Does Past Success Lead Analysts to become Overconfident?

Gilles Hilary

Lior Menzly

Graduate School of Business

University of Chicago

The authors would like to thank workshop participants at the University of Chicago, the 7th Annual Tel Aviv University Accounting and Finance Conference and IAAER 2000 as well as Raymond Ball, Daniel Bens, Robert Bloomfield, Brian Bushee, Steven Levitt, Katherine Schipper, Richard Thaler, Richard Willis, George Wu and numerous colleagues (especially Christopher Malloy and Cade Massey) in the Ph.D. program for their insightful comments. We are also grateful to the IAEER for awarding the "Aoki Outstanding Papers" Prize to this paper.

Does Past Success Lead Analysts to become Overconfident?

Abstract

This paper provides evidence that analysts who made a sequence of accurate predictions of earnings relative to the median forecast tend to be relatively more inaccurate and more out of consensus in their subsequent earnings prediction. This phenomenon is economically and statistically meaningful. The results are robust to different estimation techniques and different control variables. Our findings are consistent with an attribution bias that leads accurate analysts to become too confident in the quality of their information. However, we also offer alternative explanations that would rationally fit the data.

JEL: M40, G20

I. Introduction

This paper studies whether analysts become overconfident in their ability to predict future earnings after achieving a series of good predictions. We find that analysts who predict earnings more accurately than the median analyst in the previous four quarters tend to be relatively more inaccurate and more out of consensus in their subsequent earnings prediction. This result is consistent with our hypothesis that analysts exhibit overconfidence after achieving a series of successful predictions. However, we also offer alternative explanations that would rationally fit our empirical findings.

Analysts constitute a key institution in financial markets. Their earnings forecasts are often used as a proxy for market expectations and differences in opinion. In addition, this is one of the rare settings where researchers have a large data set on actual decisions made by individuals. Not surprisingly, it has been a fertile ground for behavioral research. In particular, prior literature suggests that analysts (1) make upwardly biased forecasts,[1] (2) overreact to positive information, and (3) under-react to negative information.[2] Yet, these findings were recently challenged on several grounds.[3] For example, Abarbanell and Lehavy [2000] posit that firms manage earnings by taking "big baths" on an irregular basis, and that analysts are unable or unwilling to accurately forecast these post-management earnings. This creates skewed data and an appearance of optimism or overreaction. Whether analysts forecasts are significantly affected by cognitive biases remains therefore an important but unresolved question.

---

[1] DeBondt and Thaler [1990], for example, report that the projected earnings exceed actual earnings and that actual earnings change less than analysts' forecasts do. Similarly, Abarbanell and Bernard [1992] find that forecasts tend to exceed the subsequently observed earning figures, and that actual earnings changes tend to move less than forecasted earnings.

[2] Easterwood and Nutt [1998] show that analysts' reactions are contingent on the nature of the earnings information provided. They divide firms into three groups based on prior performance (earnings change), and then, for each group, they regress the forecast error against prior performance. They conclude that analysts underreact to negative information and overreact to positive information.

[3] Using a GMM estimation technique, Keane and Runkle [1998] show that controlling for discretionary asset write-downs undermines the upward bias results of Ababarnell and Bernard [1992] and DeBondt and Thaler [1990]. Gu and Wu [2000] as well as Kothari, Sabino and Zach [1999] claim that some apparent irrationality in analysts' behavior may be due to skewed data. In addition, O'Brien and McNichols [1997] posit that the observed forecast distribution is truncated. They argue that analysts have an incentive not to release unfavorable forecasts in order to nurture their relations with company management. Consequently, the observed forecasts would seem to be overly optimistic even though the true distribution is unbiased.

In this paper, we offer a different approach to study this question. We consider the interdependence between accuracy and deviation from the consensus. Extant empirical literature has studied these two characteristics, but not in a unified framework. In addition, we focus on short-term cyclical variations whereas past research has considered either fixed analysts' characteristics or long-term trends. For example, Shinha, Brown and Das [1997] find systematic differences in forecast accuracy across analysts. Yet, they consider analysts' skill as a fixed parameter that varies cross-sectionally but not over time. Other empirical studies have considered the evolution of analysts' characteristics over time. One stream of literature (e.g. Mikhail et al [1997], Clement [1999], or Jacob et al [1999]) looks at the effect of experience on forecast accuracy and finds mixed results. Another stream considers the effect of experience on "herding" (i.e. the willingness to deviate from consensus). Kubrick et al [2000] find that analysts' deviate more from consensus as they gain experience. Zuckerman and Philips [2001] find that analysts are more likely to conform to the norm in the middle of their career. We depart from these papers by focusing on short-term dynamics rather than a long-term trend.

Our results indicate that after a short series of accurate predictions, analysts are more likely to be inaccurate. They also take additional risk by deviating from the consensus forecast on their subsequent prediction. This phenomenon is both statistically and economically meaningful. The results are robust to the use of different econometric techniques and to numerous control variables. They suggest a counter-intuitive implication. If two analysts possess identical skill and experience but only one of them had a recent series of superior predictions, investors may want to rely more on the subsequent forecast of the less accurate analyst.

These joint results are consistent with our research hypothesis that analysts are subject to over-confidence. [4] An individual is said to exhibit overconfidence if he or she overestimates the precision of his or her own information relative to public signals. In our setting, after analysts have had a series of good predictions, they become overconfident in their ability to predict future earnings. This leads them to put more weight on their own predictions, and less on public

[4] Note that there could be other behavioral reasons that would lead to a reduction in risk. For example, Kahneman and Tversky's ([1979]) framework (loss aversion and shift in reference) implies that after changing their reference point, analysts perceive losing new gains as having a greater negative effect on utility than obtaining these gains in the first place. In addition, the fact that individuals would make ex ante irrational decisions in order to avoid ex post

signals, such as market reactions and other analysts' forecasts. As a consequence, their next prediction is likely to be more out of consensus, and more inaccurate on average. Thus, this phenomenon is a short-term one that recurrently appears and disappears. It follows a cyclical pattern and its intensity varies with the series of performance. Random past success creates overconfidence, which, in turn, increases the probability of poor subsequent forecasts, and thus reduces overconfidence. We motivate this analysis by relying on two well-established behavioral principles: self-attribution and "static" overconfidence. According to self-attribution theory, individuals attribute positive events to their own abilities and negative events to external forces. On the other hand, overconfident subjects put too much emphasis on their private information (see Kramer et al [2001] for an experimental research, Barber and Odean [2001] for an example of large sample test). The combination of the two cognitive principles yields the dynamic notion of overconfidence. A similar analytical framework has been suggested by Gervais and Odean [1998], but, to our knowledge, no large sample empirical research has been conducted to investigate this question.

However, we cannot totally reject rationality because the existence of irrationality is always conditional on the objective function of the analysts or of the brokerage houses.[5] In fact, analysts could rationally maximize their utility given a reward function, in a way that is consistent with any given forecast behavior. Our results can also be explained by making ad hoc assumptions about the reward function of the brokerage house or of the analysts. For example, consider a reward function that is path dependent with respect to the number of good predictions. If the payoff on the last prediction is asymmetric, an analyst who has already achieved, say, four good predictions in a row has an incentive to take a higher risk on the next one. Even though this strategy induces a higher expected level of error, it could also entail a higher expected payoff for the analyst (for instance, becoming a "media star"). Nevertheless, the combination of our two tests places limits on compensation structures that are consistent with rationality: if we assume that the analysts' reward function is increasing in accuracy and that deviation from consensus is rewarded if it is associated with lower error, this function would have to be both

---

remorse about a decision that led to a bad outcome (the no regret principle of DeBondt and Thaler [1995]) may also lead analysts to reduce their risk.

[5] Note also that forecasting earnings is only one of the tasks performed by analysts, writing reports or making buy/sell recommendations are also important. See Francis and Philbrick [1993] as an example of analysis of the impact of multitasking on analysts' behavior.

path dependent and asymmetric (i.e., the payoff is increasing in number of good predictions faster than the penalty).

This paper contributes to the literature in three ways. First, it bears on the behavioral literature. It examines whether a combination of two major behavioral principles (overconfidence and self-attribution) have explanatory power in the specific setting of analyst forecasts. It offers a large sample test that complements previous research that was largely based on small sample experimental work. Second, it develops a research design that integrates previous criticisms. We focus on the change of an analyst's current performance relative to other analysts' performance. Therefore, earning surprises and unexpected write-offs equally affect all analysts and are less likely to bias our estimates. By using medians instead of means, we control for skewness in the data. Third, departing from prior analyst literature, we examine the short-term dynamics behind analyst forecasts (i.e., the influence of past success on current forecasts). Although theoretical papers have considered this issue, to the best of our knowledge, it has not been tested empirically.

The rest of the paper is organized as follows. In the next section, we present the theoretical foundations of our analysis and competing hypotheses. In section III, we provide the empirical design followed, in section IV, by the main results. We explore related issues in section V and conclude in section VI.

II. Description of the framework

In this section, we first present the underlying theoretical foundations for our framework, before describing our analysis and presenting competing implications.

II.1 Theoretical Foundations

The behavioral research has documented, among others, the following two major cognitive biases.

Self-attribution

The behavioral literature has evidenced numerous egocentric biases in cognitive processes.[6] A large body of this research indicates that individuals do not employ the same causal explanations to account for their successes and failure (e.g. Fitch [1970], Weiner and Kukla [1970], Kukla [1972]). Specifically, according to the self-attribution theory, individuals attribute too strongly events that confirm the validity of their own actions to high ability, and events that disconfirm their actions to external noise (Hastorf, Schneider and Polekfa [1970]). For example, studies by Johnson, Feigenhaum and Weiby [1964] and Beckman [1970] show that teachers tend to claim responsibility for a student's performance when there is improvement. When students perform poorly, the blame is attributed to various external causes such as the child's motivation or situational factors. Miller [1976] finds that tendencies toward self-attribution are stronger when the task is important ("ego-involving") for the subject.

Overconfidence

DeBondt and Thaler [1995] argue that "perhaps the most robust finding in the psychology of judgment is that people are overconfident." Overconfident people tend to overestimate their own ability (Camerer [1995]).[7] For example, in a classical study, Svenson [1981] finds that 90% of the automobile drivers in Sweden consider themselves as "above average". Moreover, individuals exhibit more overconfidence in their own ability in an ambiguous situation (Heath and Tversky [1990]).

However, in this paper, we depart from this general form and study overconfidence in the context of private and public information. In this setting, subjects are said to be overconfident if they "put too much weight on their private information" (Kraemer et al [2000]). The existence of this form of overconfidence is consistent with extant experimental literature. Hung and Plott [1999] estimate the weights of public and (free) private information in an experimental setting. They report that participants put too much weight on their private information. Huck and Oechssler [1999] find an extreme form of overconfidence. In a similar setting, they report that

[6] Ross and Sicoly [1979] have documented several examples of egocentric biases. For instance, they show that individuals tend to accept more responsibility for joint product than other contributors attribute to them.

[7] Camerer [1995] presents empirical results from "calibration curves" (curves that matches probability judgments with actual relative frequencies). He reports that "in general subjects are overconfident. They are insufficiently regressive in judging the likelihood of events. Events they say are certain happen only 80 percent of the time. Events judged to be impossible happen 20 of the time ".

the heuristic "follow your (private) signal" explains the observed behavior better than Bayes' law. Kraemer et al [2000] introduce a cost of acquisition for private information. They conclude that "about one half of the individuals act rationally, whereas the other participants overestimate the private signal value". Bloomfield, Krische and Nelson [2000] report experimental evidence that people are overconfident in their ability to interpret data (relative to the ability of a disciplined trading strategy) and under-perform as a result.

Our research question is motivated by overconfidence rather than optimism. Optimism is defined as overestimating the probability of a favored outcome, whereas overconfidence is the fact that an individual is overestimating the precision of his information relative to others' information. This paper is not concerned with whether analysts exhibit optimism or pessimism in their forecasts (i.e., the forecasts exhibit upward or downward bias), but rather with the dynamic aspects of overconfidence.

A Bayesian analogy alludes to the following two types of overconfidence. One possibility is that the agent starts with an abnormally strong prior about her skills, implying a longer update. In the limiting case, where the agent is absolutely sure of her skills (i.e. the variance of the prior is zero), there is no updating at all. A different form of overconfidence can be derived from the self-attribution bias. In this case, the overconfidence comes from an incorrect updating procedure. The temporal patterns of the two forms of overconfidence are different. In the first case, overconfidence decreases over time (or stay the same in the limiting case). In the second, overconfidence follows a cyclical pattern where its intensity varies with the series of performance.

II.2 Framework

Our framework suggests that when an analyst has had a series of good predictions, she becomes overconfident in her ability to predict future earnings. The self-attribution principle predicts that analysts who have successfully forecasted earnings in previous periods, improperly (i.e. in a non Bayesian manner) update their beliefs about their own skill. They attribute too much of their success to superior ability and too little to chance. Thus, they become overconfident. Overconfidence, in turn, induces analysts to overweight their private information and to rely less on public signals, such as other analysts' forecasts. As a consequence, their next forecast is likely to be more out of consensus, and on average more inaccurate. This framework is germane to the model developed by Gervais and Odean [1998] as well as Daniel et al [1998].[8]

The combination of the two cognitive biases yields a dynamic notion of overconfidence. Attributing past success to one's own ability leads to overconfidence. This yields a non-optimal behavior, as one is less likely to rely on public information. As a result, the likelihood of an inferior forecast on the subsequent prediction increases, which, in turn, leads to a reduction in self-confidence. In other words, overconfidence, in this case, is not a fixed characteristic (i.e. the analysts systematically over estimates her skills) but rather a recurring phenomenon whose intensity is dynamic in nature.

A major criticism of the behavioral literature is that irrational participants will either learn (e.g. Friedman [1998]) or be driven out of the markets over time. Our framework does not preclude some form of learning. It is possible that an analyst becomes better at predicting earnings, while remaining overconfident (i.e. not getting better at estimating her own skills). Past literature, however, presents conflicting results on whether analysts are able to learn at all (Mikhail et al [1997], Jacob et al [1999]). The second criticism does not apply to this framework because, in our case, overconfidence is not a permanent characteristic of an agent. Therefore, it cannot be used as a basis for driving out people. In fact, to the extent that overconfidence is correlated with past success, overconfident analysts may be the most influential ones. Thus, overconfident analysts may play an important long-term role in financial markets.

II.3 Competing hypotheses

Our behavioral framework leads to specific predictions: both deviations from consensus and degree of error are positively correlated with past success. However, alternative theories ("cash-in", "hot hand") propose opposite predictions. The following discussion describes the implications of each theory that enable us to discriminate between them.

---

[8] In their model, Daniel et al propose a theory of securities market under- and over reactions based on investors' confidence about the precision of their private information and biased self-attribution causing symmetric shifts in investors' confidence as a function of their investments outcomes.

II.3.1 Risk

** No effect (null hypothesis)*

Our null hypothesis is that past success has no impact on current forecasts. This would be consistent with a world where performance is not auto-correlated[9] and analysts' reward functions are not path dependent. Path dependency may generate a relation between past performance and current risk. For example, consider a reward function where the payoff on the last prediction is asymmetric (i.e. the payoff is increasing in the number of good predictions faster than the penalty):[10] an analyst who has already achieved several good predictions in a row has an incentive to take a higher risk on the next one. Even though this strategy induces higher expected error, it could also give a higher expected payoff to the analyst. This situation would not be consistent with the null hypothesis.

** Less risk (cash-in)*

Anderson and Holt [1997] note that "a forecaster may prefer the chance of being wrong with everybody else to the risk of providing a deviant forecast that turns out to be the only incorrect guess". Keynes [1936] remarks that "worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally." Graham [1999] presents a model where analysts' compensation structure lead them to rationally "cash in" past success.[11] Analysts with high reputation (and salary) would herd to protect their current reputation and level of pay if their past performance was abnormally good given their ability. By taking less risk, the market would not be able to update its assessment of the analyst's skill as much as with an out of consensus forecast.

** More risk ("hot hand" or " overconfidence")*

The overconfidence theory predicts that analysts would take more risk after a string of good predictions. This is analogous, in the Gervais-Odean model, to traders investing more.

---

[9] If performance is auto-correlated, analysts may want to mimic the predictions of analysts who had a good performance in the previous quarter. Her forecast may then become the reference point for other forecasts. Therefore, her risk would be lower.

[10] For example, if the analyst had a string of good predictions, then she might be more likely to make a bold prediction to become a 'media star".

[11] Graham's measure is different from ours. He measures the difference between the current forecast and previous forecasts. We measure the difference between a forecast and all first forecasts in a quarter. We do not test a direct implication of his model, but rather a hypothesis inspired by his work.

However, this would also be consistent with the analyst making a bolder (and more accurate) prediction because of persistent superior private information ("hot hand").

II.3.2 Error

** No effect (null hypothesis)*

Our null hypothesis is that, controlling for skill, past success has no effect on current forecast error. This implies no auto-correlation in analysts' performance.

** Less error ("hot hand")*

Alternatively, the "hot hand" (i.e., temporary superior information) hypothesis implies that the forecast accuracy should be positively correlated over time. This could be explained, for example, by superior information (for example, due to better access to management or to greater knowledge about a new product) for a short time period.

** More error (overconfidence)*

Our model predicts that after a string of successes, an analyst will underweight public signals. This will increase the likelihood of a poor subsequent forecast.

II.3.3 Summary

Note that neither seeing an increase in risk nor an increase in error after a previous good prediction is sufficient to support the overconfidence hypothesis. Only the combination of the two can do so. The following table summarizes the various predictions concerning error and risk (the cells show the predicted sign of the correlation between past and current performance under the different hypotheses).

Effect of stream of good predictions

	Over Confidence	*Null Hypothesis*	*"Hot hand"*	*"Cash in"*
Deviation from consensus (DEV)	+	*0*	*?*	-
Imprecision of the forecast (ERROR)	+	*0*	-	*?* *(Close to Mean Error)*[12]

[12] By reducing the risk, the prediction gets closer to the average forecast, and thus, closer to the average error.

III. Empirical design

III.1 Data

The analyst forecast data is retrieved from the Zacks database and cover the period from the last quarter of 1980 to the last quarter of 1997. We match the forecast data to the corresponding records of COMPUSTAT reported earnings. To increase the consistency of the data, we delete firms that do not have a December year-end report and we focus on quarterly predictions made one quarter ahead. If an analyst makes more than one forecast in a given quarter, only the first one is considered because the later predictions may be drawn from a different distribution. To obtain a meaningful measure of the consensus, we also concentrate on observations that have at least four analysts covering the firm in a particular quarter (Elliot, Philbrick and Wiedman [1995], Easterwood and Nutt [1998]). To be included in our sample, an analyst needs to have made at least four predictions for a given firm in the preceding quarters. Abarbanell and Lehavy [2000] stress that "a relatively small number of observations can have a disproportionate effect" with non-central distributions, and thus create the appearance of a bias. To abstract from this issue, we remove outliers by deleting observations where the difference between the prediction and the realization (scaled by price) is in the top or bottom one percent of the sample.[13] This procedure also mitigates any error in the data. These data requirements yield a sample of 46,905 quarterly observations.

III.2 Definition of dependent variables

The two dependent variables are the risk taken by the analyst (i.e. the deviation from consensus), and the error of the forecast. We calculate *DEV* as the absolute value of the difference between the analysts' forecasts and the consensus, where consensus is defined as the median forecast for a given firm and for the quarter.[14] We compute *ERROR* as the absolute value of the difference between actual and forecasted earnings. Keane and Runkle [1998] and Ababarnell and Lehavy [2000] argue that part of the bias in analyst forecasts is, at least partially, due to the existence of

[13] As a robustness test, we rerun the regressions without the filters. In the fixed effect and Fama-MacBeth specifications, the coefficients for FREQ and STREAK are always significant. The significance increases for the DEV regressions and decreases slightly for the ERROR regressions. In the pooled regressions, FREQ and STREAK remain significant at the 10% but not in the ERROR regressions.

[14] In order to avoid biases from extreme observations, we use the median instead of the mean. As a further robustness check for skewness, we rerun the regressions with logged control variables. Our results are unaffected.

assets write-offs. To mitigate this concern, we use EPS before extraordinary items. We use the diluted or non-diluted EPS (Compustat items 9 or 19) depending on what the analyst forecasts (as reported in the Zacks database).

In this paper, we define *DEV* and *ERROR* variables in terms of absolute difference. However, one should note that these particular measures of risk and error might be problematic because analysts' published predictions tend to be higher on average than the earnings. There could be two sources of inaccuracy in our measure for error.

First, the analyst's actual forecast may not be the one published. O'Brien and McNichols [1997] posit that the distribution of observed forecasts is not the distribution of forecasts made internally by analysts. They argue that analysts' firms have an incentive not to release unfavorable forecasts in order to nurture their relations with company management. Ideally, we would like to take the true (but unobservable) analysts' forecasts rather than the released number.

Second, the target that analysts aim for may not be the earnings per se, but rather a biased measure of it. The analysts may induce the bias intentionally because of a particular reward function. For example, the true predicted variable could be realized earnings plus a positive constant in order to ingratiate themselves with company management. In this case, the error in forecast should not be computed as a comparison of forecast and realized earnings, but to the true object of forecast.

III.3 Description of explanatory variables

We distinguish among the explanatory variables between the two "decision" variables and the control ones. To capture the notion of a sequence of good predictions, we define *FREQ* as the number of superior predictions in the preceding four quarters.[15] A superior prediction is a forecast whose error is below the median error in absolute value of all analysts' errors who cover the same firm in that particular quarter. Alternatively, we define *STREAK* as the number of superior predictions in a row before the current prediction is made.[16] *FREQ* proxies for the

---

[15] We choose four quarters (i.e. a year) as a compromise between reducing the period too much (which would not let us capture the sequence of the forecasts) and extending the period too much (which would let control variables, such as ACCURACY, pick up most of the effect). We perform a robustness check on this assumption in 4.2.

[16] Psychologists find that people not only overweight their successes and underweight their failures but also they overweight successes more than they underweight failures. For example, Fiske and Taylor [1991] state that "self-enhancing attribution for success are more common than self-protective attributions for failures". Therefore, we consider the number of above median performance and not the below median ones.

frequency of good predictions, whereas *STREAK* proxies for the length of the stream of superior predictions.

We provide controls for earning surprises and the level of uncertainty, which affect all analysts covering a given company. *Current Median Error* (*CME*) is the median error of all analysts who follow a given company for the current quarter. Similarly, *Current Median Deviation* (*CMD*) is the median risk of the analysts (excluding the analyst making the forecast) who follow a given company for the current quarter. Previous literature (e.g. Hwang et al [1996], Keane and Runkle [1998], Abarbanell and Lahavy [2000]) find that the consensus forecast is more upwardly biased and more inaccurate for firms with negative earnings and that losses may explain considerable portion of previously documented anomalies. To control for this potential effect, we introduce a dummy variable for losses (*LOSS*) in the regression.

We also provide controls for the skill of the analyst. *ACCURACY* (*A*) is negative one times the median relative error (over the entire period of coverage) for a given analyst. We calculate the relative error as the difference between the analyst error and the median error of all analysts for a given quarter for a given company. The existence of an intrinsic skill component is consistent with the results of Jacob et al [1999].

In addition, we employ two variables to proxy for the size of the information set available at the time the prediction is made. *TIME (T)* is the number of days between the date of the forecast and the end of the quarter. *RANK ORDER* (*O*) is the number of forecasts published before the analyst issues the prediction divided by the total number of forecasts published over the quarter.

Mikhail, Walther and Willis [1997] find that experience may lead to improvement in analysts' forecasts. To control for this possible effect, we calculate *EXPERIENCE*, the number of quarters when analysts have made predictions before they make the current forecast.

Finally, past literature has shown that firm size is correlated with numerous factors. To proxy for possible correlated omitted variables, we introduce MCAP, *Log of Market Equity (MCAP),* as the natural logarithm of the market value of firm equity, and proxies for the size of the firm. For example, it could proxy for the amount of information available to the market for a particular firm. In section VII, we explore the role of additional control variables.

III.4 Descriptive statistics

We present general descriptive statistics in Table 1 as well as a correlation table in Table 2. The results indicate several points. First, consistent with previous studies, average realized earnings are lower than forecasts. Second, when the observations are segregated based on analyst skills, both *DEV* and *ERROR* follow a U shape (even after being deflated by price at the beginning of the quarter). The group with the highest skill level also has a high average *ERROR*. This can be explained by the fact that other analysts are making even larger errors. In other words, "skilled" analysts follow firms whose earnings are more difficult to forecast (higher standard deviation in earnings). Third, if all analysts had 0.5 probability of making a superior prediction independent of prior performance, the expected value would be 2 for *FREQ* and 1.625 for *STREAK*. The average values, reported in Table 1 (1.24 and .48 respectively), are consistent with the idea that analysts' performance is decreasing after a success.

Table 3 provides a preliminary analysis of the dynamics of superior performance. Panel A reports the estimated transition matrix for *FREQ*, the unconditional (steady state) probability and the implied mean. The transition matrix reports the estimated probability for analysts to obtain a certain level of *FREQ* given its value in the previous quarter. This matrix contains an implicit assumption, that the dynamic of the system is governed by first order Markov process (i.e. lags higher than the first one are irrelevant for characterizing the dynamics). The unconditional probabilities represent the probability that an analyst obtains a certain a certain value of *FREQ* without conditioning on its previous value.[17] The implied mean of *FREQ*, calculated from the unconditional probability, represents the steady state mean implied by the transition probability matrix. It is fairly close to the estimated mean (1.3 vs. 1.24 respectively) which seems to suggest that first order Markov chain is a fairly reasonable approximation for dynamics of *FREQ*. As a benchmark, Panel B reports the transition matrix if the conditional probability of superior forecast is equal to 0.5 for every analyst. Consistent with the overconfidence hypothesis, the estimated probability of obtaining additional success is lower than under the i.i.d assumption.

Table 4 provides a descriptive statistics on the various variables conditional on the value of *FREQ* and *STREAK*. The results support the overconfidence hypothesis. As the value of

---

[17] The unconditional probabilities are computed by iterating the transition matrix until the product converges to a steady state.

FREQ or *STREAK* increases, the mean of *DEV* increases monotonically. This phenomenon persists after we deflated the value by the price at the beginning of the quarter. However, the standard deviation is high and increasing. For example, the mean of risk conditional on the value of *FREQ* increases from 0.032 to 0.087 but the standard deviation ranges from 0.10 to 0.44. Consistent with overconfidence, the value of *ERROR* also increases with *FREQ* (especially after deflating *ERROR* by price). Standard errors, however, are large. Results for *STREAK* are similar. However, considering the average without controls may be misleading. For instance, higher *FREQ* may pick up better analysts. Therefore, in the next section, we present a regression approach that provides better controls.

III.5 Methodology

To test the two groups of separate hypotheses, we first estimate two sets of pooled regressions: one for *DEV* on *FREQ* (and alternatively *STREAK*) and the other for *ERROR* on *FREQ* (and *STREAK*).

$$DEV_t^i = \alpha + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$
$$ERROR_t^i = \alpha + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

where i is an analyst indicator

To mitigate a potential heteroskedasticity problem, we deflate all variables involving *ERROR* and *DEV* (because they contain the number of shares in the denominator). Hence, all variables except *FREQ, STREAK, TIME* and *RANK ORDER* are scaled by the price at the beginning of the quarter.[18]

We employ three different estimation techniques: pooled regression, Fama-MacBeth [1973] (FM hereafter) regressions and a panel (fixed effect) regression. [19] In the pooled analysis, the regressions are estimated over the entire sample and the standard errors are calculated using the White correction. The FM procedure is implemented to control for cross-correlation in the

[18] The analysis is focused on earnings to price rather than earnings per share. Similar results are obtained without deflating the variables.

[19] As a robustness check, we run Seemingly Unrelated Regressions (SUR). This procedure is partially motivated by the relation between the error in forecast and the deviation from consensus, implied by the following expression:

$$\begin{aligned} Error^i &\equiv |\,forecast^i - earnings\,| \equiv |\,(forecast^i - consensus) + (consensus - earnings)\,| \\ &\leq |\,forecast^i - consensus\,| + |\,consensus - earnings\,| \equiv Risk^i + Mean\ Error \end{aligned}$$

Hence, an increase in the deviation from consensus leads to an increase in the maximum error. This also implies that there might be a simultaneous relation between the regressions of *ERROR* and *DEV*. Most of our results are error term. For example, macroeconomic shocks could drive cross-correlation across analysts' forecasts.[20] We also control for analyst/firm fixed effects to provide a better control for omitted variables (i.e. variables that proxy for analyst characteristics such as skill). For example, differences in analysts' skill level, prediction bias or willingness to take risk could all potentially affect our results. For each series of analyst forecasts of a given firm, we subtract its mean from each variable and rerun the regressions. This is equivalent to including dummies for each analyst and forecasted firm combination in the regressions. The standard errors are again adjusted using the White correction. This method is particularly suitable to test our hypotheses, which emphasize the short-term dynamics in analyst's forecasts. It eliminates the cross-sectional variations in the means yet it leaves the time-series dynamics of the analyst forecasts intact. However, it reduces the power of the test due to the high number of variables that needs to be estimated.

IV. Regression Results.

This section describes the regression results. We report the results from the pooled estimation in Table 5, from the Fama-MacBeth procedure in Table 6 and from the fixed effects regressions in Table 7.

In each table, column I and II respectively report the regressions, where *Dev* is used as a dependent variable, and *STREAK* and *FREQ* are the explanatory variables. Similarly column III and IV report the regressions where *ERROR* is the dependent variable.

IV.1. Overconfidence and Deviation from Consensus.

The results of the *DEV* regressions reported in columns I and II support the overconfidence hypothesis. The coefficients of *FREQ* and *STREAK* are positive and significant, both economically and statistically, which suggests that analysts increase their risk after a stream of success. The White corrected t-statistics for the pooled regression, are 3.13 and 2.22 for *FREQ*

---

essentially unaffected, but the t-statistics for *FREQ* or *STREAK* in the DEV regressions are materially improved. Wilks' tests indicate joint significance of *FREQ* or *STREAK* at the 0.01% level.

[20] In each quarter, we ran a cross sectional regression. We then calculated the FM coefficients and t-statistics as the mean and standard error of this time series.

and *STREAK* respectively. We obtain similar results in the FM procedures (reported in table 6). The t-statistics are respectively 2.65 and 2.45. In the fixed effect regressions the coefficient on *FREQ* and *STREAK* are significantly positive at the 10% level (the t-statistics are respectively 1.73 and 1.78). The economic magnitude of the phenomenon is such that analysts subject to the full overconfidence (*FREQ* equals 4) are expected to take 13 percent higher *DEV* than the average deflated *DEV*.[21]

As expected, the coefficients on *Current Median Deviation (CMD)*, are positive. In addition, *CME* (*Current Median Error*) is positive and significant. One possible explanation is that analysts take more risk on companies that are harder to predict. The coefficient on *LOSS* is negative but only marginally significant in the FM regressions. Throughout the analysis, the coefficients of *ACCURACY* are consistently negative and significant, implying a negative relation between analysts' skills and the risk they take. The coefficients on *RANK ORDER* (*O*) and *TIME (T)* are negative, but are insignificant in the FM regressions. The coefficient on size (*MCAP*) is negative (indicating that the dispersion of forecasts is smaller for bigger firms). The R^2 in the pool regressions are around 38.8 percent. [22] *EXPERIENCE* is insignificant in the pool and FM regressions, but positive in the fixed effect ones.

IV.2 Overconfidence and ERROR.

Columns III and IV in Tables 5, 6 and 7 indicate that analysts increase their risk after a stream of successes. The results of the *ERROR* regressions support the idea that prediction error increases after an analyst's stream of success. The coefficients on *FREQ* and *STREAK* are both significant at the 1% and have the predicted positive sign.[23] In fact, the significance of the coefficients on

---

[21] We multiply the value of the coefficient (0.0000743 from Table 5) by (4 -1.235), that is the maximum value of *FREQ* minus its expected value and divide the product by the average deflated risk (0.0016 from Table 1). The result for *STREAK* is around 19%.

[22] Auto-correlation does not appear to be an important issue. The Durbin-Watson (DW) statistics is between 1.9 and 2.1. Therefore, we cannot reject the hypothesis of serial independence. It is also possible to correct for autocorrelation in the FM procedure along the lines outlined in Fama-French [1999]. The first order autocorrelation of the coefficient series is relatively small (0.107 and –0.046 for *FREQ* in the *DEV* and *ERROR* regression respectively), yet, the standard errors are large (around 0.12). The significance level does not change much even after adjusting the standard errors for a 0.34 (that is 0.107 plus two standard errors) first-order autocorrelation. This adjustment only implies that we now require t-statistics around 2.3 (instead of the usual 2.0. The reason is that with first-order autocorrelation around 0.347, the variance of the average slopes, calculated assuming serial independence of the quarterly slopes, are too small since it represents $(1-0.347^2) = 0.88$ of the true variation. So the standard errors should be inflated by 13.6%, which is equivalent to requiring a t-statistics around 2.27 instead of 2.

[23] Introducing *CURRENT DEV* as control variable leads to similar results. The coefficient on *CURRENT DEV* is significantly negative (t-statistic=-78.8). Other coefficients are essentially unchanged, although the t-statistics for all

FREQ and *STREAK* in the *ERROR* regressions is extremely robust throughout the paper to various robustness checks that we perform. The economic magnitude is such that analysts subject to the full overconfidence (*FREQ* equals 4) are expected to have 9 percent higher *ERROR* than the average deflated *ERROR*.[24]

The results in the *ERROR* regressions can be linked to the ones for *DEV* (columns I and II). Analysts who take more risk make more mistakes on average. For example, a univariate regression of *ERROR* on *DEV* (results not reported) shows a strong positive relation (t-statistics = 42.2). As expected, current errors are positively correlated to *Current Median Errors* and to the existence of losses. The coefficient on *Current Median Risk* is negative, but insignificant in the FM regressions. The coefficient on *ACCURACY* (A) is significantly negative (good analysts tend to make fewer mistakes) in the pool and FM regression. The coefficients on *RANK ORDER* (*O*) indicate that earlier predictions are subject to more errors. *TIME* is positive in the pooled regressions, but the significance vanishes in the FM regressions. The coefficient on EXPERIENCE is negative, but only marginally significant in the pool regression. This result is generally consistent with Jacob et al [1999] who find no effect of learning-by-doing.[25] The R^2 reaches 79%, which suggests that the regression succeeds in capturing the expected error in prediction.

The combination of the *DEV* and *ERROR* regressions supports the hypothesis that analysts become overconfident after a series of successful predictions.

---

variables are increasing (including for *FREQ* and *STREAK*). Similar results are true for introducing *CURRENT ERROR* in the DEV regressions.

[24] Analyst subject to full overconfidence is expected to have an *ERROR* higher by 7 percent in the case of STREAK.

[25] These results suggest that analysts do not improve over time, but that the inferior ones are "weed-out" as time passes. Further analysis would be required, however, to verify this point.

V. Further Analysis and Robustness Checks[26]

IV.1 Additional Controls for Information Availability and Sensitivity Tests.

Clement [1999] finds that forecast accuracy is negatively associated with the number of firms and industries followed. We therefore control for the number of analysts following a firm, the number of firms and sectors followed by an analyst, the number of analyst working for the analyst's employer or the number of sectors covered by the employer. Those variables are often not significant. In addition, we control for period specific effects and seasonality by including quarter (three variables) and year dummies (seventeen variables) as well as a time trend. Our results, especially concerning, *FREQ* and *STREAK* are not affected.

In calculating *FREQ* and *STREAK*, we have considered four lagged predictions. We posit that the overconfidence should be a short-term phenomenon. The choice of the length of the period represents a trade-off between obtaining more variations in the *FREQ* and *STREAK* variables at the expense of diluting the effect. Nevertheless, we re-compute *FREQ* and *STREAK* using different numbers of lagged periods (from 2 to 8 quarters). All coefficients for *FREQ* or *STREAK* are positive. In the Error regressions, *FREQ* is significant in all regressions, *STREAK* up to four quarters. Results in the *RISK* regressions are more sensitive to the length of period and are typically not significant in shorter periods.

V.2 Overconfidence versus Under-confidence

Psychologists have found that individuals both overweight their success and underweight their failures. The literature also found that the phenomenon is asymmetric: individuals tend to overweight success more than they underweight failures. For example, Fiske and Taylor [1991] state that "self-enhancing attribution for success are more common than self-protective attributions for failures." Thus, under-confidence should be less significant than over-confidence.

---

[26] Results are not reported.

This prediction is consistent with the results from Table 3. Under the null hypothesis that the chance to be above the expected level of precision is equal to the chance of being below, the steady state distribution of *FREQ* would be symmetric and center around 2. Under the hypothesis that both over and under-confidence exist to the same extent, the distribution would remain symmetric but would be platykurtotic ("thinner tails"). Yet, the actual distribution of *FREQ* is left skewed. It less likely for *FREQ* to reach higher values than expected, but the probability of having a low value of *FREQ* remains high.

To further test this hypothesis, we compute an ordinal specification for *FREQ* and *STREAK*. We replace *FREQ* by four dummy variables for each of its values. If over and under-confidence are equally present, the value of the coefficients on the dummy variables should be equal. However, the estimation shows that the coefficients are generally increasing with the value of *FREQ* or *STREAK*.

We also re-compute *STREAK* using the length of the period during which the analyst had an error below the median one. Both in the pooled and in the fixed effect regressions, the coefficient is negative. However, compared to the regressions where *STREAK* measures superior performances, the coefficient has a smaller magnitude (about two to three times smaller) and a lower significance (the coefficient ceases to be significant fixed effect regressions) in the *DEV* regressions. Results in the *ERROR* regressions, although a bit less significant, are essentially unchanged.

V.3 Relative Error and Past Error.

Although we control for the relative performance of other analysts, we use the absolute level of error as a dependent variable instead of a relative one.[27] As a robustness check, we employ a dummy variable that measures the relative performance of the analyst instead of using the absolute level of error in the *ERROR* regressions.[28] This variable takes the value of one if the error is above the median error, zero otherwise. Consistent with previous results, this logistic regression indicates that *FREQ* and *STREAK* are significantly positive.

---

[27] We do so to abstract from the fact that analysts may mimic the prediction of a fellow analyst they believe to have superior information. Therefore, tests based on relative error may incorrectly measure the accuracy of the prediction.

[28] This test also provides an additional control for factors affecting all analysts at a given point.

To control for the magnitude as opposed to the frequency of past success, we add *Past Error (PE)* (mean past error over the last four prediction) as an additional variable. We also include *Past Risk, Past Median Error* and *Past Median Risk* (all calculated as the mean of the variables over the last four periods) as control variables. The sign and significance of *FREQ* and *STREAK* are not affected. *PE* is insignificant in most regressions, except in the *ERROR* fixed effect ones, where it is negative as expected. *PR* is significantly negative in all the ERROR regressions. In the pool and FM regressions, it is positive. However, in the fixed effect ones, which offer a better control for risk aversion, the coefficient is significantly negative.

V.4 Positive vs. Negative Error

Perhaps the most documented bias in analyst behavior is that they tend to make upwardly biased forecasts on average. DeBondt and Thaler [1990]), for instance, concludes that analysts are generally overly optimistic. To test whether this bias is related with our findings,[29] we create two variables, constructed in a similar way to *FREQ* and *STREAK*, that sum up the number of times an analyst made a negative error (forecast was below realization) in the past four periods.[30] The new variable is significantly negative in the *ERROR* fixed effect regressions and in the *RISK* pool (but not in the FM) regressions. This might be interpreted as weak evidence that pessimistic analysts might have lower subsequent errors. Results for *FREQ* and *STREAK* are essentially unaffected.

VI.5 Overconfidence Decay.

To try to study the decay of overconfidence, we consider the differences in results obtained with the two specifications: *FREQ* and *STREAK*. *STREAK* implicitly assumes that the overconfidence decay is immediate after one bad performance. *FREQ*, on the other hand, implicitly assumes a symmetric pattern where having one prediction above median performance has the same effect as having one below median (with opposite signs). Table 5, 6 and 7 indicate that, although the

[29] For example, a given analyst may be systematically relatively optimist or relatively pessimistic. An optimistic analyst would do better when there is good news and worse when there is bad news. If the good (or bad) news is serially correlated, last period performance may be an indicator of the current period performance.

[30] There exists an additional motivation for this robustness test. Since *ERROR* is calculated in absolute terms, it does not differentiate between above or below realization. However, optimistic analysts may behave differently from pessimistic ones. It is possible that they purposely bias their forecasts (for example to obtain favors from management). Thus, the benchmark for measuring the analyst accuracy should be different from earnings realization.

results are quite similar for the two models, the magnitude and the significance of the coefficients are slightly larger for *FREQ* than for *STREAK*. This would suggest some persistence in overconfidence.

To further understand this issue, we create a dummy variable, which is equal to one if *FREQ* is increasing compared to last period and zero otherwise. We add the product of the dummy variable and *FREQ* in the regressions. The product is insignificant in the *DEV* regressions, but is negatively significant in the *ERROR* regressions (with the exception of the pooled regression). *FREQ* remains of significant. These results suggest that the decay is somewhat slower than the rise but that difference is probably not significant.

VI. Conclusion

We rely on two major behavioral principles to devise our research question: do analysts become overconfident in their ability to predict future earning after a series of good predictions? Overconfidence in this setting implies that analysts overweight their own estimate, and rely less on public signals. Therefore, they are more likely to be out of consensus and to have a larger prediction error on their subsequent forecast.

To test these hypotheses, we regress both departures from the forecast consensus and forecast errors on the number of "good" predictions (i.e. error lower than the median error) in the last four quarters and control variables. We find that both deviations from consensus and prediction errors are positively correlated with past success. Additional control variables such as experience do not affect this bias. By focusing on relative performance and by controlling for earnings surprise and write-offs, this paper offers an original research design to investigate analysts' forecasts dynamics. We document a new bias: analysts' behavior follows a short-term cyclical pattern. We explain this results by relying on overconfidence as a short-term phenomenon that recurrently appears and disappears. Its intensity varies with the series of performance. However, we cannot totally reject rationality because the existence of irrationality (for example, over-optimism or overconfidence) is always conditional on the objective function of the analysts or of the brokerage houses. Therefore, our results can also be explained by making ad hoc assumptions on the reward function of the brokerage house or of the analysts.

However, our two tests limit compensation structures that are consistent with rationality: assuming that the analysts' reward function is increasing in accuracy and that deviation from consensus is rewarded if it is associated with lower error, our results indicate that this function would have to be both path dependent and asymmetric.

Yet, explaining apparent bias in analysts' forecasts by their reward function leaves important questions open: why do analysts face such a reward function and is this function a result of frictions imposed by market structure or a result of irrationality exhibited by brokerage houses or investors. For example, suppose brokerage houses' goal is to generate commission through trades. Suppose also that investors are fixated on analysts' forecasts[31] or alternatively that is too costly for them to acquire private information beyond analysts' forecasts. Analysts might be tempted to capitalize on recent success to move investors' expectations by issuing an out-of-consensus forecast. Thus, a possible extension of this paper is testing whether the market reacts differently to overconfident analysts. However, an assumption of analyst rationality calls, in turn, for modeling the economics (or the psychology) behind the reward function. For example, O'Brien and McNichols [1997] provide a rational explanation for apparent over-optimism. Yet, little is known about the economics of analyst's forecasts: it is not clear why an organization (i.e. the brokerage house) would devote important resources to forecast earnings and then disseminate this information for free.

[31] To our knowledge, no one has proposed an investor fixation on analysts' forecasts. However, cases of functional fixations are not unknown in the literature. Hand (1990), for example, posits that investors are functionally fixated on earnings.

Bibliography

Abarbanell, Jeffrey S., and Victor L. Bernard, "Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior", *Journal of Finance* 47, 1992, p1181-1207

Abarbanell, Jeffrey S., and Reuven Lehavy, "Biases Forecasts or Biased Earnings? The Role of earnings Management in Explaining Apparent Optimism and Inefficiency in Analysts' Earnings Forecasts", *University of North Carolina and UC Berkeley Working Paper*, 2000

Anderson, Lisa R., and Charles A. Holt, "Information Cascades in the Laboratory", *The American Economic Review*, December 1997, p847-862.

Barber, Brad, and Terrance Odean, "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics*, February 2001, p261-292.

Beckman, L., "Effect on Student's Performance on Teachers' and Observers' Attribution of Causality", *Journal of Educational Psychology*, 1970, 61, p76-82.

Bloomfield, Robert J., Susan D. Krische, and Mark W. Nelson, "Confidence and Investors' Reliance on Disciplined Trading Strategy", Cornell University Working Paper, 2000

Bloomfield, Robert J., Robert Libby, and Mark W. Nelson, "Underreactions, Overreactions and Moderated Confidence", Journal of Financial Markets, Vol 3, 2000, p113-137.

Camerer Colin, "Individual Decision Making", *Handbook of Experimental Economics*, Edited by John H. Kagel and Alvin E. Roth, Princeton University Press, 1995, p591

Clement, Michael B., "Analyst Forecast Accuracy: Do Ability, Resources and Portfolio Complexity Matter? ", *Journal of Accounting and Economics* 27, 1999, p285-303.

Daniel, Kent, David Hirshleifer and Avanidhar Subrahmanyam, "Investor Psychology and Security Market Under- and Overreactions," Working Paper, 1998.

DeBondt, Werner F.M. and Richard M. Thaler, "Does the Stock Market Overreact?", *Journal of Finance*, Vol XL, no3, July 1986, 793.

DeBondt, Werner F.M. and Richard M. Thaler, "Do Security Analysts Overreact?", *The American Economic Review* 80, 1990, 52-57.

DeBondt, Werner F.M. and Richard M. Thaler, "Financial Decision-Making in Markets and Firms: A Behavioral Perspective", *Hanbooks of Operations Research and Management Science*, 1995, 385-410.

Easterwood, John C. and Stacey R. Nutt, "Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism?", Forthcoming *Journal of Finance*.

Elliott, John A., Donna R. Philbrick, and Christine I. Wiedman, "Evidence from Archival Data on the Relation Between Security Analysts' Forecast errors and Prior Forecast Revisions", *Contemporary Accounting Research* 11, 1995, 919-938.

Fama, Eugene F., and Kenneth R. French, "Forecasting Profitability and Earnings", *Journal of Business*, Vol 73, n2, 2000, p161-175

Fama, Eugene F., and James D. MacBeth, "Risk, Return and Equilibrium: Empirical Tests", *Journal of Political Econom*y n81, May 1973, p607-636

Fitch, Gordon, "Effect of Self-Esteem, Perceived Performance, Choice on Causal Attributions", *Journal of Personality and Social Psychology*, Vol 16, 1970, p311-315.

Francis, Jennifer and Donna Philbrick, "Analysts' Decisions as Products of a Multi-Task Environment", *Journal of Accounting Research* 31, 1993, p216-230

Friedman, Dan, "Monty Hall's Three Doors: Construction and Deconstruction of Choice Anomaly", *The American Economic Review*, 88, 1998, p933-946.

Gervais, Simon and Terrance Odean, "Learning to Be Overconfident", *Wharton School and University of California Working Paper*, 1998.

Griffin, Dale and Amos Tversky, "The Weighting of Evidence and the Determinants of Overconfidence", *Cognitive Psychology* 24, 1992.

Graham, John R. "Herding among Investment Newsletters: Theory and Evidence", *Journal of Finance*, 1997 February, p237-268.

Gu, Zhaoyang and Joanna Shuang Wu, "Earnings Skewness and Analyst Forecast Bias", *Carnegie Mellon and University of Rochester Working Paper*, 1999

Hand, John R. M., "A Test of the Extended Functional Fixation Hypothesis", *The Accounting Review*, 1990 October, Vol 85 No 4, p740-763

Hastorf, A., D. Schneider and J. Polefka, Person Perception, Reading, Addison-Wesley.

Heath Chip and Amos Tversky, "Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty," Contemporary Issues in Decision Making, North Holland, 1990, p93-123

Huck, Steffen and Joerg Oechssler, "Informational Cascade in the Laboratory: Do they Occur for the Right Reasons", *Journal of Economic Psychology (forthcoming)*, 2001

Hung, Angela and Charles Plott, "Information Cascades: Replication and an Extension to Majority Rule and Conformity Rewarding Institutions", *American Economic Review*, Forthcoming

Hwang, L., C. Jan and Sudipa Basu, "Loss Firms and Analysts' Earnings Forecasts Errors", *Journal of Financial Statements Analysis*, 1996, 1, p18-30.

Jacob, John, Thomas Z. Lys and Margareth A. Neale, "Expertise in Forecasting Performance of Security Analysts", *Journal of Accounting and Economics* 28, 1999, p51-82

Johnson, T.J., Feigenbaum, R. and Weiby, M., "Some Determinants and Consequences of the Teachers' Perception of Causation", *Journal of Educational Psychology*, 1964, 55, p237-246.

Kahneman, Daniel and Amos Tversky, " Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, March 1979.

Kahneman, Daniel, Paul Slovic and Amos Tversky, " Judgement under Uncertainty: heuristics and biases", 1982.

Keane, Michael P. and David E. Runkle, "Are Financial Analysts' Forecasts of Corporate Profit Rational?" *Journal of Political Economy* 106, 1998, p768-805

Keynes, John Maynard, "The General Theory of Employment, Interest and Money", Imprint New York, Harcourt, Brace, 1936.

Kothari, S.P., Jowell S. Sabino, and Tzachi Zach, "Implication of Data Restrictions on Performance Measurement and Tests of Rational Pricing", *M.I.T. and University of Rochester Working Paper*, 1999

Kraemer Carlo, Markus Noeth and Martin Weber, "Information Aggregation with Costly Information and Random Ordering: Experimental Evidence", *Universitaet Manheim Working Paper*, 2000

Kukla, Andy, "Attribution Determinants of Achievement-Related Behavior", *Journal of Personality and Social Psychology*, Vol 21, 19972, p197-207

McNichols, Maureen and Patricia C. O'Brien, "Self-Selection and Analyst Coverage", *Journal of Accounting Research* 35, 1997, p167-199

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, "The Development of Expertise: Do Security Analysts Improve their Performance with Experience?", *Journal of Accounting Research*, 1997, p131-57

Miller, Dale, T. "Ego Involvement and Attributions for Success and Failures", *Journal of Personality and Social Psychology*, Vol 434, 1976, p901-906

Ross, Michael and Fiore Sicoly, Egocentric Biases in Availability and Attribution", *Journal of Personality and Social Psychology*, Vol. 37, 1979, p322-336.

Sinha, Praveen, Lawrence D. Brown and Somnath Das, "A Re-Examination of Financial Analytss' Differential Earnings Forecast Accuracy", *Contemporary Accounting Research* 14, Spring 1997, pp1-42

Weiner, Bernard and Andy Kukla, "An Attributional Analysis of Achievement Motivation", *Journal of Personality and Social Psychology,* Vol 15, 1970, p1-20

Table 1: Descriptive Statistics for Analysts

Variable	Mean	Med	StdDev	Q3	Q1
DEV	0.0433	0.0150	0.1584	0.0450	0.0000
DIFLDEV	0.0016	0.0006	0.0057	0.0017	0.0000
ERROR	0.1633	0.0600	0.3340	0.170	0.020
DIFLERR	0.0059	0.0024	0.0103	0.0067	0.0008
FORECAST	0.5263	0.4400	0.5850	0.700	0.250
EARN	0.4906	0.4300	0.6633	0.710	0.220
FREQ	1.2353	1.0000	1.0080	2.000	0.000
STREAK	0.4825	0.0000	0.8632	1.000	0.000

Notes

DEV is the absolute value of the difference between the analyst forecast and the consensus (defined as the median of other analysts' predictions). *ERROR* is the absolute value of the difference between forecasted and realized earnings. *DIFLDEV* and *DIFLERR* are *DEV* and *ERROR* deflated by the price of the stock at the beginning of the quarter. *FORECAST* is the forecast made by the analyst and *EARN* the realized earning per share (Compustat items #19 and 9). *FREQ* is the number of above average predictions in the last four predictions. *STREAK* is the number of above average predictions in a row before the current prediction. The mean, standard deviation, median and all other statistics are computed for the entire sample that begins in the last quarter of 1980 and finishes in the last quarter of 1997. This yields 46,905 observations.

Table 2: Pearson correlation matrix

	DEV	*DIFLDEV*	*ERR*	*DIFLERR*	*FREQ*	*STREAK*	*FORECAST*	*EARN*
DEV	1.000							
DIFLDEV	0.804	1.000						
ERROR	0.212	0.122	1.000					
DIFLERR	0.148	0.191	0.773	1.000				
FREQ	0.067	0.051	0.064	0.041	1.000			
STREAK	0.047	0.038	0.042	0.030	0.645	1.000		
Forecast	-0.115	-0.213	0.173	-0.044	0.094	0.057	1.000	
EARN	-0.156	-0.227	-0.160	-0.339	0.073	0.042	0.832	1.000
ACCURACY	-0.040	-0.042	-0.066	-0.081	0.219	0.163	-0.005	-0.004

Notes

DEV is the absolute value of the difference between the analyst forecast and the consensus (defined as the median of other analysts' predictions). *ERROR* is the absolute value of the difference between forecasted and realized earnings. *DIFKDEV* and *DIFLERR* are *DEV* and *ERROR* deflated by the price of the stock at the beginning of the quarter. *FREQ* is the number of above average predictions in the last four predictions. *STREAK* is the number of above average predictions in a row before the current prediction. *FORECAST* is the forecast made by the analyst and EARN the realized earning per share (Compustat items #19 and 9). The correlation estimates are computed for the entire sample that begins in the last quarter of 1980 and finishes in the last quarter of 1997. All correlation estimates are significant with p-values lower than .0001 (except the correlation between *ACCURACY* and *FORECAST* as well as between *ACCURACY* and *EARN*, where the p-values are respectively 0.21 and .38).

Table 3: Transition Probability of FREQ

Panel A: *Estimated* Transition Probability Matrix for FREQ

	FREQ				
Lag *FREQ*	0	1	2	3	4
0	71.1%	28.9%			
1	20.6%	55.4%	24.0%		
2		32.5%	48.7%	18.9%	
3			45.2%	44.1%	10.8%
4				56.1%	43.9%
Unconditional Probability	25.2%	35.5%	26.3%	11.0%	2.1%

Implied mean of FREQ: **1.3**

Panel B: Transition Matrix Under 50% Conditional Probability of Superior Prediction

	FREQ				
Lag *FREQ*	0	1	2	3	4
0	50.0%	50.0%			
1	12.5%	50.0%	37.5%		
2		25.0%	50.0%	25.0%	
3			37.5%	50.0%	12.5%
4				50.0%	50.0%
Unconditional Probability	6.3%	25.0%	37.5%	25.0%	6.3%

Implied mean of FREQ: **2.0**

Notes

The table provides, in Panel A, the estimated transition matrix for *FREQ*, the unconditional (steady state) probability and the implied mean. *FREQ* is the number of above average predictions in the last four predictions. As a benchmark, Panel B reports the transition matrix if the conditional probability of superior forecast is equal to 0.5 for every analyst. The unconditional probabilities are computed by iterating the transition matrix until convergence. The implied mean of *FREQ* is calculated from the unconditional probability.

Table 4: Statistics of *DEV* and *ERROR* conditional on the value of *FREQ*.

Panel A: Statistics of *DEV* and *ERROR* conditional on the value of *FREQ*

	FREQ (no. obs)	*DEV*	*ERROR*	*DIFLDEV*	*DIFLERR*
Mean	0	0.032	0.134	0.0013	0.0054
Median	(12561)	0.030	0.130	0.0013	0.0058
Std		0.095	0.295	0.0039	0.0101
Mean	1	0.039	0.161	0.0015	0.0058
Median	(16987)	0.040	0.170	0.0016	0.0067
Std		0.126	0.343	0.0036	0.0102
Mean	2	0.053	0.181	0.0019	0.0062
Median	(11919)	0.050	0.200	0.0020	0.0071
Std		0.214	0.349	0.0076	0.0104
Mean	3	0.058	0.196	0.0020	0.0066
Median	(4636)	0.060	0.220	0.0022	0.0077
Std		0.133	0.342	0.0058	0.0106
Mean	4	0.087	0.216	0.0030	0.0069
Median	(803)	0.080	0.250	0.0026	0.0083
Std		0.444	0.399	0.0192	0.0105

Table 4 (Cont.)

Panel B: Statistics of *DEV* and *ERROR* conditional on the value of *STREAK*

	STREAK (no. obs)	*DEV*	*ERROR*	*DIFLDEV*	*DIFLERR*
Mean	0	0.040	0.155	0.0015	0.0057
Median	32294	0.040	0.160	0.0016	0.0065
Std		0.147	0.321	0.0046	0.0102
Mean	1	0.047	0.171	0.0017	0.0060
Median	9387	0.050	0.180	0.0018	0.0069
Std		0.163	0.356	0.0070	0.0100
Mean	2	0.054	0.195	0.0019	0.0067
Median	3232	0.050	0.210	0.0021	0.0076
Std		0.126	0.362	0.0041	0.0115
Mean	3	0.060	0.199	0.0021	0.0067
Median	1190	0.065	0.220	0.0023	0.0075
Std		0.129	0.372	0.0054	0.0114
Mean	4	0.087	0.216	0.0030	0.0069
Median	803	0.080	0.250	0.0026	0.0083
Std		0.444	0.399	0.0192	0.0105

Notes

The mean, standard deviation and median are computed for each group for the entire sample that begins in the last quarter of 1980 and finishes in the last quarter of 1997. The descriptive statistics are calculated for groups based on *FREQ* (Panel A) and *STREAK* (Panel B). *FREQ* is the number of above average predictions in the last four predictions. *STREAK* is the number of above average predictions in a row before the current prediction. *DEV* is the absolute value of the difference between the analyst forecast and the consensus (defined as the median of other analysts' predictions). *ERROR* is the absolute value of the difference between forecasted and realized earnings. *DIFLDEV* and *DIFLERR* are risk and error deflated by the price of the stock at the beginning of the quarter.

Table 5: Pooled Regressions of *DEV* and *ERROR*
(Standard errors are calculated using White Correction)

The Estimated Models:

$$DEV_t^i = \alpha + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

$$ERROR_t^i = \alpha + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

	Regression			
Variable	Dev	Dev	Error	Error
Intercept	0.681	0.724	2.579	2.726
	(3.43)	(3.65)	(5.18)	(5.21)
FREQ	0.074		0.187	
	(3.13)		(4.42)	
STREAK		0.085		0.123
		(2.22)		(2.62)
Current Mean Error (CME)	216.720	216.717	769.593	769.745
	(2.49)	(2.49)	(9.13)	(9.13)
Current Mean Dev (CMD)	775.433	776.280	-409.192	-405.982
	(7.86)	(7.87)	(-2.14)	(-2.11)
Time (T)	-0.007	-0.007	0.008	0.008
	(-3.33)	(-3.37)	(3.63)	(3.52)
Rank Order (O)	-0.224	-0.227	-0.402	-0.415
	(-2.16)	(-2.20)	(-3.55)	(-3.66)
EXPERIENCE	-0.004	-0.004	-0.006	-0.005
	(-1.32)	(-1.24)	(-2.00)	(-1.70)
LOSS	-3.089	-3.087	4.805	4.814
	(-2.03)	(-2.02)	(3.20)	(3.20)
Accuracy (A)	-289.813	-283.963	-1026.153	-994.266
	(-4.25)	(-4.20)	(-10.36)	(-10.24)
Log Market Equity (MCAP)	-0.076	-0.075	-0.172	-0.168
	(-4.71)	(-4.69)	(-4.15)	(-4.08)
Number of observations	46909	46909	46909	46909
R-square	38.83%	38.84%	78.48%	78.46%

Notes:

The dependent variable *DEV* (D^i_t) is defined as the absolute error of the forecast at time t for a given analyst and company. The dependent variable *ERROR* (E^i_t) is defined as the absolute error of the forecast at time t for a given analyst and company. The columns report the results of the *DEV* and *ERROR* regressions. The first and third columns include *FREQ* as independent variable whereas column two and four include *STREAK*. *FREQ* is the number of above average predictions in the last four predictions. *STREAK* is the number of above average predictions in a row before the current prediction. *Current Median ERROR (CME)* is the median error of the analysts (excluding the analyst making the forecast) who follow a given company for the current quarter. Similarly, *Current Median DEV (CMD)* is the median Dev of the analysts (excluding the analyst making the forecast) who follow a given company for the current quarter. *TIME (T)* is the number of days between the date of the forecast and the end of the quarter. *RANK ORDER (O)* is the ratio of the number of forecasts published before the analyst issues her prediction to the total number of forecasts published over the quarter. *EXPERIENCE* is the number of quarters when analysts have made predictions before they make the current forecast. *LOSS* is an indicator of whether the company reports negative earnings. *ACCURACY (A)* is minus the median (over the entire period of coverage) of the difference between the analyst error and the median error of all analysts for a given quarter and a given company. *Log of Market Equity (MCAP)* is the natural logarithm of the market value of firm equity. For readability, all coefficients are multiplied by 100 in the table. The sample covers the period 1980 to 1997. To be included the observation has to be made within the quarter and has to be the first forecast of the analyst for the company in the quarter. For a given company, quarters when less than four analysts issue a forecast are excluded. An observation is included if the analyst has made at least four previous predictions on a company.

Table 6: Fama MacBeth Regressions of *DEV* and *ERROR*

The Estimated Models:

$$DEV_t^i = \alpha + \beta FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

$$ERROR_t^i = \alpha + \beta FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \phi_2 \cdot A_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

	Regression			
Variable	Dev	Dev	Error	Error
Intercept	0.969	0.939	1.613	1.677
	(1.89)	(2.08)	(3.33)	(3.62)
FREQ	0.055		0.126	
	(2.65)		(5.09)	
STREAK		0.071		0.100
		(2.45)		(3.72)
Current Mean Error (CME)	82.220	82.238	894.898	894.877
	(3.12)	(3.12)	(34.05)	(34.09)
Current Mean Dev (CMD)	930.471	930.810	-59.951	-56.856
	(25.38)	(25.02)	(-1.54)	(-1.46)
Time (T)	-0.005	-0.004	0.000	0.000
	(-1.25)	(-1.27)	(-0.10)	(-0.08)
Rank Order (O)	-0.076	-0.054	-0.788	-0.777
	(-0.40)	(-0.31)	(-3.37)	(-3.44)
EXPERIENCE	-0.007	-0.007	-0.010	-0.009
	(-0.85)	(-0.84)	(-1.18)	(-1.08)
LOSS	-0.775	-0.772	2.089	2.096
	(-1.76)	(-1.76)	(4.35)	(4.35)
Accuracy (A)	-329.343	-326.061	-1033.018	-1012.786
	(-4.53)	(-4.31)	(-13.63)	(-13.38)
Log Market Equity (MCAP)	-0.077	-0.074	-0.077	-0.076
	(-2.61)	(-2.72)	(-3.26)	(-3.22)
Number of Quarters	58	58	58	58

Notes:

The regressions report estimation of the Fama MacBeth coefficients. The coefficients are calculated as follows: for every quarter we estimate the regression using OLS. Then we compute the average and the t-statistics based on the time series of the estimated coefficients from the previous stage. The dependent variable *DEV* (D^i_t) is defined as the absolute error of the forecast at time t for a given analyst and company. The dependent variable *ERROR* (E^i_t) is defined as the absolute error of the forecast at time t for a given analyst and company. The columns report the results of the *DEV* and *ERROR* regressions. The first and third columns include *FREQ* as explanatory variable whereas column two and four include *STREAK*. *FREQ* is the number of above average predictions in the last four predictions. *STREAK* is the number of above average predictions in a row before the current prediction. *Current Median ERROR (CME)* is the median error of the analysts (excluding the analyst making the forecast) who follow a given company for the current quarter. Similarly, *Current Median DEV (CMD)* is the median Dev of the analysts (excluding the analyst making the forecast) who follow a given company for the current quarter. *TIME (T)* is the number of days between the date of the forecast and the end of the quarter. *RANK ORDER (O)* is the ratio of the number of forecasts published before the analyst issues her prediction to the total number of forecasts published over the quarter. *EXPERIENCE* is the number of quarters when analysts have made predictions before they make the current forecast. *LOSS* is an indicator of whether the company reports negative earnings. *ACCURACY (A)* is minus the median (over the entire period of coverage) of the difference between the analyst error and the median error of all analysts for a given quarter and a given company. *Log of Market Equity (MCAP)* is the natural logarithm of the market value of firm equity. For readability, all coefficients are multiplied by 100 in the table. The regressions are run on a quarterly basis over the period 1980 to 1997. To be included the observation has to be made within the quarter and has to be the first forecast of the analyst for the company in the quarter. For a given company, quarters when less than four analysts issue a forecast are excluded. An observation is included if the analyst has made at least four previous predictions on a company.

Table 7: Fixed Effect Regressions with White Correction

$$DEV_t^i = \alpha^i + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$
$$ERROR_t^i = \alpha^i + \beta\, FREQ_t^i + \gamma_1 CME_t^i + \gamma_2 CMD_t^i + \nu_1 T_t^i + \nu_2 O_t^i + \phi_1 \cdot EXPERIENCE_t^i + \psi_1 LOSS_t^i + \psi_2 MCAP_t^i + \varepsilon_t^i$$

Variable	Regression			
	Dev	Dev	Error	Error
Intercept	0	0	0	0
	(0.00)	(0.00)	(0.00)	(0.00)
FREQ	0.065		0.119	
	(1.73)		(3.22)	
STREAK		0.066		0.097
		(1.78)		(2.65)
Current Mean Error (CME)	225.299	225.223	755.261	755.127
	(2.42)	(2.42)	(8.49)	(8.49)
Current Mean Dev (CMD)	763.429	763.653	-521.319	-520.960
	(8.12)	(8.13)	(-3.41)	(-3.41)
Time (T)	-0.010	-0.010	0.012	0.012
	(-2.93)	(-2.93)	(3.69)	(3.69)
Rank Order (O)	-0.460	-0.461	-0.233	-0.235
	(-3.03)	(-3.03)	(-1.50)	(-1.51)
EXPERIENCE	0.013	0.013	-0.010	-0.009
	(2.32)	(2.32)	(-1.60)	(-1.52)
LOSS	-3.819	-3.818	5.058	5.060
	(-2.08)	(-2.08)	(2.87)	(2.88)
Log Market Equity (MCAP)	-0.160	-0.159	-0.665	-0.664
	(-2.50)	(-2.48)	(-5.20)	(-5.19)
Number of observations	46,909	46,909	46,909	46,909
R-square	35.96%	35.97%	76.49%	76.49%

Notes:

The dependent variables are *ERROR* (E_t^i) and *DEV* (D_t^i). We consider the following explanatory variable: *FREQ*, *STREAK*, *Current Median Error* (*CME*), *Current Median DEV* (*CMD*), *TIME* (*T*), *RANK ORDER* (*O*), *EXPERIENCE*, *LOSS* and *Log of Market Capitalization* (*MCAP*). The variables are defined in Tables 5. The two first columns report the regressions of *DEV*. The first column includes *FREQ* as explanatory variable whereas column two uses instead *STREAK*. Similarly, column three and four report the results for the *ERROR* regressions. All regressions control for analyst/company fixed effects. For readability, all coefficients are multiplied by 100 in the table. The sample covers the period 1980 to 1997. The t-statistics, reported in parenthesis, are calculated using White correction for standard errors.