

The forecasts of firms' earnings made by financial analysts serve as a major source of information for brokerage houses and money management funds, who rely on the accuracy of these forecasts in providing services for their clients. This paper explores how the way in which an analyst is compensated affects his forecasts.<sup>1</sup> In particular, we posit that an analyst's compensation may be tied both directly to absolute forecast accuracy and to how the accuracy of his forecast compares with those of other analysts; that is, to the analyst's relative forecast accuracy.

Relative performance compensation emerges naturally in environments where potential employers and clients are trying to infer analysts' abilities from their forecasts.<sup>2</sup> Firms with better analysts win more clients, generating more brokerage commissions and profits. Because the relative accuracy of an analyst's forecast helps firms and investors unravel the analyst's forecasting ability, competition among firms for better analysts leads to relative performance compensation. Indeed, brokerage houses might want to design performance compensation contracts that induce analysts to over-emphasize their private information and hence reveal their ability more quickly.

There is significant evidence suggesting that analysts' compensation depends on the relative accuracy of their forecasts. For example, the *Institutional Investor (II)* publishes an annual poll of analysts ranked by their forecasting record, and in many cases an analyst's salary is affected by his rank on the *II* poll (see Stickel (1990, 1992)).<sup>3</sup> In the 1970s, Merrill Lynch hired top-ranked analysts using the *II* poll as a guide.<sup>4</sup> Generally, it is estimated that analysts earn up to 2 or 3 times their basic salary in bonuses if they get a high rating in the *II* or an equivalent survey.

We allow for very general functional forms of relative performance compensation and derive the implications of the different curvatures for the forecast of the last analyst to report earnings. If analysts are compensated on the basis of relative performance, then when reporting their forecasts,

---

<sup>1</sup>Dugar and Nathan (1995) discuss some other incentives that may impact on analysts' forecasts. Lin and McNichols (1998) documents evidence suggesting that analysts working for investment banks that under-write equity offerings tend to issue more favorable recommendations. In general, there is also a widely-held belief that analysts issue optimistic forecasts in order to retain access to management (e.g., "Incredible Buys: Many companies press analysts to steer clear of negative ratings," (*Wall Street Journal*, July 19, 1995); and "Analyst takes heat for a 'Sell' Rating on Archer Daniels Midland," (*Wall Street Journal*, August 31, 1995) details how a favored analyst was shunned by Archer Daniels Midland (ADM) for switching to a 'sell' recommendation of ADM).

<sup>2</sup>In a labor market context, Holmstrom (1982) explored how performance contracts distort the actions of workers when the market is trying to learn workers' abilities. See, also Scharfstein and Stein (1990), and Zweibel (1995). Chevalier and Ellison (1997) document empirically how mutual fund manager's investments evolve over time, presumably reflecting learning by investors who direct investment funds according to performance.

<sup>3</sup>A similar recognition, *The Blue Chip Economic Forecasting Award*, is bestowed on the *best* forecaster of macroeconomic trends.

<sup>4</sup>"Who Really Moves The Market," *Fortune*, Oct. 27, 1997.

analysts will balance their aims of minimizing forecast errors against looking good (relative to other forecasters). For example, the fear of being fired upon relative under-performance<sup>5</sup> might make the analyst averse to the possibility of substantially under-performing other analysts, implying concave relative performance compensation on this range. Alternatively, being “right”, when everyone else is “wrong”, might lead to a significant increase in the number of clients desiring the analyst’s services, implying a convex compensation relationship over some range of relative forecasts. One or two accurate predictions that depart sharply from the consensus can give an analyst press attention and, hence a career (see Henry (1989) and Laster, Bennett and Geoum (1999)). If the marginal gain (loss) to being better (worse) than the average is small, then compensation may be “anti-symmetric” concave; while anti-symmetric convex compensation would be consistent with significant bonuses for being “right” when others are “wrong”, and with being fired if one’s performance is significantly worse than others.

We consider an environment in which forecasts are released sequentially and analysts observe the existing “consensus” forecast before reporting their own forecasts. In practice, forecasts are produced and become publicly available in an almost continuous fashion. For example, subscribers (and contributors) to the Institutional Brokers Estimate System (I/B/E/S) EXPRESS receive the latest forecasts of earnings made by a panel of active security analysts. This product is delivered electronically and, in most cases, it is accompanied by the mean (or “consensus”) and other summary statistics, so that analysts have a good idea of what the market consensus is, even if they don’t know what particular analysts are reporting. Analysts also disseminate their forecasts through the media, newsletters, *etc.*

We show that the consequence of concave or anti-symmetric concave relative performance compensation is that the last reporting analyst will strategically bias his report toward the consensus.<sup>6</sup> The consequence is that forecasts will be “clustered” more closely around the consensus than analysts’ information would suggest. We next show that if, instead, relative performance compensation is a convex or anti-symmetric convex function of relative performance, then the last analyst’s forecast will place excessive weight on his or her private information; such analysts will “distinguish” themselves from the consensus. An implication of our findings is that using a simple average to represent the market consensus could be inaccurate: the order of release and the possible bias in forecasts must be taken into account.

In our environment, forecasts are biased, not because analysts are irrational as De Bondt and

---

<sup>5</sup>Mikhail *et. al.*, (1999) finds that analysts who are relatively less accurate than their peers are more likely to lose their jobs, but that absolute forecast accuracy does not affect layoff probabilities.

<sup>6</sup>The consensus forecast and the outstanding mean forecast will be used interchangeably in this paper.

Thaler (1990) and Abarbanell and Bernard (1992) argue, but rather that the market provides them incentives to rationally distort their forecasts. Keane and Runkle (1998) argue that “Since financial analysts’ livelihoods depend on the accuracy of their forecasts . . . we can plausibly argue that these numbers accurately measure the analysts’ expectations.” But this argument follows only when financial analysts are compensated solely on the basis of absolute accuracy — and when the market uses relative forecast accuracy to glean information about an analyst’s ability, analysts will also be rewarded for relatively more accurate forecasts.

We then use individual analysts’ forecasts of earnings per share (EPS) to test these theoretical predictions regarding the direction of forecast bias. We find overwhelming evidence that the last analyst to report a forecast of a firm’s EPS “exaggerates” his private information by announcing a contrarian forecast, overshooting true EPS away from the consensus. If the last forecast is unbiased, then given that it is less than the consensus it should be equally likely to exceed actual EPS as fall short. In our sample, however, it falls short 68% of the time. Similarly, conditional on the last forecast exceeding the consensus, the last forecast overshoots true EPS 56 percent of the time. This finding is all the more remarkable in light of the fact that the factors highlighted by Keane and Runkle (1998) — correlated errors in analysts’ signals because a primary source of information to analysts, firm management, is common to all analysts; common unanticipated shocks to earnings; consistently over-optimistic forecasts to retain access to management — would, *ceteris paribus* lead to *clustering* of forecasts. Our findings are consistent with convexity of analysts’ relative compensation scheme, and are sharply inconsistent with concave relative compensation.

We then test to see whether investors correctly unravel the bias in the last analyst’s forecast. In particular, we investigate whether the difference between the last forecast and the consensus can predict the market’s response to the firm’s earnings announcement. If investors do not unravel the bias, then if the last forecast exceeded (lagged) the consensus, they will systematically over-predict (under-predict) earnings and will be disappointed (excited) by the earnings announcement, leading to a negative (positive) excess returns. We find modest evidence that investors in fact do not unravel the bias for firms that are followed by relatively few analysts.

Recent studies that explored how economic environments can lead to biased financial forecasts include Trueman (1994). In a model based on heterogeneous forecasting abilities, Trueman (1994) showed that when investors attempt to infer ability from forecasts, less able analysts will tend to mimic the forecasts of others, leading them to herd toward the consensus forecast. In a similar model, Trueman (1990) argues that an analyst will be reluctant to revise her forecast upon receiving new information since such action will mean the initial forecast is less precise and the analyst

will be perceived as incompetent. Ehrbeck and Waldmann (1996) investigated this phenomenon using agency models in which less able forecasters balance their aims of minimizing forecast errors and making small revisions in their forecasts like more able forecasters. However, they find that forecasters of interest rates on 91-day U.S. Treasury bills revise their forecasts by too much and that large revisions tend to be correlated with large forecast errors<sup>7</sup>. This finding can be rationalized if we again posit that forecasters have a convex relative performance compensation scheme.

Laster, Bennett and Geoum (1999) assume that analysts are compensated according to mean squared error plus a bonus for the most accurate forecast among all forecasters. They find that independent forecasters of GDP growth are more likely to report outlying forecasts. Lamont (1995) and Hong *et al.* (1998) find evidence that such incentives may vary over time: analysts who have been in the business longer tend to make bolder and less accurate forecasts.

Our empirical results appear inconsistent with the theoretical clustering predictions of Trueman (1994) and contrast *sharply* with the findings of Keane and Runkle (1998). Keane and Runkle argue that once the fact that analysts receive information from common sources is properly accounted for in the estimation procedure, that “The evidence strongly supports the view that professional stock market analysts make rational (unbiased) forecasts of earnings per share.” Since positive correlation in signals leads to clustering, the difference between our findings and Keane and Runkle’s is startling. There are many plausible reasons for why Keane and Runkle fail to uncover our relationship. Most importantly, Keane and Runkle do not distinguish among forecasters. With convex compensation, robust predictions about the nature of bias in analyst’s forecasts only obtain for the *last* forecaster, because earlier forecasters must take into account how their forecasts affect subsequent forecasts. Also, self-selection may mean that there is something special about the last forecaster — for example, the last forecaster may be the one with the most to gain from differentiating himself from others. We only document the bias in the last forecast and are unconcerned with identifying the properties of who *is* last. Finally, Keane and Runkle consider forecasts for only 21 firms each of which is heavily followed (an average of 37 analysts per firm), whereas our sample has 237 firms with varying degree of analysts’ interests.

The rest of the paper is organized as follows. After presenting a simple model of sequential forecasting in section 2, we analyze the effect of relative performance compensation for announced forecasts in section 3. In section 4, we describe the data for empirical analysis, develop testable hypotheses and present our empirical findings. Section 5 concludes. All proofs are in an appendix.

---

<sup>7</sup>Kandel and Pearson (1995) also found evidence of over-revision by security analysts and showed that analysts’ differential interpretation of public signals could explain the relation between volume of trade and stock returns around public announcements.

## 2 A Model of Sequential Forecasting

Consider  $L$  identical security analysts who sequentially report forecasts of the earnings per share,  $\theta$ , of a firm. We take the sequence of release of forecasts as given and index analysts by order in which they report their forecast<sup>8</sup>:  $j = \{1, 2, \dots, L\}$ . Investors have no strategic role in this model — we only look at investors to explore whether they correctly process the information in the publicly available forecasts of security analysts. The prior distribution of earnings per share,  $\theta$ , is known, but the actual value is realized after the last forecast. Denote the  $j$ th analyst's forecast by  $f_j$ .

Analyst  $j$  is compensated according to both his absolute performance, measured by individual error; and his relative performance, measured by the difference between own error and the error in the consensus (mean) forecast. Because we do not know the larger (unmodeled) game<sup>9</sup> that gives rise to these preferences, we derive the consequences for strategic forecast choice imposing only weak functional form restrictions on compensation:

$$w_j = \bar{w} - \lambda|f_j - \theta| + R(|f_m - \theta| - |f_j - \theta|), \quad (1)$$

where  $\bar{w} > 0$  is a fixed reward,  $f_m$  is the mean of all other forecasts, excluding  $f_j$ , and  $\lambda$  is a positive constant.<sup>10</sup>  $|f_j - \theta|$  denotes the absolute performance and the relative compensation is given by  $R(|f_m - \theta| - |f_j - \theta|)$ . Analyst compensation is increasing in relative performance: *i.e.*,  $R'(\cdot) > 0$ . All of our theoretical findings save one would extend were we to relax the linear structure on absolute performance compensation and assume only that absolute performance compensation is strictly monotonically increasing in accuracy.<sup>11</sup> Without loss of generality, we normalize  $R(0)$  to be zero.

We assume that the  $j$ -th analyst observes all preceding forecasts of other analysts in addition to a private signal,  $s_j$  of the firm's earnings per share:  $j$ 's information set,  $\Omega_j$ , includes at least  $\{f_1, f_2, \dots, f_{j-1}, s_j\}$ . Analyst  $j$ 's signal is given by

$$s_j = \theta + \eta_j, \quad (2)$$

---

<sup>8</sup>We do not model the strategic choice of when to release a forecast. Our analysis focuses on the last forecaster and is consistent with the possibility that analysts may differ in the form of their relative compensation, and that the forecaster who chooses to go last does so because he gains the most from distinguishing his forecast from others.

<sup>9</sup>Larger games that give rise to absolute and relative performance compensation include the learning about analyst ability discussed in the introduction, tournaments to overcome moral hazard among analysts, etc. We then derive and test the theoretical implications of different induced preferences, which, in turn, impose restrictions on possible larger games/primitive preferences.

<sup>10</sup>The parameter  $\lambda$  can be interpreted as the relative weight of absolute performance in an analyst's payoff function.

<sup>11</sup>Proposition 1(iv), which we do not test, requires more structure.

where  $\eta_j$  is a symmetrically and independently distributed random variable with mean zero. Note that we assume that signal errors are independent; we discuss how correlated signals affect our analysis later. Given this information set, analyst  $j$ 's posterior beliefs about  $\theta$  are captured by a density  $g(\theta|\Omega_j)$  that is symmetrically distributed about its posterior mean,  $\hat{\theta}_j$ , and strictly positive on its connected support. Let  $G(\theta|\Omega_j) = \int_{x \leq \theta} g(x|\Omega_j) dx$ .

Except for the case where compensation functions are linear, we characterize only the forecast of the last analyst. Our analysis of the last analyst permits a variety of possible informational structures, including allowing for informational arrival over time. For example, we could let  $\theta = \sum_{i=0}^{L+1} \theta_i$ , where innovations  $\theta_i$ , are mean zero and independently and symmetrically distributed so that analyst  $j$ 's posterior is symmetrically distributed around  $\hat{\theta}_j = \sum_{i=0}^j \theta_i$ . This formulation allows later forecasters to see more innovations to earnings than earlier forecasters.

The information environment would have to be specified more carefully in order to characterize the forecasts of analysts other than the last. In particular, if *inferences* by later forecasters about  $\theta$  are affected by the forecasts of earlier analysts then we would have to worry about the incentives of earlier analysts to try to “fool” subsequent forecasters into believing that they observed a different private signal, thereby influencing their posteriors and thus their announcements. This caveat is irrelevant for the last analyst, because he has no other forecaster to “fool.”

### 3 Relative Performance Compensation and the Pattern of Forecasts

To analyze how relative performance compensation affects forecast patterns, we consider, as a benchmark, a linear relative performance compensation function. We show that if  $R(\cdot)$  is linear, then an analyst's forecast is unaffected by those made by others. Even though the reports of other analysts still affect an analyst's compensation — more accurate forecasts by others reduce an analyst's compensation — each analyst maximizes expected compensation by reporting his mean squared error (MSE) minimizing forecast:  $f_j = \hat{\theta}_j, \forall j = 1, 2, \dots, n$ . In this case, forecast differences will either reflect differences in analysts's signals or their (mis)interpretation of similar signals.

**Lemma 1** *If an analyst's relative compensation function is linear, then each analyst's equilibrium forecast equals the posterior mean of his beliefs about  $\theta$ :  $f_j = \hat{\theta}_j, \forall j = 1, 2, \dots, n$ .*

We now show that when  $R''(\cdot) \neq 0$ , the outstanding forecasts (or the consensus of earlier forecasts) of other analysts will affect the last agent's forecast: Analysts balance their aims of



minimizing forecast errors against looking good *after* the outcome is realized.

The next lemma details the set of feasible reports by the last analyst given his information and the outstanding consensus forecast,  $f_m$ .

**Lemma 2** *If the consensus forecast is less than the mean earnings expected by the last analyst,  $f_m < \hat{\theta}_L$ , then his report will exceed the consensus,  $f_L > f_m$ . Conversely, if  $\hat{\theta}_L < f_m$  then  $f_L < f_m$ .*

Lemma 2 ensures that if the last analyst's expectation of  $\theta$  exceeds  $f_m$ , then so will his forecast. For every  $f'_L < f_m$ , since his posterior is symmetrically distributed about  $\hat{\theta}_L$ , there exists some forecast exceeding  $f_m$  that dominates  $f'_L$  in terms of expected reward and offers the same expected relative performance. Conversely, if he expects earnings to be less than  $f_m$ , then he reports a forecast less than  $f_m$ .

Having eliminated dominated strategies, we can now characterize (qualitatively) the pattern of forecasts by the last analyst. Depending on the nature of the relative compensation function, the last analyst will choose either to “under-emphasize” or to “exaggerate” his information.

**Concave R:**  $R''(|f_m - \theta| - |f_L - \theta|) < 0$ .

If the last analyst's compensation is concave in relative performance, this makes the analyst averse to locating too far from the consensus, lest his forecast be inaccurate. The last analyst is willing to forego some of the compensation that derives from smaller absolute errors in order to be closer to the consensus: If he is wrong, then there is comfort in ‘numbers’, in having everyone else be wrong, too. The consequence is that the last analyst puts less weight on his information (or private signal) and biases his forecast,  $f_L$ , towards the consensus forecast,  $f_m$  :

**Proposition 1** *Suppose  $R(\cdot)$  is concave. Then*

- (i) *If  $f_m = \hat{\theta}_L$ , then  $f_L = \hat{\theta}_L$ .*
- (ii) *If  $f_m < \hat{\theta}_L$  then  $f_m < f_L < \hat{\theta}_L$ .*
- (iii) *If  $f_m > \hat{\theta}_L$  then  $f_m > f_L > \hat{\theta}_L$ .*
- (iv) *The bias in  $f_L$  increases in  $(\hat{\theta}_L - f_m)$  :  $0 < \frac{\partial(\hat{\theta}_L - f_L)}{\partial(\hat{\theta}_L - f_m)} < 1$ .*

Proposition 1 states that the last analyst will bias his forecast in the direction of the consensus report: he gets a higher expected payoff by reporting such a forecast. Hence if analysts' payoffs are

a concave function of relative performance, then forecasts can become “clustered” even if individual private information suggests otherwise: The last analyst biases his forecast closer to  $f_m$  because the reward from relative performance more than compensates for losses in payoff due to a worse absolute performance. Indeed, as  $f_m$  diverges further from the last analyst’s expectation,  $\hat{\theta}_L$ , his forecast,  $f_L$ , diverges further from  $\hat{\theta}_L$  but at a rate less than one.<sup>12</sup> Equivalently, if (say) the consensus forecast is revised, this will cause the last analyst to revise his forecast in the direction of revision even though his beliefs about earnings are unaltered.

Proposition 1 forms the basis of our predictions about a concave  $R(\cdot)$ . The conservative bias introduced in the forecast implies that if  $f_L$  exceeds  $f_m$ , it is more likely that the last forecast will be less than  $\theta$ . Conversely, if  $f_L$  is less than  $f_m$ , then  $f_L$  is more likely to over-estimate the actual earnings. This generates a relationship between the pattern of forecasts and the direction of error in the last forecast.

**Convex R:**  $R''(|f_m - \theta| - |f_j - \theta|) > 0$ .

Compensation may be convex if one or two accurate forecasts that depart from the consensus draw a substantial clientele. We now show that convex relative performance compensation causes the last analyst “exaggerates” his information (or private signal) by reporting an  $f_L$  that overshoots  $\hat{\theta}_L$  away from  $f_m$ .

**Proposition 2** *Suppose that  $R(\cdot)$  is convex and  $f_m \neq \hat{\theta}_L$  be the consensus forecast. Then*

- (i) *If  $f_m < \hat{\theta}_L$  then  $f_m < \hat{\theta}_L < f_L$ .*
- (iii) *If  $f_m > \hat{\theta}_L$  then  $f_m > \hat{\theta}_L > f_L$ .*

If the relative compensation function is convex, the last analyst has an incentive to bias his forecast to “locate” away from  $f_m$ : a relatively accurate forecast yields a much higher payoff than the losses incurred if his forecast is less accurate. This induces the last analyst to take forecasting risks by reporting contrarian forecasts. The last analyst wants to “stand out”, so he exaggerates his private information and biases his forecast away from  $f_m$ . In particular, the last analyst will report an over-optimistic forecast of the firm’s EPS if  $\hat{\theta}_L > f_m$ , and an over-pessimistic forecast if  $\hat{\theta}_L < f_m$ . In other words,  $f_L$  is more likely to be greater (less) than the actual earnings if  $f_L$  is greater (less) than  $f_m$ .

---

<sup>12</sup>Similar predictions do *not* obtain for convex relative performance compensation, absent specific parameterizations.

Characterizations of forecasts by analyst  $j$ ,  $1 < j < L$  are more difficult because *even when*  $j$ 's forecast does not affect their inferences, it *still* affects the relevant consensus forecast and hence their announcement.<sup>13</sup>

### Anti-symmetric Relative Compensation Functions

**Definition:** The relative compensation function is anti-symmetric about zero if:

$$R(|f_m - \theta| - |f_j - \theta|) = \begin{cases} R(|f_m - \theta| - |f_j - \theta|), & \text{if } |f_m - \theta| > |f_j - \theta| \\ -R(|f_j - \theta| - |f_m - \theta|), & \text{if } |f_m - \theta| < |f_j - \theta|, \end{cases}$$

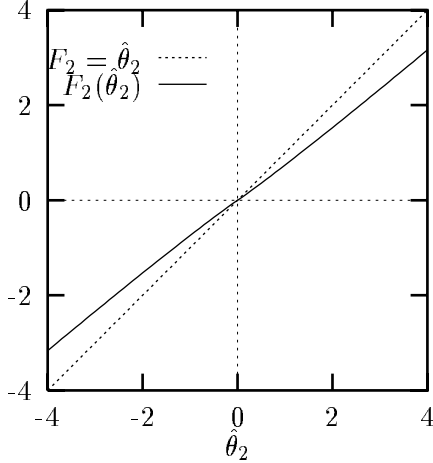
with  $R(0) = 0$  and  $R'(\cdot) > 0$ . That is,  $R(z) = -R(-z)$ ,  $\forall z > 0$ : Analysts are penalized for bad relative performance just as much as they are rewarded for a good relative performance. The relative performance compensation function is *anti-symmetric-concave* if  $R(\cdot)$  is a concave function of positive relative performances; the relative performance compensation function is *anti-symmetric-convex* if  $R(\cdot)$  is a convex function of positive relative performances.

One can show that the last analyst has an incentive to issue an exaggerated forecast if the relative performance function is “anti-symmetric convex”; and an incentive to bias his forecast toward the consensus if the relative performance function is “anti-symmetric concave.” That is, if relative performance compensation is *anti-symmetric-concave*, the last analyst “locates” closer to  $f_m$ ; and he “locates” further away from  $f_m$  if  $R$  is *anti-symmetric-convex*. The corresponding lemmas and propositions are proved in the same way. The key is to exploit the symmetry in the analysts’ beliefs about  $\theta$  and the fact that  $R'(z) = R'(-z)$  for all  $z$ . The intuition is that relative to the consensus, the last analyst anticipates being right more often than being wrong, so his forecast reflects more heavily the curvature assumptions on the payoffs for good relative performances.

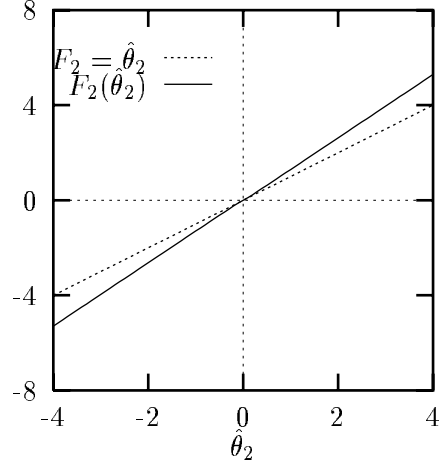
**Example:** Let  $\theta$  be uniformly distributed on  $[-4 + \hat{\theta}_L, 4 + \hat{\theta}_L]$ ,  $\lambda = \frac{1}{4}$  and  $f_m = 0$ . Figure 1(a) shows the equilibrium forecast of the last analyst as a function of  $\hat{\theta}_L$  when  $R$  is given by the *anti-symmetric-concave* function

$$R(z) = \begin{cases} \sqrt{z}, & \text{if } z > 0; \\ -\sqrt{-z}, & \text{if } z < 0. \end{cases}$$

<sup>13</sup>Consider 3 analysts, suppose that  $f_1 < \hat{\theta}_2 = \hat{\theta}_3$ , and that relative compensation is convex. Analyst 2 wants to bias his forecast away from  $f_1$  only if that bias raises  $|f_2 - 0.5(f_1 + f_3(f_2, \hat{\theta}_3))|$ . But analyst 3 will bias his forecast away from  $0.5(f_1 + f_2)$ , so that if analyst 2 biases his forecast further away from  $f_1$ , it may cause analyst 3 to further bias his forecast away from  $f_1$ . But this may reduce  $|f_2 - 0.5(f_1 + f_3(f_2, \hat{\theta}_3))|$  — predictions depend finely on the particular parameterizations.



(a) Optimal  $f_L$  for an *anti-symmetric* concave  $R$



(b) Optimal  $f_L$  for an *anti-symmetric* convex  $R$

Figure 1: *Anti-symmetric* concave  $R$  vs. *anti-symmetric* convex  $R$

Optimal  $f_L$  plotted against  $\hat{\theta}_L$  for (a) *anti-symmetric* concave  $R$  and (b) *anti-symmetric* convex  $R$ , when  $f_m = 0$ ,  $\lambda = \frac{1}{4}$ , and  $\theta$  is uniformly distributed on  $[-4 + \hat{\theta}_L, 4 + \hat{\theta}_L]$ . The 45° line represents  $f_L = \hat{\theta}_L$ .

where  $z = |f_m - \theta| - |f_L - \theta|$  is the relative forecast error. Again, the last analyst biases his forecast to “locate” closer to  $f_m = 0$ , albeit by a lesser amount than if the relative compensation function were a concave function of below par relative performances. Figure 1(b) illustrates the last analyst’s incentives to “over-emphasize” his information when  $R(z)$  is given by the *anti-symmetric*-convex function,

$$R(z) = \begin{cases} z^2, & \text{if } z > 0; \\ -z^2, & \text{if } z < 0. \end{cases}$$

The equilibrium forecast of the last analyst is biased away from  $f_m$ , albeit by a lesser amount than when the relative compensation function is a convex function of below par relative performances.

## 4 Empirical Analysis

### Data and Sample Selection

Individual analysts forecasts, forecast publication dates and the actual EPS reported by the firms are provided by the Institutional Brokers Estimate System (I/B/E/S). The database features  $n_q^{th}$  quarter ahead ( $n_q = 1, 2, \dots, 8$ ) forecasts, as well as  $n_y^{th}$  year ahead ( $n_y = 1, 2, \dots, 10$ ) forecasts and long-term growth forecasts of EPS reported by over 2,000 individual analysts for about 13,000 US firms. Data on daily returns, share prices and shares outstanding are taken from the Center for

Research in Security Prices (CRSP).

We focus on one-quarter ahead forecasts in the I/B/E/S file between 1988–I and 1995–IV for all firms with fiscal year-end in December that are followed over that period by at least 2 analysts, have actual earnings available and have daily returns on the CRSP tapes. We choose one-quarter ahead forecasts because there are more observations per firm and fewer forecast revisions than longer-term forecasts. We do not consider forecasts prior to 1988, because the lag between the date of an analyst’s forecast and the date the forecast was entered in the I/B/E/S database was significant. O’Brien (1988) reported an average publication lag over the period 1975 to mid-1982 of 34 trading days. For our sample, improvements in technology have reduced this turn-around time to less than 24 hours (see I/B/E/S Research Bibliography). Our sample selection ensures that the publication dates are close to the actual dates at which analysts released their forecasts; and, importantly, the last analyst *observes* the earlier forecasts.

This leaves us with a sample size of 237 firms with 7584 firm-quarters. Analysts sometimes revise their initial forecasts for a firm-quarter before the actual earnings are announced by the firm. We use only the last forecast reported by an analyst in any given firm-quarter, so the number of forecasts in a firm-quarter equals the number of analysts following the firm in that quarter.

Our theory imposes predictions only on the last analyst. If multiple analysts provide forecasts on the last day, we both do not know which one was ‘last’, nor what their information is. For firms followed by many analysts, the forecasts of analysts on the last date are unlikely to be affected significantly by beliefs about what other analysts are reporting on that date. Still, to avoid taking stands, we restrict our focus to firm-quarters where a single forecast was reported on the last date. This reduces the number of observations (firm-quarters) by 1702. However, it is important to note that our empirical findings are qualitatively unaffected if we include observations with multiple forecasts on the last day, and interpret the last forecast as the mean of those forecasts.

For the purpose of analyzing market reaction, we require all dates of announcements be known and in the required order — publication dates of forecasts must precede the date of announcement of actual EPS. For some firm-quarters, the date of announcement of actual EPS is either missing or the publication date of the last forecast is later than date of release of actual EPS. In either case, we delete that firm-quarter. Finally, some observations were dropped due to apparent data-entry errors. Using the (leverage points) method of detecting influential observations (see Davidson and MacKinnon (1993, pp.32)), we deleted six observations believed to be errors in data entry.<sup>14</sup> Table 1

---

<sup>14</sup>For example, in the third quarter of 1988, Pan Energy Corporation reported an EPS of \$1.98 whilst the mean forecast from the I/B/E/S file for the same firm-quarter was \$50.35. Similar data entry errors were reported by

summarizes the final sample and shows the type of industries/sectors represented in the sample.

Table 2 gives the distribution of analysts, forecasts and the average number of forecasts per analyst across all firms. An average of 9 analysts follow a firm each quarter, and the average number of forecasts per analyst is about 1.31, suggesting that analysts infrequently revise their one-quarter ahead forecasts. Although the distributions of analysts and forecasts between quarters are similar, on average the number of analysts and forecasts per firm-quarter increases slightly from the first to the fourth quarter — more forecasts are reported as the fiscal year-end approaches. Forecasts are also more likely to be revised as the fiscal year-end approaches.

#### 4.1 Hypotheses Development

We now develop our testable hypotheses regarding the direction of bias in last analyst’s forecast. The objective of our empirical work is not to estimate the relative compensation function, but rather to determine which form of compensation better describes analysts’ behavior. Once the nature of the relative compensation is determined several empirical implications of the effect of relative compensation on analysts’ forecasts can be explored.

The empirical analysis uses only the last forecast reported by each analyst in a firm-quarter. In practice, an analyst may receive new information at any time and revise his forecast upon the arrival of new information. We focus only on an analyst’s last forecast, because more recent forecasts contain more information and are more precise than older forecasts (O’Brien (1988)). In any event, quarterly forecasts are rarely revised so this procedure should not alter the results. The empirical issue is what is the “appropriate” consensus forecast with which the last analyst’s forecast is being compared. Choosing the wrong measure for “consensus” only reduces our chances of finding that the last forecast is systematically affected by the outstanding consensus. Using the mean of the forecasts of other analysts to construct the consensus seemed most appropriate because the mean forecast is highlighted in the I/B/E/S database and by other researchers.

Our analysis shows that if  $R$  is linear in relative performance, then all forecasts are independent of each other — analysts report their true expectation. For firm quarter  $\tau$ , let  $f_{L\tau}$  be the last forecast of the firm’s earnings reported by an analyst and denote the *outstanding mean forecast*, i.e. the mean of the forecasts excluding  $f_{L\tau}$ , by  $f_{m\tau}$ . The associated (independence/unbiased) hypothesis is

---

Kandel and Pearson (1995) and O’Brien (1988). For example, O’Brien deleted all entries with forecast errors larger than \$10.00 in absolute value.

Table 1: Sample Selection

	Number of Firms	Firm-quarters
1. Firms with at least 2 analysts following, fiscal year end in December and actual EPS available for all quarters in 1988–1995	279	8928
2. Firms in (1) with daily returns on CRSP tapes	237	7584
3. Firms-quarters in (2) with at least 2 analysts reporting their last forecast at different dates	237	7571
4. Sample after deleting observations with multiple forecasts on last forecast date (1702 deleted)	237	5869
5. Firm-quarters with publication date of last forecast in correct order (i.e., before date of actual EPS)	237	5842
6. Finally, six (6) influential observations deleted (possible errors in data entry)	237	5836
Sample of firms by Sectors		
Finance	34	
Health	13	
Consumer Non-Durables	20	
Consumer Services	19	
Consumer Durables	20	
Energy	22	
Transportation	10	
Technology	13	
Basic	41	
Capital/Machinery	22	
Utility	31	
—	—	
Total	237	

Table 2: Distribution of analysts and forecasts

		min	max	Percentiles			Sample moments	
				25	50	75	Mean	Std. dev.
Total sample: $N = 5836$	analysts	2	32	6	9	12	9	4.50
	forecasts	2	67	7	11	17	13	7.33
	fcts. per analyst	1	3.3	1.11	1.25	1.44	1.31	0.27
1st Qtr: $N = 1363^{\text{a}}$	analysts <sup>b</sup>	2	30	5	8	11	8	4.37
	forecasts	2	50	6	9	14	11	6.93
	fcts. per analyst	1	2.3	1.10	1.21	1.40	1.26	0.25
2nd Qtr: $N = 1528$	analysts	2	27	6	8	1	9	4.36
	forecasts	2	47	7	11	15	12	6.88
	fcts. per analyst	1	3.3	1.08	1.25	1.43	1.29	0.27
3rd Qtr: $N = 1402$	analysts	2	28	7	9	13	10	4.51
	forecasts	2	67	8	12	17	13	7.51
	fcts. per analyst	1	2.9	1.11	1.28	1.50	1.33	0.28
4th Qtr: $N = 1543$	analysts	2	32	7	10	14	10	4.47
	forecasts	2	47	9	13	19	14	7.52
	fcts. per analyst	1	2.8	1.14	1.28	1.50	1.35	0.28

a.  $N$  represents the number of observations (firm-quarters). All firm-quarters have at least two forecasts reported by distinct analysts on different dates.

b. Forecasts is the total number of forecasts reported by all analysts in that firm-quarter.



$\mathbf{H}_I$ : All forecasts reflect analysts' information in an unbiased manner, and no correlation exists between the direction of error in  $f_{L\tau}$  and the "location" of  $f_{L\tau}$  with respect to  $f_{m\tau}$ .

To test  $\mathbf{H}_I$ , a simple frequency test can be employed: If the last forecast is not affected by the consensus, then the probability that the last forecast exceeds actual earnings should be not vary with whether or not the last forecast exceeded the public information consensus forecast. We can group the data into mutually exclusive events:

$$\begin{aligned} X_A &\equiv \theta_\tau > f_{L\tau} > f_{m\tau}; & X_B &\equiv \theta_\tau < f_{L\tau} < f_{m\tau}; & X &= X_A \cup X_B \\ Y_A &\equiv f_{m\tau} \leq \theta_\tau < f_{L\tau}; & Y_B &\equiv f_{L\tau} < \theta_\tau \leq f_{m\tau}; & Y &= Y_A \cup Y_B \\ Z_A &\equiv \theta_\tau > f_{m\tau} > f_{L\tau}; & Z_B &\equiv \theta_\tau < f_{m\tau} < f_{L\tau}; & Z &= Z_A \cup Z_B \end{aligned}$$

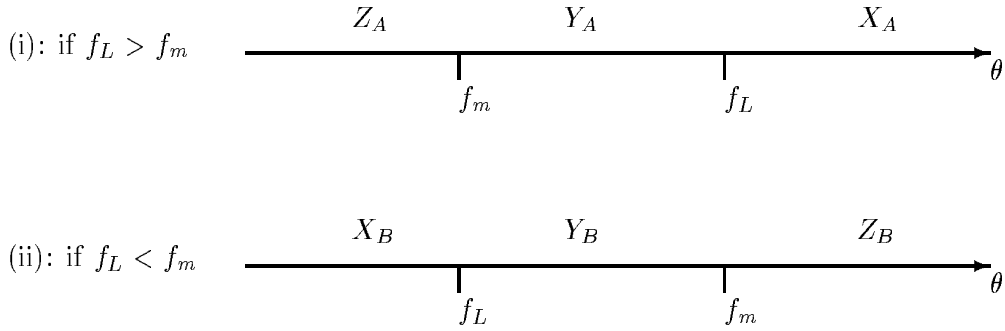


Figure 2: Realization of actual EPS,  $\theta$ , relative to  $f_L$  and  $f_m$

The description of these events is illustrated in Figure 2: the top panel shows the events  $X_A$ ,  $Y_A$  and  $Z_A$  when  $f_L > f_m$ , and the lower line segment depicts the events  $X_B$ ,  $Y_B$  and  $Z_B$  when  $f_L < f_m$ .

If the last forecast is unbiased, then

$$\Pr(\theta_\tau > f_{L\tau} | f_{L\tau} > f_{m\tau}) = \Pr(\theta_\tau > f_{L\tau} > f_{m\tau} | f_{L\tau} > f_{m\tau}) = \Pr(X_A | f_{L\tau} > f_{m\tau}) = 0.5$$

$$\Pr(\theta_\tau < f_{L\tau} | f_{L\tau} < f_{m\tau}) = \Pr(\theta_\tau < f_{L\tau} < f_{m\tau} | f_{L\tau} < f_{m\tau}) = \Pr(X_B | f_{L\tau} < f_{m\tau}) = 0.5$$

$$\Pr(X = X_A \cup X_B) = 0.5.$$

To understand the frequency tests suppose that given the posterior distribution of the last analyst, the probability that actual EPS is less than the consensus,  $f_{m\tau}$ , is 0.42. Lemma 2 ensures

that if the last analyst believes that actual EPS is more likely to exceed the consensus, then so will his forecast, implying that the event  $Z = Z_A \cup Z_B$  should occur 42% of the time. An upper bound on the frequency of event  $Z$  should be 0.5, which will hold when the last analyst has no additional information so that the consensus forecast always corresponds to the last analyst's expectation of earnings. The frequency with which event  $Z$  occurs captures the quality of the consensus forecast — the closer to 0.5, the better is the consensus. If  $R(\cdot)$  is a linear function of relative performance, then the last analyst's forecast should be equally likely to exceed  $\theta_\tau$  (as it is to fall short) if it exceeds the consensus forecast: that is, we should not be able to reject the hypothesis that  $\Pr(X) = 0.5$ .

Additional evidence can be gleaned from fixed-effects regressions: If the last forecast is unbiased,  $(f_{L\tau} - \theta_\tau)$  should not be systematically related to  $(f_{L\tau} - f_{m\tau})$ .

If, instead,  $R(\cdot)$  is a concave, or anti-symmetric concave function of relative performance, then the last forecaster will bias his forecast further toward  $f_{m\tau}$ . In this case, if  $f_{L\tau} > f_{m\tau}$ , the last forecaster's expectation of  $\theta_\tau$  exceeds  $f_{L\tau}$ , so that more often than not,  $\theta_\tau$  should exceed  $f_{L\tau}$ . Conversely, if  $f_{L\tau} < f_{m\tau}$ , more often than not,  $\theta_\tau$  should fall short of  $f_{L\tau}$ . Our joint hypothesis for a concave relative compensation function will be given by

- H<sub>c</sub>** :
- (i)  $\Pr(\theta_\tau > f_{L\tau} | f_{L\tau} > f_{m\tau}) = \Pr(\theta_\tau > f_{L\tau} > f_{m\tau} | f_{L\tau} > f_{m\tau}) = \Pr(X_A | f_{L\tau} > f_{m\tau}) > 0.5$
  - (ii)  $\Pr(\theta_\tau < f_{L\tau} | f_{L\tau} < f_{m\tau}) = \Pr(\theta_\tau < f_{L\tau} < f_{m\tau} | f_{L\tau} < f_{m\tau}) = \Pr(X_B | f_{L\tau} < f_{m\tau}) > 0.5$
  - (iii)  $\Pr(X = X_A \cup X_B) > 0.5$ .
  - (iv)  $(f_{L\tau} - \theta_\tau)$  will be negatively correlated with  $(f_{L\tau} - f_{m\tau})$ .

Finally, if  $R(\cdot)$  is a convex, or anti-symmetric convex function of relative performance, then the last forecaster will bias his forecast past  $\theta_\tau$ , away from  $f_{m\tau}$ , reducing the frequency with which the last forecaster undershoots. That is, if  $f_{L\tau} > f_{m\tau}$ , then  $f_{L\tau}$  exceeds the last forecaster's expectation of  $\theta_\tau$  so that more often than not,  $\theta_\tau$  should fall short of  $f_{L\tau}$ . Conversely, if  $f_{L\tau} < f_{m\tau}$ , more often than not,  $\theta_\tau$  should exceed  $f_{L\tau}$ . If analysts' payoffs are convex in relative performance, or driven by some other incentives that induce over-emphasizing of own information (and/or under-utilize the information in other forecasts), then the following (alternative) hypothesis holds:

- H<sub>v</sub>** :
- (i)  $\Pr(\theta_\tau > f_{L\tau} | f_{L\tau} > f_{m\tau}) = \Pr(\theta_\tau > f_{L\tau} > f_{m\tau} | f_{L\tau} > f_{m\tau}) = \Pr(X_A | f_{L\tau} > f_{m\tau}) < 0.5$
  - (ii)  $\Pr(\theta_\tau < f_{L\tau} | f_{L\tau} < f_{m\tau}) = \Pr(\theta_\tau < f_{L\tau} < f_{m\tau} | f_{L\tau} < f_{m\tau}) = \Pr(X_B | f_{L\tau} < f_{m\tau}) < 0.5$
  - (iii)  $\Pr(X = X_A \cup X_B) < 0.5$ .
  - (iv)  $(f_{L\tau} - \theta_\tau)$  will be positively correlated with  $(f_{L\tau} - f_{m\tau})$ .

## 4.2 Tests

### Evaluating Bias in last forecast: Frequency tests

We first perform the simple frequency tests of whether the direction of error in the last forecast is independent of the “location” of the last forecast relative to the outstanding mean. There are several important advantages of these tests:

- They are non-parametric and unrelated to the scale of errors across firms: no assumption is made about the structure of the relationship between the error in the last forecast and the difference in  $f_{L\tau}$  and  $f_{m\tau}$ .
- Outliers do not have disproportionate effects on frequencies. Consequently, the tests are robust to failing to exclude unusual “one-time” events such as large discretionary write-downs of assets that analysts do not seek to predict (see, for example, the concerns detailed in Keane and Runkle (1998)).

The flip side to the robustness of the frequency test to outliers is that we must distinguish small differences in forecasts. While earnings per share are continuously distributed, in practice, both earnings per share and forecasts by analysts are reported in cents, so that some unobserved rounding occurs. This creates a problem with inference whenever the last forecast equals the consensus or reported EPS. Without rounding, both of these events should happen with probability zero. With rounding, we cannot discern whether the last analyst’s expectation exceeds the consensus or not; and whether it exceeds true EPS or not. Consequently, for our frequency tests, we must drop an observation when either of these events occurs, reducing our sample size to  $N = 5234$ .

Table 3 presents the unconditional and conditional relative frequencies with which the last forecast exceeds the actual earnings per share:

- *On average, the last forecaster does not appear to make systematically overly-optimistic forecasts in order to curry favor with the firm that he follows.* In the sample considered,  $\Pr(f_{L\tau} > \theta_\tau) = 0.43$ .
- The probability that the last analyst’s forecast exceeds the actual EPS is greater for smaller firms<sup>15</sup>, and for firms that are followed by few analysts.

---

<sup>15</sup>Firm size equals share price at the time of the last forecast times outstanding shares.

Table 3: Bias Test for Last Forecast

Sample	$\Pr(f_{L\tau} > \theta_\tau)$	$\Pr(f_{L\tau} > \theta_\tau   f_{L\tau} < f_{m\tau})$	$\Pr(f_{L\tau} > \theta_\tau   f_{L\tau} > f_{m\tau})$
Panel A: Entire Sample			
Total	0.43 [0.42, 0.44]	0.32 [0.30, 0.34]	0.56 [0.54, 0.58]
Panel B: Segmented by Quarters			
1st Qtr.	0.39 [0.36, 0.42]	0.27 [0.24, 0.31]	0.52 [0.48, 0.56]
2nd Qtr.	0.43 [0.40, 0.46]	0.34 [0.30, 0.37]	0.54 [0.50, 0.58]
3rd Qtr.	0.43 [0.40, 0.46]	0.33 [0.29, 0.36]	0.56 [0.52, 0.61]
4th Qtr.	0.47 [0.45, 0.50]	0.34 [0.31, 0.37]	0.62 [0.58, 0.66]
Panel C: Segmented by Number of Analysts <sup>a</sup>			
$n \in [2, 6]$	0.47 [0.44, 0.49]	0.36 [0.32, 0.39]	0.59 [0.55, 0.63]
$n \in [7, 10]$	0.45 [0.42, 0.47]	0.34 [0.31, 0.37]	0.58 [0.54, 0.61]
$n \geq 11$	0.39 [0.37, 0.41]	0.28 [0.26, 0.31]	0.52 [0.49, 0.56]
Panel D: Segmented by Firm-Size <sup>b</sup>			
$\leq 25\%$	0.46 [0.43, 0.49]	0.35 [0.31, 0.38]	0.61 [0.57, 0.65]
25% – 75%	0.43 [0.41, 0.45]	0.33 [0.31, 0.36]	0.55 [0.52, 0.58]
$> 75\%$	0.40 [0.37, 0.42]	0.28 [0.25, 0.32]	0.53 [0.49, 0.57]

NOTE:–

Table provides the unconditional and conditional relative frequencies of the direction of error in the last forecast. 95% confidence intervals are reported in square brackets.

a.  $n$  denotes that number of analysts following the firm. For example,  $n \in [2, 6]$  implies that forecasts for the firm's quarterly EPS were provided by 2 to 6 analysts, and so on.

b. Firm size is calculated by multiplying share price (at the time of publication of the last forecast) by the number of shares outstanding, in a firm-quarter.

- The probability that the last analyst’s forecast exceeds the actual EPS is higher in the fourth quarter than the first.

One must be careful interpreting the observation that unconditionally, that the last forecast tended to fall short of true earnings per share, especially for large firms (which also tend to be followed by many analysts). Over this period stock prices, especially stock prices of large firms, climbed substantially. The most likely reason that the last forecast tended to fall short of true earnings per share, is that analysts (and investors) were (pleasantly) surprised by the high earnings growth experienced over this time period. This consistent pleasant surprise is reflected in the rise in stock prices.<sup>16</sup>

Our central finding that the last analyst’s forecast is severely biased is highlighted by:

- *With very high probability, the last analyst’s forecast overshoots the true level of earnings in the direction **away** from the consensus forecast:  $Pr(f_{L\tau} > \theta_\tau | f_{L\tau} > f_{m\tau}) = 0.56$ , and  $Pr(f_{L\tau} < \theta_\tau | f_{L\tau} < f_{m\tau}) = 0.68$ .*
- *This qualitative pattern of overshooting in the direction away from the consensus continues to hold when we condition on financial quarter, firm size and analyst coverage.*

This is overwhelming evidence that the last analyst’s forecast is influenced by the “location” of the consensus forecast and is consistent only with convex or anti-symmetric convex relative performance compensation. The last analyst’s forecast exhibits a contrarian bias **both** when he has positive and negative private information relative to the consensus: If his forecast exceeds the consensus, then it tends to exceed actual earnings; and if his forecast is less than the consensus then it generally falls short of actual earnings. It is very important to recognize that this bias always holds, and is essentially unrelated to the fact that in the sample, forecasts on average fall short of earnings. Indeed, even though on average, the last forecast exceeds true earnings only 43% of the time, conditional on exceeding the consensus, it exceeds true earnings 56% of the time. When we condition further on firm size or analyst following, the mean probability that the last forecast will exceed true earnings is shifted up or down, but the qualitative bias properties are also preserved: There are parallel shifts in the conditional probabilities of overshooting earnings in the direction away from the consensus.

---

<sup>16</sup>Analyses on a year by year basis appear to confirm this hypothesis. For the years 1989–1992 where stock prices rose less, one cannot reject the null that  $Pr(f_{L\tau} > \theta) = 0.5$  In the other years, the forecasts consistently fell short of actual earnings.

One can also gain insights by looking at the relative frequencies of the mutually exclusive events  $X$ ,  $Y$  and  $Z$  defined in Section 4.1. The information is, of course, contained in Table 3, but looking at  $X = X_A \cup X_B$  directly offers one measure to control for the variations in the sub-populations in the unconditional probability that the last forecast exceeds EPS. This eases comparisons of the bias across groups. The frequency with which these events occur are given in Table 4. The results show that  $\Pr(X = X_A \cup X_B) = 0.37$ , and that there is remarkably little variation when we further condition on financial year, the number of analysts following the firm or firm size. While variations in the number of analysts following a stock shift up or down the mean probability that the consensus exceeds actual EPS, and hence the frequency with which event  $X_A$  occurs, there is a virtually equal offsetting shift in the frequency with which event  $X_B$  occurs.

Were the last forecast unbiased event  $X$  should occur half the time: That event  $X$  occurs so infrequently provides overwhelming evidence that relative performance compensation matters, and that it takes a convex (or anti-symmetric convex) form. The last analyst clearly “over-emphasizes” his own information in his forecast and/or ignores information in earlier forecasts to which they have access. Of course our empirical findings have nothing to say about analysts other than the last, and it may be that it is those analysts with the most convex relative compensation functions who choose to report later because it is easier to distinguish their forecast from the consensus if they are last.

*This evidence is particularly striking when we recognize that the obvious potential sources of bias tend to lead to clustering of forecasts; that is, the event  $X$  occurring **more frequently** — **even more than half of the time**. Other research has highlighted the possibility that analysts tend to issue consistently overly-optimistic forecasts in order to retain informational access to firm management. This would lead both the consensus forecast and the last forecast to be high, too often, making it **more** likely that both are on the same side of EPS, so that the event  $X$  would occur *more* frequently. In fact, in our sample, the opposite result obtains — the consensus and last forecast are typically too low. This too would tend to increase the frequency with which  $X$  occurs, but the bias in the last forecast overwhelms this effect: event  $X$  occurs only 37% of the time.*

A second, almost surely significant factor that would lead to clustering is that analysts receive much of their information from the same source — firm management — so if there is any error in what firm management reports to analysts, this will lead to correlation in signal errors, and hence to clustering; hence concave relative performance compensation should be difficult to reject. To see this suppose that analyst  $j$  sees the signal  $s_{j\tau} = \theta_\tau + \nu_\tau + \epsilon_{j\tau}$ , where  $\nu_\tau$  is a common error to the report. Then again, if  $\nu_\tau$  is positive (negative) this will lead to both the consensus and last forecast

Table 4: Frequency of events  $X$ ,  $Y$  and  $Z$

Sample	$\Pr(Z = Z_A \cup Z_B)$	$\Pr(Y = Y_A \cup Y_B)$	$\Pr(X = X_A \cup X_B)$
Total	0.42	0.21	0.37
Panel B: Segmented by Number of Analysts <sup>a</sup>			
$n \in [2, 6]$	0.41	0.21	0.38
$n \in [7, 10]$	0.42	0.20	0.38
$n \geq 11$	0.43	0.20	0.37
Panel C: Segmented by Firm-Size <sup>b</sup>			
$\leq 25\%$	0.40	0.23	0.37
$25\% - 75\%$	0.41	0.21	0.38
$> 75\%$	0.44	0.19	0.37
Panel D: Segmented by Financial Year			
1988	0.42	0.20	0.38
1989	0.38	0.24	0.38
1990	0.43	0.22	0.35
1991	0.41	0.23	0.36
1992	0.41	0.19	0.40
1993	0.45	0.17	0.38
1994	0.43	0.20	0.37
1995	0.41	0.21	0.38

Table provides relative frequencies of the events  $X$ ,  $Y$  and  $Z$  defined in Section 4.1.

**a.**  $n$  denotes that number of analysts following the firm. For example,  $n \in [2, 6]$  implies that forecasts for the firm's quarterly EPS were provided by 2 to 6 analysts, and so on.

**b.** Firm size is calculated by multiplying share price (at the time of publication of the last forecast) by the number of shares outstanding, in a firm-quarter.

to be high (low), too often. That is, with correlated signals, one would expect that consensus and last forecast will be on the same side of the true earnings (with the last forecast closer to the true earnings) more frequently than not, implying that the event  $X$  will occur more frequently. So, too, if, for example, after all forecasts have been made, there is an unanticipated shock to earnings then forecasts will tend to be systematically clustered on the same side of earnings, implying that the event  $X$  would occur more frequently. These factors, for which we do not account, make it all the more surprising to find that the event  $X$  occurs so infrequently. Phrased differently, were we able to control for correlated signals or common unanticipated earnings shocks, the evidence that the last forecaster biases his forecast away from the consensus would only be further reinforced — evidence that strongly suggests some form of convex relative performance compensation.

Recall that event  $Z = Z_A \cup Z_B$ , where  $Z_{A\tau} \equiv \theta_\tau > f_{m\tau} > f_{L\tau}$  and  $Z_{B\tau} \equiv \theta_\tau < f_{m\tau} < f_{L\tau}$  occurs when the last analyst’s expectation turns out to be on the wrong side of the consensus relative to true EPS. The likelihood of event  $Z$  captures the quality of the outstanding mean forecast — Lemma 1 implies that event  $Z$  should occur no more than half the time, and the less frequently it occurs, the worse is the quality of the consensus report. Since  $Z$  occurs about 42% of the time, it appears that the consensus is of low quality. The frequency of event  $Z$  does not vary significantly with firm size or number of analysts, so we do not report them here. Finally, the relative frequency with which the two forecast measures are on the opposite sides of the actual earnings per share, event  $Y = Y_A \cup Y_B$ , is 0.21. If the consensus is relatively good, *i.e.* event  $Z$  is relatively likely, then unless the over-shooting bias is extreme, event  $Y$  should not occur “that frequently”. That event  $Y$  is so likely reflects both the relatively inaccurate consensus and an extreme over-shooting bias: Despite having access to the outstanding consensus forecast, the last analyst’s forecast is closer to true EPS only marginally more frequently than is the outstanding consensus.

### Fixed-effects Regressions

In this section, we offer an alternative characterization of the bias in last forecast using fixed-effects regression. In particular, we explore how the error in the last analyst’s report is related to the difference between the last and consensus forecast. Using the notation  $f_{Li\tau}$  and  $f_{mi\tau}$  to represent the last and outstanding mean forecasts for firm  $i$  in quarter  $\tau$ , we estimate a fixed-effects model of the form

$$(f_{Li\tau} - \theta_{i\tau}) = \beta_0 + \nu_i + \sum_{k \in \text{year}} \gamma_k + \beta_1(f_{Li\tau} - f_{mi\tau}) + \epsilon_{i\tau}, \quad (3)$$



where  $\nu_i$  represent individual firm dummies, to control for heterogeneity across firms, and  $\gamma_k$  is a year dummy. The results in Table 5 show a positive relationship between the error in the last forecast and the difference between the last and outstanding mean forecasts.<sup>17</sup> On average, if the last forecast exceeds the outstanding consensus by \$1.00, the last forecast over-shoots the actual earnings per share by \$0.43!

Table 5: Fixed-effects Regression

$$\text{Regression: } (f_{Lit} - \theta_{it}) = \beta_0 + \nu_i + \sum_{k \in \text{year}} \gamma_k + \beta_1(f_{Lit} - f_{mit}) + \epsilon_{it}$$

	$\beta_1$	$R^2$	$N$
Entire sample	0.445 (10.700)***	0.021	5836
Segmented by number of analysts			
$n \in [2, 6]$	0.518 (4.857)***	0.023	1710
$n \in [7, 10]$	0.350 (5.928)***	0.018	2026
$n \geq 11$	0.430 (6.984)***	0.039	2100
Segmented by firm size			
Small	0.841 (10.084)***	0.072	1459
Medium	0.107 (1.760)*	0.008	2918
Large	0.300 (5.646)***	0.031	1459

NOTE: t-statistics are given in parentheses: (\*), (\*\*) and (\*\*\*) indicate significant at 90%, 95% and 99% levels. Small firms are those in the lowest quartile by firm size; and large firms are those in the highest quartile. Firm size equals share price at of last forecast times shares outstanding. The constant term and estimates of the firm and year dummies are not reported here.

<sup>17</sup>In all but one of the sub-sample regression, a standard Hausman test does not reject the hypothesis that the error components specification is correct at the 5% level. A random effects GLS estimation, which is more efficient, give similar results as the fixed-effects estimates. We do not report the constant term and estimates of the firm and year dummies in Table 5.

## Returns around earnings announcement

The analysis in the last two sub-sections documents that the last analyst’s reported forecast over-emphasizes his private information: the forecast tends to over-shoot (under-estimate) the actual EPS if it is greater (less) than the outstanding consensus.

This raises the question: Are investors systematically fooled by the bias in the forecast of the last analyst? For example, a large positive deviation,  $(f_{L\tau} - f_{m\tau})$ , in the last forecast away from the consensus, may mislead investors who do not account for the bias in this forecast into over-estimating by how much earnings will go up. Investors who treat the last report as unbiased will over-react to their forecasts of earnings. Then, if the ‘correct’ information does not leak out until near the earnings announcement date, a large positive  $(f_{L\tau} - f_{m\tau})$  should be associated with subsequent disappointment and a downward revision in prices around the earnings announcement, and hence negative excess returns. Conversely, a large negative  $f_{L\tau} - f_{m\tau}$  should be associated with positive excess returns around the earnings announcement. In this section, we explore whether investors are systematically fooled by analyzing excess returns around the earnings announcement.

Following Imhoff and Lobo (1984), we measure the information content in the last forecast by the standardized forecast difference (SFD) in the two forecast measures given by

$$SFD = \frac{(f_{L\tau} - f_{m\tau})}{P_\tau},$$

where  $P_\tau$  is the price of the firm’s share at the *time of publication of the last forecast* in the firm-quarter  $\tau$ .<sup>18</sup> We rank the SFDs from the most negative to the most positive. Finally, we look at the percentage mean cumulative abnormal return (MCAR) over the different windows for portfolios of extreme positive and negative SFDs. In particular, we look at portfolios consisting of SFDs in the top decile (10%) and quintile (20%); and at portfolios of SFDs in the bottom decile and quintile. We do this for several reasons:

- The mean outstanding forecast is a noisy measure of the beliefs of investors prior to the last forecast. Hence, we are more confident of getting the direction of the revision “correct” for more extreme portfolios.
- The predicted bias in forecast is larger if there is a greater difference between the last and consensus forecast, so the degree to which naive investors would be fooled is greater for more extreme deviations in forecasts.

---

<sup>18</sup>This method, suggested by Christie (1987) as a way of reducing scale effects, was also adopted Lys and Sohn (1990) and Alexander and Ang (1997), among others.

Consequently, it is in these extreme portfolios that we are most likely to find any evidence that investors do not correctly unravel the bias in reports.

To obtain excess returns, we first estimate a market model for each firm–quarter:

$$r_{it} = \alpha_0 + \alpha_1 r_{mt} + \epsilon_t, \quad (4)$$

where  $r_{it}$  is the daily return on security  $i$  and  $r_{mt}$  is the return on the CRSP value weighted market index. For each firm–quarter, the market model regression (equation (4)) uses returns for the period  $t = -170$  through to  $t = -21$ , where time  $t = 0$  denotes the day of actual earnings announcement by the firm. Hence, each market model is estimated with 150 data points. The least squares estimates,  $\hat{\alpha}_0$  and  $\hat{\alpha}_1$ , are then used to calculate the cumulative abnormal returns (CAR) for different event windows around the earnings announcement date for each firm–quarter  $\tau$  :

$$CAR_{\tau} = \sum_{t=\underline{t}}^{\bar{t}} [r_{\tau t} - (\hat{\alpha}_0 + \hat{\alpha}_1 r_{mt})]. \quad (5)$$

where  $\underline{t}$  and  $\bar{t}$  are starting and end points of the period over which abnormal returns are cumulated.

In our analysis, it is important to exclude the market’s reaction to the information in the last forecast. For instance, if the last forecast exceeds the consensus, then this good news will be reflected in the price and lead to immediate positive excess returns; cumulating those immediate returns may cause us to fail to find evidence of bias, since the market’s failure to incorporate the bias leads to predicted negative excess returns. To disentangle the impact of new information from the reaction to the last forecast, we drop all observations for which the last forecast was not at least two days from the beginning of our event window. Our findings are not systematically affected by other choices.

Note that the empirical regularity for earnings surprise is that, on average, daily abnormal returns evolve monotonically throughout a long event window (see e.g. Foster, Olsen and Shevlin (1984)). Since we are searching for evidence that investors are systematically fooled — which implies a non–monotonic relationship in abnormal returns as investors over–react to the last forecast — we are “battling” the standard empirical regularity.

The empirical regularities regarding percentage mean cumulative abnormal returns presented in Table 6 are qualitatively similar to those that obtain for other windows. Qualitatively, the following empirical regularities emerge:

- Aggregating across numbers of analysts following a firm

Table 6: Cumulative Abnormal Returns: Percentage mean

Percentage mean cumulative abnormal returns ( $\text{MCAR}_{[t, \bar{t}]}$ ) cumulated over periods  $[t, \bar{t}]$  for portfolios classified by smallest and largest SFD. Date of earnings announcement is  $\bar{t} = 0$ .

Window	Analysts ( $n$ )	Portfolio by SFD	MCAR (%)	t-stat	$N$	
[-2, 0]	All	Bottom 10%	0.561**	2.368	469	
		Top 10%	0.314*	1.755	470	
	$n \in [2, 6]$	Bottom 10%	0.445	1.301	156	
		Top 10%	-0.073	-0.255	157	
	$n \in [7, 10]$	Bottom 10%	0.925**	2.144	166	
		Top 10%	1.035***	3.224	167	
	$n \geq 11$	Bottom 10%	0.266	0.589	146	
		Top 10%	0.258	0.761	147	
	[-2, 1]	All	Bottom 10%	0.554**	2.026	469
			Top 10%	0.321	1.585	470
$n \in [2, 6]$		Bottom 10%	0.799*	1.846	156	
		Top 10%	-0.012	-0.038	157	
$n \in [7, 10]$		Bottom 10%	0.737	1.608	166	
		Top 10%	1.088***	3.087	167	
$n \geq 11$		Bottom 10%	0.124	0.231	146	
		Top 10%	0.176	0.444	147	
[-2, 0]		All	Bottom 20%	0.264*	1.751	939
			Top 20%	0.190	1.552	940
	$n \in [2, 6]$	Bottom 20%	0.438*	1.821	313	
		Top 20%	-0.119	-0.601	314	
	$n \in [7, 10]$	Bottom 20%	0.278**	1.076	332	
		Top 20%	0.495*	1.954	333	
	$n \geq 11$	Bottom 20%	0.096	0.334	293	
		Top 20%	0.403	1.781	294	
	[-2, 1]	All	Bottom 20%	0.177**	1.009	939
			Top 20%	0.138	1.008	940
$n \in [2, 6]$		Bottom 20%	0.722**	2.500	313	
		Top 20%	-0.101	-0.456	314	
$n \in [7, 10]$		Bottom 20%	0.037	0.131	332	
		Top 20%	0.333	1.442	333	
$n \geq 11$		Bottom 20%	-0.263	-0.777	293	
		Top 20%	0.144	0.537	294	

NOTE:- (\*), (\*\*) and (\*\*\*) indicate significant at 90%, 95%, and 99% levels.  $\text{SFD} = \frac{f_{L,t} - f_{M,t}}{P}$  is standardized forecast difference.  $P$  is stock price at time of publication of last forecast.

- For the SFD decile portfolios, the MCAR for the lowest SFD is significantly greater than zero, but the MCAR for the highest SFD was not. While the percentage mean excess return was always greater for the lowest SFD decile than the highest, which is consistent with investors being systematically fooled by the last analyst, the difference is not significant.
- For the SFD quintile portfolios, aggregating across numbers of analysts following a firm, the MCARs were generally insignificantly different from zero.
- For firms followed by fewer analysts (2 to 6 analysts)
  - The MCARs for the lowest SFD were even greater, always significantly different from zero; and the MCARs for the highest SFD were negative, but insignificantly different from zero.
  - The difference in MCARs for the lowest SFD and highest SFD was statistically different from zero for the (larger) quintile portfolios, but not for most of the (smaller) decile portfolios.
- For firms followed by more than six analysts,
  - The MCARs were generally not significantly different from zero; and, in fact, in a majority of windows the MCARs for the portfolios of the highest SFDs slightly exceeded those of the lowest SFDs.

Thus, if few analysts follow a firm, there is some evidence consistent with the hypothesis that investors are “fooled” by the last analyst’s forecast; in particular, investors may attach undue weight to an especially negative forecast. There is no evidence that investors are “fooled” if more analysts follow the firm. One interpretation of these findings is that investors have fewer sources of information about firms that are followed by few analysts, so that the impact of a biased last forecast is greater. However, the supporting evidence is far from compelling.

## 5 Conclusion

This paper presents a compensation-based model to explain the strategic bias in individual analysts’ forecast of earnings per share (EPS). When analysts are compensated both on the basis of the accuracy of their forecasts and how the accuracy of their forecasts compare with other ana-

lysts, then earlier forecasts affect later announcements. For different forms of relative performance compensation, we derive the implications for the pattern of forecasts.

We find overwhelming evidence that the last forecast is not an unbiased reflection of that analyst’s information: The last analyst to report a forecast of a firm’s EPS reports a strongly contrarian forecast, biasing his forecast *away* from the consensus forecast — 62% of the time, the last forecast overshoots actual EPS in the direction away from the outstanding mean forecast. This is consistent with some form of convex relative performance and implies that a better estimate of EPS than either the last forecast or the average of all forecasts could be obtained by accounting for the bias in the last forecast (even just by averaging the last and consensus forecasts).

We then test to see whether investors correctly unravel the strategic bias in the last forecast. Excess returns cumulated around the earnings announcement date offer modest evidence that for firms followed by few analysts, the final forecast may “fool” investors.

## Appendix

### Proof of Lemma 1:

If compensation is linear in relative performance, then an analyst’s compensation can be written as

$$w_j = \bar{w} - (1 + \lambda)|f_j - \theta| + |f_m - \theta|.$$

Consider the last analyst to report his forecast. Since the  $(L-1)$  preceding forecasts have been announced,  $f_L$  will not affect  $f_m$ , the mean of all other forecasts. Hence, maximizing  $E(w_L|\hat{\theta}_L)$  is equivalent to minimizing

$$E \left[ |f_L - \theta| \mid \hat{\theta}_L \right] = \int_{\theta} (|f_L - \theta|) dG(\theta|\hat{\theta}_L)$$

To show that the optimal forecast is  $f_L = \hat{\theta}_L$ , we prove that the expected forecast error is greater if the last analyst reports any other forecast.

Suppose he reports  $f_L = \hat{\theta}_L - \xi$  (a lesser forecast), where  $\xi > 0$ . Then the difference in expected errors is

$$\begin{aligned} \int_{\theta} (|\hat{\theta}_L - \theta| - |\hat{\theta}_L - \xi - \theta|) dG(\theta|\hat{\theta}_L) &= \int_{\theta \leq \hat{\theta}_L - \xi} \xi dG(\theta|\hat{\theta}_L) + \int_{\hat{\theta}_L - \xi}^{\hat{\theta}_L} (2\hat{\theta}_L - 2\theta - \xi) dG(\theta|\hat{\theta}_L) \\ &\quad - \int_{\theta > \hat{\theta}_L} \xi dG(\theta|\hat{\theta}_L) \\ &\leq \int_{\theta \leq \hat{\theta}_L} \xi dG(\theta|\hat{\theta}_L) - \int_{\theta > \hat{\theta}_L} \xi dG(\theta|\hat{\theta}_L) \end{aligned} \quad (6)$$

since  $\theta > \hat{\theta}_L - \xi$  in the second term on the RHS of (6), which implies that  $2\hat{\theta}_L - 2\theta - \xi \leq \xi$ . By the symmetry of  $g(\theta|\hat{\theta}_L)$  about  $\hat{\theta}_L$ , the last term in (6) vanishes. Hence,

$$\int_{\theta} |\hat{\theta}_L - \theta| dG(\theta|\hat{\theta}_L) \leq \int_{\theta} |\hat{\theta}_L - \xi - \theta| dG(\theta|\hat{\theta}_L)$$

Similarly,  $\forall \xi > 0$ , we get  $\int_{\theta} (|\hat{\theta}_L - \theta|) dG(\theta|\hat{\theta}_L) < \int_{\theta} (|\hat{\theta}_L + \xi - \theta|) dG(\theta|\hat{\theta}_L)$ . The last analyst has a larger expected forecast error if he announces a forecast greater than his posterior expectation. Hence the optimal forecast is  $f_L = \hat{\theta}_L$ .

Now consider the  $(L-1)$ th analyst. Since the  $(L-2)$  preceding forecasts have been announced and the last analyst will report his true signal anyway<sup>19</sup>, the forecast of the  $(L-1)$ th analyst will not affect the mean of the other forecasts: that is,  $\frac{\partial f_m}{\partial f_{(L-1)}} = 0$ . Hence, maximizing  $E(w_{L-1}|\hat{\theta}_{L-1})$  is equivalent to minimizing

$$E(|F_{L-1} - \theta| | \hat{\theta}_{L-1}).$$

Repeating the analysis above, it is straightforward to show that the optimal forecast of the  $(L-1)$ th analyst will then be  $f_{L-1} = \hat{\theta}_{L-1}$ . By backward induction, we get  $F_j^* = \hat{\theta}_j, \forall j = 1, 2, \dots, L$ .  $\square$

### Proof of Lemma 2:

Given his posterior beliefs, the last analyst's expected payoff will be given by

$$E \left[ w_L(f_L) | \hat{\theta}_L \right] = \bar{w} - \lambda \int_{\theta} (|f_L - \theta|) dG(\theta|\hat{\theta}_L) + \int_{\theta} R(|f_m - \theta| - |f_L - \theta|) dG(\theta|\hat{\theta}_L).$$

Suppose  $f_m < \hat{\theta}_L$ . We first show that  $f_L \not\leq f_m$ . To do this, we prove that for every forecast  $f_L < f_m$ , there exists some other  $f_L > f_m$  that yields a higher payoff. Consider the following two reports by the last analyst:  $f_L = f_m + \xi$  and  $f_L = f_m - \xi$ , where  $\xi > 0$ . The difference in payoffs will be given by

$$\begin{aligned} E \left[ w_L(f_m + \xi) - w_L(f_m - \xi) | \hat{\theta}_L \right] &= \int_{\theta \leq f_m - \xi} (-2\lambda\xi + R(-\xi) - \\ &R(\xi)) dG(\theta|\hat{\theta}_L) + \int_{f_m - \xi}^{f_m} (2\lambda(\theta - f_m) + R(-\xi) - R(2f_m - 2\theta - \xi)) dG(\theta|\hat{\theta}_L) \\ &+ \int_{f_m}^{f_m + \xi} (2\lambda(\theta - f_m) + R(2\theta - 2f_m - \xi) - R(-\xi)) dG(\theta|\hat{\theta}_L) \\ &+ \int_{\theta \geq f_m + \xi} (2\lambda\xi + R(\xi) - R(-\xi)) dG(\theta|\hat{\theta}_L) \end{aligned} \quad (7)$$

Clearly, the sum of the first and fourth terms on the RHS of (7) is positive by the symmetry of  $g(\theta|\hat{\theta}_L)$  about  $\hat{\theta}_L$ . Secondly, the third term is at least equal to the second term since  $g(\theta|\hat{\theta}_L)$  is single-peaked at  $\hat{\theta}_L$ .

Therefore,

$$E \left[ w_L(f_m + \xi) - w_L(f_m - \xi) | \hat{\theta}_L \right] > 0$$

and  $f_m + \xi$  dominates  $f_m - \xi$ ,  $\forall \xi > 0$ .

It is also straightforward to show that  $f_L \neq f_m$ : that is, it is not optimal for the last analyst to mimic the mean of earlier forecasts by reporting  $f_L = f_m$ . Because,

$$E \left[ w_L(f_m + \epsilon) - w_L(f_m) | \hat{\theta}_L \right] = \int_{\theta \leq f_m} (-\lambda\epsilon + R(-\epsilon)) dG(\theta|\hat{\theta}_L)$$

---

<sup>19</sup>By assumption,  $f_{L-1}$  will not affect  $\hat{\theta}_L$  since analysts cannot mislead others.

$$\begin{aligned}
& + \int_{f_m}^{f_m+\epsilon} (\lambda(2\theta - 2f_m - \epsilon) + R(2\theta - 2f_m - \epsilon)) dG(\theta|\hat{\theta}_L) \\
& + \int_{\theta > f_m+\epsilon} (\lambda\epsilon + R(\epsilon)) dG(\theta|\hat{\theta}_L).
\end{aligned}$$

Dividing through by  $\epsilon$  and taking the limit as  $\epsilon \rightarrow 0$ , we get the derivative of  $w_L(f_L)$  at  $f_L = f_m$  to be

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_m + \epsilon) - w_L(f_m)}{\epsilon} \middle| \hat{\theta}_L \right] = - \int_{\theta \leq f_m} (\lambda + R'(0)) dG(\theta|\hat{\theta}_L) + \int_{\theta > f_m} (\lambda + R'(0)) dG(\theta|\hat{\theta}_L)$$

Using the symmetry of  $g(\theta|\hat{\theta}_L)$  about  $\hat{\theta}_L$ , we get derivative of  $w_L(f_L)$  w.r.t  $f_L$  at  $f_L = f_m$  to be

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_m + \epsilon) - w_L(f_m)}{\epsilon} \middle| \hat{\theta}_L \right] = 2 \int_{f_m}^{\hat{\theta}_L} (\lambda + R'(0)) dG(\theta|\hat{\theta}_L) > 0.$$

That is  $w_L(f_L)$  is increasing at  $f_L = f_m$ . Hence,  $f_L > f_m$  if  $\hat{\theta}_L > f_m$ . The proof that  $f_L < f_m$  if  $\hat{\theta}_L < f_m$  is analogous.  $\square$

### Proof of Proposition 1:

(i) Let  $f_m = \hat{\theta}_L$ . Then

$$\begin{aligned}
E \left[ w_L(\hat{\theta}_L + \epsilon) - w_L(\hat{\theta}_L) \middle| \hat{\theta}_L \right] & = - \int_{\theta \leq \hat{\theta}_L} (-\lambda\epsilon + R(-\epsilon)) dG(\theta|\hat{\theta}_L) \\
& + \int_{\hat{\theta}_L}^{\hat{\theta}_L+\epsilon} (-\lambda(2\hat{\theta}_L - 2\theta + \epsilon) + R(2\theta - 2\hat{\theta}_L - \epsilon)) dG(\theta|\hat{\theta}_L) \\
& + \int_{\theta > \hat{\theta}_L+\epsilon} (\lambda\epsilon + R(\epsilon)) dG(\theta|\hat{\theta}_L). \tag{8}
\end{aligned}$$

Dividing equation (8) through by  $\epsilon$  and taking the limit as  $\epsilon \rightarrow 0$ , we get the derivative of  $w_L(f_L)$  at  $f_L = \hat{\theta}_L$  to be

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L + \epsilon) - w_L(\hat{\theta}_L)}{\epsilon} \middle| \hat{\theta}_L \right] & = - \int_{\theta \leq \hat{\theta}_L} (\lambda + R'(0)) dG(\theta|\hat{\theta}_L) + \int_{\theta > \hat{\theta}_L} (\lambda + R'(0)) dG(\theta|\hat{\theta}_L) \\
& = 0
\end{aligned}$$

Hence if  $f_m = \hat{\theta}_L$ , then the optimal  $f_L = \hat{\theta}_L$ .

(ii) Now suppose  $f_m < \hat{\theta}_L$ , then by Lemma 2,  $f_L > f_m$ . So it remains to show that  $f_L \notin [\hat{\theta}_L, \infty)$ . To do this, we show that  $w_L(f_L)$  is decreasing in  $f_L$  over  $[\hat{\theta}_L, \infty)$ . Consider  $f_L = \hat{\theta}_L + \xi$ , where  $\xi \geq 0$ . The derivative of  $w_L(f_L)$  at this point is given by

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L + \xi + \epsilon) - w_L(\hat{\theta}_L + \xi)}{\epsilon} \middle| \hat{\theta}_L \right] & = - \int_{\theta \leq f_m} (\lambda + R'(f_m - \hat{\theta}_L - \xi)) dG(\theta|\hat{\theta}_L) \\
& - \int_{f_m}^{\hat{\theta}_L+\xi} (\lambda + R'(2\theta - \hat{\theta}_L - \xi - f_m)) dG(\theta|\hat{\theta}_L) \\
& + \int_{\theta > \hat{\theta}_L+\xi} (\lambda + R'(\hat{\theta}_L + \xi - f_m)) dG(\theta|\hat{\theta}_L). \tag{9}
\end{aligned}$$



Consider the second term on the RHS of (9). Since  $\theta < (\hat{\theta}_L + \xi)$ , we have  $R'(2\theta - \hat{\theta}_L - \xi - f_m) \geq R'(\hat{\theta}_L + \xi - f_m)$  by the concavity of  $R(\cdot)$ . Also  $R'(f_m - \hat{\theta}_L - \xi) > R'(\hat{\theta}_L + \xi - f_m)$ . Therefore,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L + \xi + \epsilon) - w_L(\hat{\theta}_L + \xi)}{\epsilon} \middle| \hat{\theta}_L \right] &< - \int_{\theta \leq \hat{\theta}_L + \xi} (\lambda + R'(\hat{\theta}_L + \xi - f_m)) dG(\theta | \hat{\theta}_L) \\ &+ \int_{\theta > \hat{\theta}_L + \xi} (\lambda + R'(\hat{\theta}_L + \xi - f_m)) dG(\theta | \hat{\theta}_L). \end{aligned}$$

Exploiting the symmetry of  $g(\theta | \hat{\theta}_L)$  about  $\hat{\theta}_L$ , we get

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L + \xi + \epsilon) - w_L(\hat{\theta}_L + \xi)}{\epsilon} \middle| \hat{\theta}_L \right] < - \int_{\hat{\theta}_L - \xi}^{\hat{\theta}_L + \xi} (\lambda + R'(\hat{\theta}_L + \xi - f_m)) dG(\theta | \hat{\theta}_L) < 0,$$

$\forall \xi \geq 0$ . So the last analyst's payoff is decreasing in  $f_L \in [\hat{\theta}_L, \infty)$ . Combining this result with Lemma 2, it follows that the last analyst will "locate" between  $f_m$  and  $\hat{\theta}_L$ .

(iii) The proof that  $\hat{\theta}_L < f_L < f_m$  if  $\hat{\theta}_L < f_m$  is similar to part (ii).

(iv) Let  $f_m < \hat{\theta}_L$  and  $f_L$  be the optimal forecast of the second analyst given  $\hat{\theta}_L$  and  $f_m$ . Then it will suffice to show that if the consensus forecast were to be revised to  $f_m^* = f_m - \delta$ , ( $\delta > 0$ ), then the last analyst will report  $f_L^* \in (f_L - \delta, f_L)$ .

By Proposition 1, we know that  $f_L \in (f_m, \hat{\theta}_L)$ , and by the optimality of  $f_L(\hat{\theta}_L)$ ,  $\frac{\partial w_L}{\partial f_L} = 0$  at  $f_L(\hat{\theta}_L)$ . But

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_L + \epsilon) - w_L(f_L)}{\epsilon} \middle| \hat{\theta}_L \right] &= - \int_{\theta \leq f_m} (\lambda + R'(f_m - f_L)) dG(\theta | \hat{\theta}_L) \\ &- \int_{f_m}^{f_L} (\lambda + R'(2\theta - f_m - f_L)) dG(\theta | \hat{\theta}_L) \\ &+ \int_{\theta > f_L} (\lambda + R'(f_L - f_m)) dG(\theta | \hat{\theta}_L) = 0. \end{aligned}$$

Let  $H(f_L)$  be defined as

$$\begin{aligned} H(f_L) &= - \int_{\theta \leq f_m} (\lambda + R'(f_m - f_L)) dG(\theta | \hat{\theta}_L) - \int_{f_m}^{f_L} (\lambda + R'(2\theta - f_m - f_L)) dG(\theta | \hat{\theta}_L) \\ &+ \int_{\theta > f_L} (\lambda + R'(f_L - f_m)) dG(\theta | \hat{\theta}_L) = 0. \end{aligned} \tag{10}$$

Now, suppose the consensus forecast were  $f_m^* = f_m - \delta$  ( $\delta > 0$ ). Then the derivative of  $w_L$  at  $f_L$  will be

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_L + \epsilon) - w_L(f_L)}{\epsilon} \middle| \hat{\theta}_L, f_m^* \right] &= - \int_{\theta \leq f_m - \delta} (\lambda + R'(f_m - \delta - f_L)) dG(\theta | \hat{\theta}_L) \\ &- \int_{f_m - \delta}^{f_L} (\lambda + R'(2\theta - f_m + \delta - f_L)) dG(\theta | \hat{\theta}_L) + \int_{\theta > f_L} (\lambda + R'(f_L - f_m + \delta)) dG(\theta | \hat{\theta}_L). \end{aligned} \tag{11}$$

By a simple change of variable, we can write the RHS of equation (11) as

$$\begin{aligned}
& - \int_{\theta \leq f_m} (\lambda + R'(f_m - f_L - \delta)) dG(\theta | \hat{\theta}_L) & - \int_{f_m}^{f_L + \delta} (\lambda + R'(2\theta - f_m - f_L - \delta)) dG(\theta | \hat{\theta}_L) \\
& & + \int_{\theta > f_L + \delta} (\lambda + R'(f_L + \delta - f_m)) dG(\theta | \hat{\theta}_L)
\end{aligned}$$

which is  $H(f_L + \delta)$ . By the optimality of  $f_L$ ,

$$\lim_{\delta \rightarrow 0} E \left[ \frac{H(f_L + \delta) - H(f_L)}{\delta} \right] < 0$$

which implies that  $H(f_L + \delta) < H(f_L) = 0$ . Hence,

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_L + \epsilon) - w_L(f_L)}{\epsilon} \middle| \hat{\theta}_L, f_m^* \right] < 0$$

and the last analyst will shade his forecast in the direction of revision in the outstanding consensus.

Finally, it is straightforward to show that

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(f_L - \delta + \epsilon) - w_L(f_L - \delta)}{\epsilon} \middle| \hat{\theta}_L, f_m^* \right] > 0$$

which implies that the last analyst will bias his forecast by less than  $\delta$  in response to the new mean. That is,  $0 < \frac{\partial(\hat{\theta}_L - f_L)}{\partial(\hat{\theta}_L - f_m)} < 1$ .  $\square$

### Proof of Proposition 2:

(i) Suppose  $f_m < \hat{\theta}_L$ , then by Lemma 2,  $f_L > f_m$ . So we need to show that  $f_L \notin (f_m, \hat{\theta}_L]$ . To do this, we prove that  $w_L(f_L)$  is increasing in  $f_L$  over  $(f_m, \hat{\theta}_L]$ . Consider two reports by the last analyst:  $f_L = \hat{\theta}_L - \xi + \epsilon$  and  $f_L = \hat{\theta}_L - \xi$ , where  $0 \leq \xi < \hat{\theta}_L - f_m$ . The payoffs from these two reports can be used to compute the derivative of  $w_L(f_L)$  at  $f_L = \hat{\theta}_L - \xi$ . This is given by

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L - \xi + \epsilon) - w_L(\hat{\theta}_L - \xi)}{\epsilon} \middle| \hat{\theta}_L \right] &= - \int_{\theta \leq f_m} (\lambda + R'(f_m - \hat{\theta}_L + \xi)) dG(\theta | \hat{\theta}_L) \\
& - \int_{f_m}^{\hat{\theta}_L - \xi} (\lambda + R'(2\theta - \hat{\theta}_L + \xi - f_m)) dG(\theta | \hat{\theta}_L) \\
& + \int_{\theta > \hat{\theta}_L - \xi} (\lambda + R'(\hat{\theta}_L - \xi - f_m)) dG(\theta | \hat{\theta}_L). \tag{12}
\end{aligned}$$

Consider the second term on the RHS of (12). Since  $\theta < (\hat{\theta}_L - \xi)$ , we have  $R'(2\theta - \hat{\theta}_L + \xi - f_m) \leq R'(\hat{\theta}_L - \xi - f_m)$  by the convexity of  $R(\cdot)$ . Also  $R'(f_m - \hat{\theta}_L + \xi) \leq R'(\hat{\theta}_L - \xi - f_m)$ . Hence,

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L - \xi + \epsilon) - w_L(\hat{\theta}_L - \xi)}{\epsilon} \middle| \hat{\theta}_L \right] &> - \int_{\theta \leq \hat{\theta}_L - \xi} (\lambda + R'(\hat{\theta}_L - \xi - f_m)) dG(\theta | \hat{\theta}_L) \\
& + \int_{\theta > \hat{\theta}_L - \xi} (\lambda + R'(\hat{\theta}_L - \xi - f_m)) dG(\theta | \hat{\theta}_L).
\end{aligned}$$

Again, by the symmetry of  $g(\theta|\hat{\theta}_L)$  about  $\hat{\theta}_L$ , we get

$$\lim_{\epsilon \rightarrow 0} E \left[ \frac{w_L(\hat{\theta}_L - \xi + \epsilon) - w_L(\hat{\theta}_L - \xi)}{\epsilon} \middle| \hat{\theta}_L \right] > \int_{\hat{\theta}_L - \xi}^{\hat{\theta}_L + \xi} (\lambda + R'(\hat{\theta}_L - \xi - f_m)) dG(\theta|\hat{\theta}_L) \geq 0,$$

$\forall 0 \leq \xi < \hat{\theta}_L - f_m$ . The last analyst's payoff is therefore increasing in  $f_L$  over  $(f_m, \hat{\theta}_L]$ . Hence, the optimal  $f_L$  will be such that  $f_m < \hat{\theta}_L < f_L$ .

(ii) The proof that  $f_L < \hat{\theta}_L < f_m$  if  $\hat{\theta}_L < f_m$  is analogous.  $\square$

## References

- [1] Abarbanell, J. and V. Bernard (1992), “Tests of Analysts’ Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior,” *Journal of Finance* 47, 1181–1207.
- [2] Alexander, John C. and James S. Ang (1985), “The Market Response to Low-Balling Earnings Expectations,” *Working paper*, Florida State University.
- [3] Chevalier, Judith and Glenn Ellison (1997), “Risk Taking by Mutual Fund as a Response to Incentives,” *Journal of Political Economy*, 106(6), 1167–1200
- [4] Christie, Andrew A. (1987), “On Cross-Sectional Analysis in Accounting Research,” *Journal of Accounting and Economics*, 9, 231–58.
- [5] Davidson, Russell and James MacKinnon (1993), *Estimation and Inference in Econometrics*, New York, Oxford University Press.
- [6] De Bondt, Werner F. M. and Richard H. Thaler (1990), “Do Security Analysts Overreact?” *American Economic Review* 80, 52–57.
- [7] Dugar, Amitabh and Siva Nathan (1995), “The Effects of Investment Banking Relationships on Financial Earnings Forecasts and Investment Recommendations,” *Contemporary Accounting Research*, 12, 131–60.
- [8] Ehrbeck, Tilman and Robert Waldmann (1996), “Why Are Professional Forecasters Biased? Agency versus Behavioral Explanations,” *Quarterly Journal of Economics*, Feb 1996, 41–64.
- [9] Foster, George, Chris Olsen and Terry Shevlin (1984), “Earnings Releases, Anomalies, and the Behaviour of Security Returns,” *Accounting Review*, 56, 574–603.
- [10] Henry, George (1989), “Wall Street Economists: Are they Worth Their Salt?” *Business Economics*, Oct., 44–48.
- [11] Holmstrom, Bengt (1982), “Managerial Incentives Problems — A Dynamic Perspective,” in *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki: Swedish School of Economics.
- [12] Hong, Harrison, Jeffrey Kubik and Amit Solomon (1998), “Security Analysts’ Career Concerns and Herding of Earnings Forecasts,” *Discussion paper*, Stanford University, July.
- [13] Hsiou-wei, Lin and Maureen F. McNichols (1998), “Underwriting Relationships, Analysts’ Earnings Forecasts and Investment Recommendations,” *Journal Of Accounting And Economics*, 25 (1), 101-127
- [14] I/B/E/S Research Bibliography (1996), “Annotated Bibliography of Earnings Expectation Research,” Institutional Brokers estimate System, Inc., 5th Edition.

- [15] Imhoff, Eugene and Gerald Lobo (1992), “The Effect of *Ex Ante* Earnings Uncertainty on Earnings Response Coefficients,” *Accounting Review*, 67 (2), 427–439.
- [16] Kandel, Eugene and Neil D. Pearson (1995), “Differential Interpretation of Public Signals and Trade in Speculative Markets,” *Journal of Political Economy*, 103 (4), 831–872.
- [17] Keane, Michael P. and David E. Runkle (1998), “Are Financial Analysts’ Forecasts of Corporate Profits Rational?” *Journal of Political Economy*, 106 (4), 768–805.
- [18] Lamont, Owen (1995), “Macroeconomic Forecasts and Microeconomic Forecasters,” NBER Working paper # 5284
- [19] Laster, David, Paul Bennett and In Sun Geoum (1999), “Rational Bias in Macroeconomic Forecasts,” *Quarterly Journal of Economics*, 114 (1), 293–318.
- [20] Lys, Thomas and Songkyu Sohn (1990), “The Association between Revisions of Financial Analysts’ Earnings Forecasts and Security–Price Changes,” *Journal of Accounting and Economics*, 13, 341–363.
- [21] Mikhail, Michael, Beverly Walther and Richard Willis (1999), “Does Forecast Accuracy Matter to Security Analysts?” *The Accounting Review*, forthcoming
- [22] O’Brien, Patricia (1988), “Analysts’ Forecasts as Earnings Expectations,” *Journal of Accounting and Economics*, 10, 53–83.
- [23] Scharfstein, David and Jeremy C. Stein (1990), “Herd Behavior and Investment”, *American Economic Review*, June 1990, 465–79.
- [24] Stickel, Scott (1990), “Predicting Individual Analyst Earnings Forecasts,” *Journal of Accounting Research*, 28(2), 409–417.
- [25] — (1992) “Reputation and Performance among Security Analysts,” *Journal of Finance*, 47, 1811–1836.
- [26] Trueman, Brett (1990), “On the Incentives for Security Analysts to Revise Their earnings Forecasts,” *Contemporary Accounting Research*, 7(1), 203–222.
- [27] — (1994), “Analyst Forecasts and Herding Behavior,” *Review of Financial Studies*, 7(1), 97–124.
- [28] Zweibel, Jeffrey (1995), “Corporate Conservatism and Relative Compensation”, *Journal of Political Economy*, 103(1), 1–25.

**TUFTS UNIVERSITY  
ECONOMICS DEPARTMENT**

**DISCUSSION PAPERS SERIES 1999**

- 99-09 Edward KUTSOATI and Dan BERNHARDT; Can Relative Performance Compensation Explain Analysts' Forecasts of Earnings?
- 99-08 Don FULLERTON, Inkee HONG, and Gilbert E. METCALF; A Tax on Output of the Polluting Industry is Not a Tax on Pollution: The Importance of Hitting the Target.
- 99-07 Marcelo BIANCONI; A Dynamic Monetary Model with Costly Foreign Currency.
- 99-06 Marcelo BIANCONI; The Effects of Alternative Fiscal Policies on the Intertemporal Government Budget Constraint.
- 99-05 Lynne PEPALL and Dan RICHARDS; The Simple Economics of "Brand-Stretching."
- 99-04 Drusilla K. BROWN, Alan V. DEARDORFF, and Robert M. STERN; U.S. Trade and Other Policy Options and Programs to Deter Foreign Exploitation of Child Labor.
- 99-03 Margaret S. MCMILLAN and William A. MASTERS; A Political Economy Model of Agricultural Taxation, R&D, and Growth in Africa.
- 99-02 MASTERS, William A. and Margaret S. MCMILLAN; Ethnolinguistic Diversity, Government Expenditure and Economic Growth Across Countries.
- 99-01 MCMILLAN, Margaret S.; Foreign Direct Investment: Leader or Follower?

Discussion Papers are available on-line at

<http://www.tufts.edu/as/econ/papers/papers.html>