How Much Do Analysts Influence Each Other's Forecasts?

Jonathan B. Cohn
University of Texas at Austin
jonathan.cohn@mccombs.utexas.edu

Jennifer L. Juergens
Drexel University
jlj54@drexel.edu

September 2014

Abstract

This paper develops and applies a new approach to disentangling the influence of analysts on each other's earnings forecasts from the effects of correlated information shocks. We estimate that, on average, each cent a new forecast by an analyst is above (below) another analyst's most recent forecast causes the other analyst to revise her forecast upwards (downwards) by between 0.21 and 0.36 cents. More reputable analysts are more influential, while those that tend to be optimistic are less influential and are influenced more by the forecasts of other analysts. We do not find support for career concerns-driven herding or anti-herding. Finally, we find that more influential analysts are more likely to subsequently be ranked as All-Stars and to move from a less to more prestigious brokerage house, and less likely to leave the analyst profession, suggesting that influence is a desirable characteristic.

We would like to thank Aydoğan Altı, Keith Brown, Jie Cai, Alex Edmans, Grant Farnsworth, Heber Farnsworth, Wes Gray, Revansiddha Khanapure, Kelvin Law, Michelle Lowry, Roni Michaely, Uday Rajan, Vito Sciaraffia, Dennis Sheehan, Irena Stefanescu, Sheridan Titman, Joel Vanden, seminar participants at Drexel University, Penn State University, and the University of Texas at Austin, and participants at the 2012 Financial Intermediation Research Society conference and 2013 Mid-Atlantic Research Conference for helpful comments.

1. Introduction

A securities analyst may take many pieces of information into account when forecasting a company's earnings. One important and potentially influential piece of information is the set of recent earnings forecasts issued by other analysts for the same company. These other forecasts should influence an analyst's own earnings estimates if she believes that they are incrementally informative. Many have also argued that analysts are influenced by other analysts' forecasts because of career concerns, which create incentives to herd with other forecasts (e.g. Hong, Kubik, and Solomon, 2000; Jegadeesh and Kim, 2010).

But how much do analysts actually influence each other's forecasts? This question has proved difficult to answer. There is significant evidence that analysts tend to revise their forecasts towards those of other analysts (e.g., Hong, Kubik, and Solomon, 2000; Clement and Tse, 2005). While this could indeed reflect inter-analyst influence, it could also reflect different analysts responding asynchronously to common information shocks, such as the provision of new earnings guidance by a company's management. Being able to cleanly disentangle and quantify inter-analyst influence would afford us a better understanding of how analysts formulate earnings forecasts. It would also facilitate testing of hypotheses relating to both information diffusion among analysts and the effects of career concerns on forecasting.

This paper develops a new approach to estimating the effects of inter-analyst influence. We focus on sequences of three forecasts issued by two different analysts for the same company and earnings period.[1] A natural way to measure how much an analyst's forecast moves towards that of another analyst is to calculate the difference between the other analyst's forecast and the first analyst's most recent prior forecast, and then compute the fraction of this "gap" that is

[1] For the majority of our tests, we focus on daily influence and therefore eliminate forecasts which are released on the same day as any other forecast. We have time-stamped forecasts for a subsample of observations which allows us to investigate the intraday influence of analysts on a minute-by-minute basis.

closed by the first analyst's subsequent revision. To isolate the amount of this gap closure driven by the influence of the other analyst's forecast, we focus on the length of time between the first analyst's initial forecast in the sequence and the other analyst's new forecast. The longer this interval, the more common information one would expect to arrive between the two forecasts. As the interval goes to zero, the amount of common information between the two forecasts goes to zero as long as the arrival rate of common information does not go to infinity.[2]

The fraction of the gap that the first analyst closes as this interval goes to zero then is purged of the contaminating effects of common information shocks, and reflects only influence, as long as these shocks do not systematically cluster between forecasts that are close together in time. This is our identifying assumption. It is easy to see why large information shocks (e.g. an earnings announcement) might cluster *before* forecasts by two different analysts that are close together in time, as a large shock would induce both analysts to revise their forecasts quickly. However, it is unclear why information shocks should cluster *between* forecasts by two different analysts when these forecasts are close together in time, especially as forecasts take some time to prepare.

Translating this insight into a test is straightforward. Using data on analyst earnings forecasts collected from I/B/E/S, we construct a sample of forecast trios consisting of a forecast by one analyst, the next forecast by a different analyst for the same company and earnings period, and the first analyst's first subsequent revision. In our primary analysis, we measure the interval between the two forecasts in each trio in number of days, though we also conduct tests measuring the interval in minutes. For each observation, we compute the gap (the difference)

---

[2] We exclude sequences of three forecasts where two of the forecasts occur on the same day, as we do not, in general, observe the sequencing of forecasts within-day. We observe "time-stamped" forecasts for a subsample of our data, and use this subsample to examine intervals as small as one minute.

between the first two forecasts in the trio, and the fraction of this gap that is closed by the first analyst's revision (the third forecast in the trio).

We then regress the fraction of the gap closed on the number of days between the first two forecasts in the trio. The intercept from this regression captures how much the first analyst revises towards the forecast of the second when the interval between the first analyst's initial forecast in the sequence and the second analyst's forecast goes to zero.[3] From the arguments above, this represents an estimate of the second analyst's influence on the first. We allow the intercept to vary by ordered analyst pair (by including ordered pair fixed effects) to address concerns about the non-random position of analysts in a sequence. The slope from the regression also contains information. Specifically, it measures how much more the first analyst revises towards the second as this interval between the first two forecasts increases, and represents an estimate of the effect (per day) of common information shocks on both the second analyst's forecast and the first's subsequent revision.

Our regressions yield intercepts of 0.21 to 0.36, and slope coefficients on the forecast interval of 0.005 to 0.006. The intercept estimates imply that a new forecast c cents above (below) an analyst's most recent forecast causes her to revise her own forecast upwards (downwards) by a considerable $0.21c$ to $0.36c$ cents on average. This is between 75% and 83% of the average fraction of the gap closed in our sample. The slope estimate implies that 45 to 70 days of common information shocks would be required to produce an effect of similar magnitude. While influence drives the majority of co-movement in forecasts, common information shocks also play a significant role.

[3] We also add the second and third powers of the interval between forecasts to reduce the likelihood that a nonlinear relationship biases our intercept estimates.

We seek validation of our tests by examining variation in the average level of influence over time. Prior to the advent of databases such as First Call and widespread adoption of the Internet in the 1990s, analysts were less likely to be able to observe each other's forecasts quickly, limiting the possibility of influence.[4] After implementation of Regulation Fair Disclosure (FD) in late 2000, which is thought to have reduced individual analysts' access to preferential information, analysts are less likely to have unique information. Our estimates of influence should be highest in the intermediate period, when forecasts were observable in almost-real time and some analysts were likely to have an informational advantage. Consistent with this argument, we find that average influence is greater between 1992 and 2000 than it is before 1992 or after 2000.

We next use our approach to examine what characteristics make an analyst influential or more likely to be influenced by others. We find that more reputable analysts based on previously-used measures such as brokerage affiliation or all-star ranking by *Institutional Investor* magazine (e.g., Gleason and Lee, 2003, Clement and Tse, 2005) are more influential. We also find that analysts who tend to be more optimistic in their forecasts in general are both less influential and more likely to be influenced by others. The fact that they are less influential could indicate that their forecasts are seen as less credible, perhaps because analysts who bias their forecasts upwards do so in order to curry favor with managers. However, this does not explain why they are more influenced by the forecast of others.

We do not find that the tendency to be influenced varies with an analyst's experience. The career concerns literature has argued that agents have incentives to "herd" (i.e., to mimic each other's actions) to avoid standing out when they are worried about others' beliefs about

[4] Although First Call was originally introduced in 1984, it consisted of only nine founding brokerage firms. The Real-Time Earnings Estimate (RTEE) database was introduced in 1987, but it was not until 1992 that an additional 20 brokerage firms were added to the RTEE, increasing firm coverage by approximately 50% (Thomson Financial).

their skills. This literature often uses an agent's experience as a proxy for the severity of her career concerns, as others update more on an agent's skill based on outcomes when she has a shorter track record (e.g., Gibbons and Murphy, 1992; Hong, Kubik, and Solomon, 2000). While the validity of this proxy could be subject to debate, our results do not appear to support the argument that career concerns cause analysts to herd on each other's forecasts.

Finally, to further explore whether career concerns might affect the incentives of analysts have to mimic each other's forecasts, we examine the relation between influence and analyst career outcomes. As a preliminary step, we show that both the tendency to influence and be influenced are persistent over time. We then show that, controlling for average forecast accuracy, analysts are more likely to be ranked as all-stars by *Institutional Investor* magazine, more likely to move to more prestigious brokerages in the future, and less likely to leave the analyst profession if they exert more influence over other analysts. We do not find that an analyst's tendency to be influenced towards or away from others predicts any of these outcomes. This last result could explain why analysts facing stronger career concerns do not appear to be more influenced by other analysts: it does not actually affect career outcomes.

A number of papers present evidence that stock analysts' earnings forecasts tend to cluster (e.g., Hong, Kubik, and Solomon, 2000; Clement and Tse, 2005). Similar behavior has been observed in stock analyst recommendations (Welch, 2000), macroeconomic forecasts (Gallo, Granger, and Jeon, 2000; Lamont, 2002), and newsletter recommendation weights (Graham, 1999). Much of the debate has centered on whether or not this clustering is driven, at least in part, by forecaster career concerns. Such career concern-driven clustering would be consistent with models in which agents want to avoid "standing out" because they face adverse consequences if they are identified as low quality (e.g., Scharfstein and Stein, 1990; Zwiebel, 1995). Supporting the presence of such incentives, Hong, Kubik, and Solomon (2000) show that analysts with the least accurate forecasts are the most likely to be terminated, especially among inexperienced analysts.[5]

Other papers have raised questions about whether stock analysts herd by showing that, once rational updating based on the information content of existing forecasts is taken into account, analysts actually appear to consciously anti-herd (Zitzewitz, 2001; Bernhardt, Campello, and Kutsoati, 2006). That is, they overweight their own information and underweight the information in other forecasts. Such behavior is consistent with the argument of Prendergast and Stole (1996) that less experienced agents facing career concerns attempt to signal the precision of their information by overweighting it. However, Chen and Jiang (2006) conclude that the data is more consistent with apparent anti-herding being driven by analysts being excessively optimistic about the precision of their own information rather than deliberate deviation. Our finding that both more and less experienced analysts are influenced to a similar degree suggests that career concerns do not appear to drive analysts to either herd or anti-herd. The remainder of the paper is organized as follows. Section 2 presents the methodology that we employ in more detail. In Section 3, we describe the data and the sample. Section 4 presents the results. Finally, Section 5 concludes.

2. Methodology

Stock analysts typically begin issuing forecasts of a company's earnings per share (EPS) for a fiscal year at least one year before the fiscal year ends. As the year progresses, they periodically issue revised EPS forecasts, creating a series of forecasts by the different analysts

[5] Sciaraffia (2013) finds that analysts decrease risk-taking after poor performance to reduce the risk of being fired.

covering the company. Consider a sequence of forecasts within this series consisting of consecutive forecasts by two different analysts, as well as the first subsequent revision of the analyst issuing the first forecast in that sequence. For reasons that will become clear shortly, we refer to the analyst issuing the first of the consecutive forecasts as the "responding analyst" (or "responder"), and the analyst issuing the second of the consecutive forecasts the "influencing analyst" (or "influencer"). We use the subscripts 'R' and 'I' to denote the responder and influencer in our analysis. Let F_R and F_I denote the responder's and the influencer's forecasts in the pair of consecutive forecasts and F_R' the responder's first subsequent revision. Let t_R, t_I, and t_R' denote the time at which each of the respective forecasts is issued, and define $\Delta t = t_I - t_R$. Then, by construction, $t_R < t_I < t_R'$, and $\Delta t > 0$.

We define the difference or the "gap" between the values of the consecutive forecasts formally as $Gap = F_I - F_R$. This is the amount by which the influencer's forecast represents an innovation relative to the responder's most recent forecast. We define $Revision = F_R' - F_R$ as the size of the revision in the responder's first forecast after the influencer's forecast in the sequence. Finally, we define $GapClosed = Revision / Gap$. This is the fraction of the gap between F_R and F_I that the responder closes when she revises her forecast by issuing F_R'. This gives us a properly scaled measure of the degree to which analyst the responder's forecast moves towards analyst the influencer's. As an example, if the responder issues a forecast of F_R = \$1.00, the influencer subsequently issues a forecast of F_I = \$1.10, and the responder then revises her forecast to F_R' = \$1.03, then we would say that the responder closed 30% of the \$0.10 gap between the influencer's forecast and her own most recent forecast.

A positive value of *GapClosed* could indicate that the responder is responding to analyst influencer's forecast or that both the responder and influencer are observing and responding to similar information arriving between t_R and t_I. If these information shocks are distributed continuously across time, however, we can estimate the effect of the influencer on the responder by examining *GapClosed* as Δt goes to zero. Specifically, we can estimate influence using the following regression:

$$GapClosed_i = \alpha + \boldsymbol{\beta}' \boldsymbol{f}(\Delta t_i) + \varepsilon_i, \tag{1}$$

where i represents a specific observation, $\boldsymbol{f}(\Delta t_i)$ is a vector of polynomial functions of Δt_i (e.g., Δt_i, Δt_i^2, Δt_i^3), and $\boldsymbol{\beta}$ is a vector of slope coefficients. We include polynomial functions of Δt_i to allow for the possibility that the arrival of information is continuous but not constant. If information arrives at a constant rate, then $\boldsymbol{f}(\Delta t_i) = \{\Delta t_i\}$ would be appropriate. The intercept α from this regression represents an estimate of the average level of influence of the influencer's forecast on the responder's revision. The slope coefficient β represents an estimate of the effect of common information shocks on multiple analysts' forecasts per day of information.

One issue that deserves immediate attention is that the timing of forecasts is not likely to be random. This could lead to biased estimates if, for example, less-skilled analysts consistently wait to issue a forecast until right after a more-skilled analyst does. We address this by including ordered analyst-pair fixed effects in the regression. This is equivalent to estimating a separate intercept for each ordered pair. We then use the equal-weighted average of these intercepts as our estimate of the effect of influence. Including analyst pair fixed effects gives us regressions of the form:

$$GapClosed_i = \alpha_{R,I} + \boldsymbol{\beta}' \boldsymbol{f}(\Delta t_i) + \varepsilon_i, \tag{2}$$

where $\alpha_{R,I}$ is a constant specific to an ordered pair (R,I).[6]

[6] By ordered pair, we mean a responding analyst and an influencer analyst. So, for any two analysts, call them analysts A and B, there can be two separate ordered pairs: one in which A is the responder and B is the influencer, and one in which B is the responder and A is the influencer. While including ordered pair fixed effects is important for addressing concerns about the timing of forecasts, the results we obtain are almost unchanged if they are omitted.

We also test how influence varies with characteristics of the analysts in a pair, the firm being covered, and the forecasts themselves. We do so by allowing the intercept term to vary with these characteristics. These regressions are of the form:

$$GapClosed_i = \boldsymbol{\beta}' \boldsymbol{g}(\boldsymbol{x_i}, \Delta t_i) + \boldsymbol{\gamma}' \boldsymbol{x_i} + \varepsilon_i, \tag{3}$$

where x_i is a vector of characteristics associated with observation i, $\boldsymbol{g}(\boldsymbol{x_i}, \Delta t_i)$ is a vector consisting of all of the products of the elements of x_i and the elements of $f(\Delta t_i)$, and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of coefficients. We drop the subscript i going forward for brevity.

By construction, there are no interceding forecasts between the responder's first forecast in a sequence and the influencer's forecast. However, there can be – and in reality often are – interceding forecasts by other analysts between influencer's forecast and the responder's revision. For example, the patterns we observe in the data may actually be of the form R-I-Y-R, where Y is the forecast of a third analyst.[7] While the intervening forecast Y may also influence the responder's revision, this does not invalidate our approach. If the forecast Y is not itself affected by the influencer's forecast, then it represents noise with respect to our estimate of the influencer's impact on the responder. If the forecast Y is itself altered by the influencer's forecast, then the appearance that the responder is affected by the influencer could result from the responder being influenced by the Y forecast, who itself is affected by the influencer's forecast. However, even in this latter case, we are still capturing the influencer's total impact in this case, even if part of it is indirect and comes through the forecast of another analyst.[8]

Finally, our empirical approach relies on the ability of analysts to observe and respond to each other's forecasts quickly, either by preparing a new forecast or by adjusting an already-planned forecast. Since the early 1990s, analyst forecast databases, such as I/B/E/S and First

[7] It is also possible that the pattern is of the form R-I-I-R. We discuss this possibility in the next section when we consider how to operationalize *GapClosed*.
[8] As a robustness check, we perform analyses in Section 4 that eliminate interceding forecasts.

Call, as well as the Internet, have made forecasts observable in almost-real time. While much of the product released by analysts consists of lengthy reports that take time to compile, analysts frequently release brief, time-sensitive "call notes," which are immediately available to clients and analyst databases. Moreover, even lengthier reports can be adjusted as necessary to take into account incremental information. Therefore, an analyst can respond quickly to forecasts of another analyst or to time-sensitive common information, such as an earnings announcement, usually within minutes to hours of the information being released.

3. Data and Sample

The primary data used in this paper comes from I/B/E/S. We collect all one-year annual earnings per share forecasts for US companies between 1983 and 2009 from the unadjusted I/B/E/S details file.[9] This gives us 1,428,647 annual forecasts made by individual analysts in this time period. For each company and reporting period in the sample, we line up the forecasts sequentially. We then identify each pair of adjacent forecasts by two different analysts, as well as the first subsequent revision by the analyst issuing the first of these two forecasts. Each three-forecast sequence of this type represents an observation. In our main tests, where we count the time between forecasts in days, we exclude cases where either the first two or last two forecasts in the sequence were issued on the same day. We only include sequences where we can detect a subsequent forecast by the first analyst in the pair.

To be consistent with the notation in the previous section, we label the analyst issuing the first and third forecasts in the sequence the responder, and the analyst issuing the second forecast in the sequence the influencer. For each observation, we measure Δt in days as the difference in

[9] Following recent studies, we use the unadjusted details file and make adjustments for splits using CRSP split-adjustment dates and factors.

the reported dates of the two consecutive forecasts, F_R and F_I. We restrict the sample to observations where Δt is less than or equal to 60 days to avoid observations with very large Δt having an outsized impact on our estimates.[10] The resulting sample consists of 247,032 observations. In supplemental analysis, we use information about the exact time of each forecast to measure Δt in minutes. We restrict the sample in this case to observations where Δt is less than or equal to 48 hours.[11]

We compute *GapClosed* as described in the previous section. One practical issue in computing *GapClosed* is how to treat cases where the influencer issues a second forecast before the responder revises (e.g., when the pattern of forecasts is R-I-I-R). There are two assumptions that we could make here. The first is that the responder may have revised her beliefs about earnings but simply has not issued a new forecast yet. In this case, we would simply use the responder's revision when it does take place as the third forecast in the three-forecast sequence. We refer to this as the "Next Forecast" approach. Second, we could assume that the absence of a new forecast by the responder after the influencer's forecast represents a reaffirmation of her prior forecast. In this case, we would treat the responder's first forecast in the sequence as the third forecast as well, and *GapClosed* would be zero. We refer to this as the "Zero Change" approach. Rather than take a stand on which assumption is proper, we use both approaches in our analysis. In addition, we also conduct analysis where we simply omit observations in which the influencer issues two forecasts before the responder revises. We refer to this as the "Omitted Observation" approach. Thus, in total, we compute *GapClosed* in three different ways.

[10] The results are almost identical if we use a 10- , 30-, or 90-day window instead.

[11] Note that there can be overlap in the forecasts across sequences. Consider, for example, a sequence of four consecutive forecasts for a stock. Suppose that the first and third forecasts in the sequence are issued by analyst A and the second and fourth by analyst B. Then one observation consists of the first three forecasts in the sequence, with A the responder and B the influencer, and another observation consists of the last three forecasts in the sequence, with B as the responder and A as the influencer.

While our measure *GapClosed* is designed to capture the extent to which an analyst closes the gap between her prior forecast and that of another analyst, it should not be surprising that this measure routinely takes on absolute values much greater than one. For example, if the responding analyst forecasts $1.00, the influencer forecasts $1.01, and then the responder revises to $1.07, we would record a *GapClosed* of 700%. It would be unclear in this case that the responder is attempting to move towards the influencer substantially more than in the case where *GapClosed* is, say, 500%. We attempt to ensure that such large values of *GapClosed* are not overly-influential by winsorizing each of our three measures of *GapClosed* at the 5th and 95th percentiles. The exact point at which we winsorize the data has little effect on our estimates. As a robustness check, we also use a transformation of *GapClosed* into percentiles, so that only the ordering of *GapClosed* in the sample and not its actual magnitude matters.

Our empirical approach is designed to filter out the effects of information shocks that hit multiple analysts covering a stock at about the same time and that could therefore cause clustering in forecasts unrelated to influence. In one robustness check, we eliminate observations where a forecast is released within ±5 days of a quarterly earnings announcement, since these are likely to be particularly informationally-intensive periods. We identify earnings announcement dates from the I/B/E/S "Actuals" file.

We also test whether influence varies with analyst characteristics. Analysts may vary by quality or access to information about the firm, incentives, or sensitivity to career concerns. We construct five analyst-level variables that capture potentially important analyst characteristics. The first is whether the analyst works for a top-tier brokerage firm. The prestige of the brokerage house for which an analyst works may be a good indicator of her skill or informedness if analysts prefer to work for more reputable brokerages, other things being equal (Fang and

Yasuda, 2009). Using the Carter-Manaster (1990) and Carter, Dark, and Singh (1998) measures of broker reputation, we classify brokerage firms as "top-tier" if this measure is greater than 9 and "lower-tier" if this measure is less than 9. Although the measure is based on investment banking business and not specifically the quality of the research department, one might expect that reputable banks would have the preference and ability to hire the most skilled analysts.

The second analyst characteristic is whether an analyst is ranked by the *Institutional Investor's* All-American Research Team poll in a given year. Each year, we hand-collect analyst names, brokerage firms, placement, and industry from *Institutional Investor* magazine, and then hand-match this data to the I/B/E/S BRAN (broker-analyst) file using all available information. We are able to uniquely match more than 99% of the analysts in the *Institutional Investor* rankings data. Numerous studies have suggested that ranking in the *Institutional Investor* poll is an indication of analyst quality, including the ability to produce more accurate (Gleason and Lee, 2003; Hong and Kubik, 2003) and timely earnings forecasts (see, for instance, Fang and Yasuda, 2011). We broadly identify analysts as either "ranked" or "unranked" in a given year based upon their inclusion in the research poll.

The third analyst characteristic is whether the analyst has an investment banking relationship with the covered firm. Following Michaely and Womack (1999), among others, we identify the brokerage firm that was the lead underwriter in the firm's IPO. If this relationship exists, the analyst is identified as a "banker" analyst, while all other analysts are "unaffiliated." It is reasonable to assume that affiliated or banker analysts have better access to information about the firm in their capacity as arms-length insiders of the firm. Data on lead underwriters is collected from Thomson/SDC's "Global New Issues" database, and is hand-matched to the

I/B/E/S broker-analyst identification file (BRAN), which was only available to academics via WRDS until 2006.

The fourth analyst characteristic is the analysts' overall level of optimism about the stocks she covers. Studies have identified systematic optimism by analysts (see, for instance, Easterwood and Nutt, 1999), which in part can be accounted for by biases in forecasts. One reason for this systematic optimism is that analysts may have incentives to avoid issuing forecasts that might upset a company's management. For each company covered by a given analyst, we identify if the analyst is on average above or below the prevailing consensus estimate in each year. We then construct a variable called *optimism* that is equal to 1 if more than 50% of an analyst's forecasts in a given year exceed the consensus, and zero otherwise.

The fifth analyst characteristic is whether or not an analyst is experienced. One might expect analysts with longer career histories to be more skilled, either because their survival is an indicator that they are skilled, or because analysts become more skilled or develop better contacts as they gain experience. Lack of experience is also often used as a proxy for the severity of an agent's career concerns, since less is known ex ante about the skill of relatively inexperienced agents (e.g., Gibbons and Murphy 1992; Hong, Kubik, and Solomon, 2000). In each month, we measure the amount of time since the analyst's first forecast reported in the I/B/E/S data. The median career duration is approximately two years (means are slightly higher and range between 3.6 to 4.1 years). We partition the sample into "experienced" and "inexperienced" depending on whether the analyst's career duration at the time of her forecast is greater than or less than two years.[12]

[12] The analyst forecast data in I/B/E/S begins in 1970, thereby ruling out any concerns that our experience measures are driven by truncation due to I/B/E/S availability.

We also we examine how influence has evolved over time. We focus on three time periods: 1982-1991, 1992-2000, and 2001-2009. The first subperiod, 1982-1991, represents a period before the Internet and First Call made analyst forecasts observable to other analysts in almost-real time; the second, 1992-2000, represents a period after forecasts became more readily-observable but before the adoption of Regulation FD. Regulation FD in principle made it more difficult for any individual analyst to acquire private information by banning management from sharing information with only selected analysts. The final subperiod, 2001-2009, represents a period after Regulation FD limited such private information sharing.

Table 1 presents basic summary statistics for our sample period. In the median, analysts revise approximately every 70 days in our sample, indicating that forecasts are updated slightly more than once per quarter. As expected, positive (negative) forecast revisions earn significantly positive (negative) 3-day announcement returns. On average, approximately 26.2% of analysts work for top-tier brokerage firms, 12.4% of analysts are ranked by *Institutional Investor*, 3.4% are affiliated with the lead underwriter on the IPO, and 25.9% have more than two years of work experience.

Insert Table 1 here

Figure 1 shows the temporal distribution of Δt in our sample. It is not surprising that there are more observations for small Δt than for large Δt, as we are constructing observations by looking for two consecutive forecasts. We have re-run all of our tests excluding observations where $\Delta t = 1$ (implying that observations with $\Delta t = 2$ are the closest to $\Delta t = 0$), and the results are virtually unchanged.

Insert Figure 1 here
----

4. Estimates of Influence

We begin our empirical analysis by examining the distribution of *GapClosed*, without attempting to separate out the effects of causal influence from common information shocks in driving it. We next implement the regression approach described in Section 2 in order to estimate the effect of causal influence on forecast revisions. Then, we examine how influence varies over time and with the direction of the influencer's forecast relative to the responder's prior forecast.

4.1. The distribution of GapClosed

Figure 2 presents histograms of our three different measures of *GapClosed*. Note that there are large masses at the endpoints of the distributions because the measures are winsorized.

Insert Figure 2 here

Observe that *GapClosed* is not distributed evenly around zero. Rather it is right-skewed, with the size of almost every bin of positive values exceeding the corresponding bin of negative values. This is not surprising, as it merely confirms that analysts do indeed cluster in the sense that their forecasts tend to move towards rather than away from those of other analysts' forecasts. All measures of *GapClosed* also tend to cluster at whole numbers, especially one, and to lesser degree at the halfway point between whole numbers. This clustering is simply a function of the fact that forecasts are issued in whole cents. For example, if the influencer's forecast is $0.02 above the responder's prior forecast, then *GapClosed* can only take values that are multiples of 0.5. Finally, observe the large mass of *GapClosed* at zero when we use the Zero Change approach. This reflects the fact that we set *GapClosed* to zero when the influencer issues two forecasts before the responder's revision, which occurs in 31.2% of the observations in our sample.

Next, we present summary statistics for the distributions of the three measures of *GapClosed* in Table 2. These summary statistics confirm our conclusion from examining Figure 2 that *GapClosed* tends to be positive. Mean *GapClosed* lies between 0.287 and 0.438, depending on the approach used to compute *GapClosed*. On average, then, an analyst closes 29%-44% of the gap between a new forecast and her prior forecast in her next revision. All of the means are statistically different than zero at the one percent level based on a simple two-tailed t-test. The medians are slightly larger than the means, and are statistically different than zero at the one percent level based on a Wilcoxon Signed-Rank test.

Insert Table 2 here

Table 2 also shows the percentage of observations for which *GapClosed* is greater than, equal to, or less than zero. Each of the three measures of *GapClosed* is positive almost twice as often as it is negative, indicating that revisions tend strongly to be in the direction of a newly-observed forecast. Binomial probability tests reject equality of the proportion of positive and negative *GapClosed* observations at the one percent level. Overall, the distribution of *GapClosed* is consistent with evidence that analysts exhibit herding, in the sense that they revise, on average, towards the forecasts of other analysts. Note that the appearance of herding does not imply influence, as it could also be driven by common information shocks.

4.2. The causal influence of analysts' forecasts

We now turn to the primary focus of our analysis, which is estimating the average causal influence of an analyst's forecast on other analyst's revisions using the approach described in

Section 2. We begin by simply plotting the mean values of each of our three measures of *GapClosed* for values of Δt from 1 day to 60 days in our sample in Figure 3. This figure captures our main results in a simple form. First, all three show that *GapClosed* is generally increasing in Δt. This is consistent with our expectation that analysts' forecasts will co-move more when there is more time for common information shocks to affect both of their beliefs about a company's earnings. Second, the figure shows that *GapClosed* approaches between 0.20 and 0.35 as Δt approaches zero, depending on the measure. This implies that an analyst responds to a new forecast that is *c* cents above (below) her prior forecast by revising her forecast upwards (downwards) by between 0.20*c* and 0.35*c* cents on average.

Insert Figure 3 here

We next implement the regression analysis discussed in Section 2. We present estimates from a number of regression specifications based on equation (2) in Section 2. The main results are shown in Table 3. In Panel A, Δt (the time between the responder's initial forecast and the influencer's subsequent forecast) is measured in days. We show results separately for each of our three measures of *GapClosed*. For each measure, we estimate the model once using only the first power of Δt, which makes it easy to interpret the relation between *GapClosed* and Δt, and once using the first three powers of Δt, which allows for a more flexible relation, reducing the likelihood that nonlinearities result in biased intercept estimates. Each regression includes analyst-ordered pair fixed effects. The reported intercept in each is the average of the pair-specific intercept terms. p-values based on standard errors clustered at the ordered pair level are shown in parentheses below each point estimate.

Insert Table 3 here

Whether we include one or three powers of Δt has virtually no impact on the intercept estimate. Moreover, the coefficient on Δt is virtually unchanged when the higher order powers of Δt, which themselves are all statistically insignificant, are included. This suggests that the relation between *GapClosed* and Δt is approximately linear.

When we use the Next Forecast approach in constructing *GapClosed*, the intercept of the regression is 0.316 and the slope coefficient is 0.006 (column 1). The intercept represents our estimate of the average level of influence that analysts exert on each other. That is, we interpret it as implying that a new forecast induces an analyst to update her forecast in such a way that she closes 31.6% of the gap with respect to that forecast when she issues her next revision. For comparison, Table 2 shows that, on average, the responder closes 39.6% of the gap with respect to a new forecast by the influencer (using the Next Forecast approach). Thus the estimate implies that influence explains approximately 80% of the overall tendency of the responder's forecast to move towards the influencer's forecast.

The coefficient on Δt of 0.006 implies that having one extra day over which common information shocks can arrive results in analysts appearing to close the gap between their forecasts and those of other analysts by an extra 0.6%. Comparing this to our estimate of influence, slightly more than 50 days of common information shocks are required to produce the same gravitation towards a new forecast that the influence of that forecast itself produces.

Column 3 shows that the intercept falls to 0.216 and the slope coefficient on Δt falls to 0.005 when we use the Zero Change approach to computing the dependent variable. The attenuation in both occurs because the Zero Change version of *GapClosed* is sometimes the same as in the Next Forecast version, where the relation between *GapClosed* and Δt is positive and has a positive intercept, and is sometimes is set to zero for various values of Δt. The intercept is largest in column 5, where we simply omit observations where the influencer issues two forecasts before the responder revises. Overall, our results suggest that a new forecast causes other analysts to revise their forecasts 21% to 36% of the distance towards that new forecast.

In Panel B of Table 3, we repeat the regression from Panel A measuring Δt in minutes rather than days. In these regressions, we only use observations where $\Delta t \leq 2{,}880$ minutes (48 hours). Note that minute-level forecast data is only available from the I/B/E/S database starting in 1993. Although the magnitudes of the intercepts are slightly lower when we measure the interval in minutes rather than days, we continue to observe that new forecasts significantly influence other analysts on average.

To further verify the robustness of our results, we next test alternative specifications either based on or related to equation (2). For brevity, we use only the Next Forecast version of *GapClosed* in these tests, though the conclusions are the same if we use either of the other two versions of the variable. The results from testing these alternative specifications are presented in Table 4. The first column presents results from a linear probability model (with ordered pair fixed effects) in which the dependent variable takes a value of one if *GapClosed* is positive and zero otherwise.[13] The intercept, which measures influence, is 59%, which exceeds the values found in Table 3.

An additional concern is that, in spite of the fact that we winsorize *GapClosed* at the 5th and 95th percentiles, there are still outliers that are driving our results. We address this possibility by using the percentile of *GapClosed* in its empirical distribution in the sample instead of the value of *GapClosed* as the dependent variable. This percentile by definition lies between zero and one. We use a simple re-normalization to make the coefficients comparable to

[13] We lump cases where *GapClosed* = *0* together with cases where *GapClosed* < *0*. We have also tried omitting them and lumping them together with cases where *GapClosed* > *0*. Because they represent only one percent of all observations, how we treat them has virtually no effect on the results.

those in Table 3.[14] The results are shown in column 2 of Table 4. The intercept of 0.327, which measures influence, is slightly larger than the comparable intercept in column 2 of Table 3, Panel A.

Insert Table 4 here

Our objective is to estimate the effect of the influencer's forecast on the responder's revision. This effect could be direct, but could also occur indirectly through other analysts. Recall that, in constructing our sample, we do not exclude observations where there are intervening forecasts by other analysts between the influencer's forecast and the responder's next revision. It is possible that part of the influencer's effect on the responder results from the influencer affecting the forecast of another analyst, whose forecast in turn affects the responder's revision. This possibility of indirect influence does not invalidate our interpretation: we are still ultimately capturing the effect of the influencer's forecast on the responder, even if it is indirect. Nevertheless, we would like to also test the direct effect in isolation. We do so by excluding from our sample any instance where there is an intervening forecast and re-estimating equation (2). The results are shown in column 3. Although our sample size is reduced substantially to 39,593 observations, the intercept of 0.261 is only slightly smaller than the intercepts in Table 3. This suggests that analysts have a significant direct effect on each other's forecasts.

Finally, recall that our identifying assumption is that information shocks do not systematically cluster between two analysts' forecasts when those forecasts are close together in time. In principle, one could attempt to filter out all observations taking place in close proximity

[14] Let *GapClosed'* denote the percentile of *GapClosed* in the empirical distribution. To make the coefficients comparable, we first estimate the following regression using OLS: $GapClosed' = \lambda + \mu\ GapClosed + \eta$. We then transform *GapClosed'* by subtracting λ and dividing the difference by μ, and use this as the dependent variable in the regressions.

to any major news about the company involved. This is difficult to achieve in practice because it requires a definition of major news and a means of systematically identifying the date on which such news arrives. However, we can easily remove observations in proximity to one type of major news event: quarterly earnings announcements. A company's quarterly earnings announcement is likely to be an important source of information regarding the firm's expected annual earnings. Column 3 of Table 4 presents coefficients from a regression where we exclude any observations where the influencer's forecast occurs within ±5 days of an earnings announcement by the company. The intercept of 0.307 is very similar to those shown in the first two columns in Table 3, suggesting that our estimates of influence are unlikely driven by analysts responding to contaminating events.

4.3. Influence over time

Next, we examine how the influence of analysts on each other's forecasts has changed over time. There have been two major shifts in the information environment in which analysts operate that are likely to affect how influential they are. First, the wide-spread use of analyst forecast databases such as I/B/E/S and First Call in the early 1990s along with the growth of the Internet made it much easier for analysts to observe each other's forecasts in almost-real time. This is likely to have led to an increase in the average level of influence, as a forecast can only affect other analysts if these analysts are aware of it.

Second, Regulation FD, implemented by the Securities and Exchange Commission in October 2000, prohibited a company's management from sharing material non-public information with selected outside parties, including analysts. If successful, this regulation should have reduced the amount of private information that any analyst is likely to have. A number of papers present evidence that this is the case.[15] This decrease in private information is likely to reduce the influence of forecasts on average, as each forecast should contain less private information that is otherwise unobserved by other analysts.

We divide our 1983-2009 sample period into three sub-periods in order to examine variation in average influence over time. The first of these sub-periods, 1983-1991, represents the period before forecast information became widely and quickly available. As the change in this information occurred gradually over time, the choice of when to end this period is arbitrary. We choose 1991 in part because it makes the lengths of the three sub-periods equal. The second of these sub-periods, 1992-2000, is the period after information about analyst forecasts became readily-accessible but before the implementation of Regulation FD. The third sub-period, 2001-2009, is the post-Regulation FD period.

We then estimate equation (2) for each of the three sub-periods and present the results in Table 5. The average level of influence appears to be fairly large in all three sub-periods. However, consistent with influence being larger in the era after forecast information became readily-available but before implementation of Regulation FD, the intercept term is larger in the 1992-2000 sub-period (column 2) than in either the 1983-1991 (column 1) or 2001-2009 (column 3) sub-periods. Though not reported in the table, differences between the 1992-2000 intercept and the 1983-1991 and 2001-2009 intercepts are statistically significant at the 1% level.

Insert Table 5 here

4.4. Early vs. late year forecasts

[15] For example, analyst forecasts became less accurate after Regulation FD (Agrawal, Chen, and Chadha, 2006; Francis, Nanda, and Wang, 2006; Findlay and Mathew, 2006), the stock price impact of forecasts declined (Gintschel and Markov, 2004), and the impact of information asymmetries on spreads decreased (Eleswarapu, Thompson, and Venkataraman, 2004).

As a final step in this section, we examine how influence changes over time within year. One concern with our estimates is that they might be contaminated by the well-established tendency of firms to "walk down" analyst estimates towards the end of the fiscal year (Richardson, Teoh, and Wysocki, 2004). Successful walking down of estimates could result in clustering of forecasts in a short period of time. We therefore examine the influence of forecasts issued in the first and second halves of a fiscal year by estimating equation (2) separately for each period. Table 6 shows the results.

Insert Table 6 here

The first two columns show the results for forecasts issued in the first half of the fiscal year, and the last two show the results for forecasts supplied in the second half. The intercept estimate is similar across the two periods. Our results suggest that walk down is unlikely to drive our estimates of influence given that the intercepts are not detectably higher for forecasts in the second half of the year.

5. Influence and Analyst Characteristics

In the previous section, we provided estimates of average analyst influence across our entire sample. However, many of the interesting questions in the analyst literature, including how information diffuses among analysts and whether career concerns impact forecasts, relate to which analysts influence other analysts and which in turn are influenced. In this section, we investigate how influence and the tendency to be influenced relate to an analyst's characteristics.

As described in Section 3, our characteristic variables are indicators for whether an analyst belongs to a Carter-Manaster top-ranked brokerage firm, whether an analyst is ranked by the *Institutional Investor* All-American Research Team, whether or not an analyst is bank-affiliated, whether an analyst tends to be optimistic in her forecasts, and whether an analyst is experienced. We capture these characteristics for both the influencer and responder in each observation. This allows us to estimate the effect of each of these characteristics on both how influential an analyst is and how much the analyst is influenced by other analysts.

Noting that all of our characteristic variables are binary (e.g., All-Star ranked or unranked), we begin by simply estimating regression equation (2) separately for analysts with values of one or zero for each characteristic separately. Table 7 shows the results. We report only the intercepts from these regressions (our measure of average influence) for brevity, as well as the differences for the two groups of analysts in each split. Panel A shows results for influencer analyst characteristics. The differences here represent estimates of how much more influential an analyst with a certain characteristic is than an analyst without that characteristic. Panel B shows results for responder characteristics. The differences here represent estimates of how much more readily-influenced an analyst with a certain characteristic is than an analyst without that characteristic.

Insert Table 7 here

The first two rows of Panel A show that an analyst's forecasts are more influential on average if she works for a top-ranked brokerage or is an *Institutional Investor* All-Star analyst. This suggests that analysts with better reputations are more influential, though it does not indicate whether they are influential because they are reputable or vice versa. Numerous papers have shown that accuracy and price impact are positively related to analyst reputation (Gleason and Lee, 2003; Clement and Tse, 2005). Panel A also shows that analysts who tend to be pessimistic are more influential than those who tend to be optimistic (row 4). One possible explanation for this difference is that analysts discount the forecasts of their more optimistic peers because these peers may face conflicts of interest which skew incentives away from maximizing the accuracy of their forecasts (Hong and Kubik, 2003). Differences in influence between analysts when conditioned on bank affiliation (row 3) or experience (row 5) are small and statistically insignificant.

Panel B shows that differences across analysts with different characteristics are less pronounced for responding analysts. The only difference that is statistically significant at the five percent level is that between optimistic and pessimistic analysts. Optimistic analysts appear not only to be less influential, but also to be more readily-influenced by other analysts. Also noteworthy is the lack of difference in tendency to be influenced between more and less experienced analysts. To the extent that inexperience is a commonly-used proxy for the severity of career concerns an analyst faces, the lack of a difference here raises questions about whether career concerns cause analysts to herd on each other's forecasts, as some have suggested.

The characteristics on which we split analysts in Table 6 not surprisingly are somewhat correlated with each other. To assess which characteristics have incremental explanatory power over influence and tendency to be influenced, we conduct multivariate analysis by estimating regression equation (3). Recall that both the intercept and slope coefficients in equation (3) are allowed to vary with analyst characteristics, with the difference in intercepts captured by the coefficient on the characteristic. Table 8 presents the results. We omit the slope coefficients and their interactions with analyst characteristics from the table for brevity. Influencer characteristics, responder characteristics, and both are included as explanatory variables in columns 1, 2, and 3, respectively.

Insert Table 8 here

The results are roughly consistent with those in Table 7. As column 1 shows, an analyst continues to be more influential on average if she works for a top-ranked brokerage, though being ranked an All-Star no longer appears to explain an analyst's influence once we control for other characteristics. Optimistic analysts continue to be substantially less influential than more pessimistic analysts. As column 2 shows, an analyst is also more influenced by other analysts when she tends to be optimistic. As in Table 7, the small and statistically insignificant coefficient on the responding analyst's inexperienced indicator does not support career concern-based herding. When both influencing and responding analyst characteristics are included in column 3, we obtain similar results when each is examined independently.

6. Influence and Career Outcomes

In this section, we examine how the tendency to influence or be influenced by other analysts relates to an analysts' career outcomes. We do so by estimating analyst-level influence and tendency to be influenced on a biennial basis. As a first step, we examine whether these characteristics persist over time.

6.*1. Persistence of influence*

While analysts in general appear to influence each other, and this influence appears to depend to some degree on analyst characteristics, we have no evidence yet that it is the same analysts who consistently influence or are influenced by other analysts. If the same analysts reliably exert influence, then analysts who are influential in one period should also be influential in subsequent periods – i.e., their influence should be persistent. The same holds for the tendency to be influenced. We next test this persistence in influence.

To study persistence, we need to estimate an analyst's influence level over different periods. We do so by estimating equation (2) separately for each analyst in the sample over each two-year window starting with 1982-1983 and ending with 2008-2009 using all observations in which the analyst is the influencing analyst in a three-forecast sequence. The intercept from this estimation represents an estimate of the analyst's average influence over the period. We use two-year windows rather than one-year to increase the number of observations we can use to estimate intercepts and therefore reduce the noise in these estimates. We exclude cases where an analyst is the influencing analyst in a sequence fewer than ten times over the given two-year window, as intercept estimates are likely to be very noisy in these cases. In each two-year window, we sort analysts into deciles based on our estimates of their average influence. We then calculate the average influence across all analysts in each decile in the next two-year window. We repeat this for all two-year windows in our sample, and then compute an average influence next-period influence for each decile.

Figure 4a shows a plot of average next-period influence for each decile. The figure shows that average next-period influence generally increases with prior period influence decile. The five highest prior period influence deciles have the five highest next period average influence estimates. A regression line fitted to the points in the graph has a slope of 0.0048, which is statistically different than zero with a p-value of 0.014 based on a simple two-sided t-test. The intercept of the regression line is 0.1198. Together, these estimates imply that the predicted next-period influence of an analyst in the tenth decile in the prior period is approximately 36% greater than the predicted next period influence of an analyst in the first decile. Overall, the results provide strong support for the persistence of influence, suggesting that some analysts are systematically more influential over time.

Insert Figure 4 here

To test whether the tendency to be influenced is also persistent, we repeat this exercise, except that we focus on the responding analyst. That is, we estimate equation (2) separately for each analyst in the sample over each two-year window using all observations in which the analyst is the *responding* analyst in a three-forecast sequence. We then form deciles based on the regression intercepts, which capture tendency to be influenced, and examine next period tendency to be influenced. Figure 4b shows a plot of average next-period tendency to be influenced for each decile. The upward-sloping relationship that the figure shows indicates that tendency to be influenced is also persistent. The intercept and slope indicate that the predicted next-period tendency to be influenced of an analyst in the tenth decile in the prior period is approximately 50% greater than that of an analyst in the first decile.

6.2. Analyst influence and career outcomes

We now examine whether career outcomes an analyst experiences vary with her recent level of influence and tendency to be influenced. We do so by examining four different outcome variables. The first is whether or not the analyst gets ranked as an All-Star analyst by *Institutional Investor* magazine. This ranking is prestigious and widely-followed, and therefore has positive implications for an analyst's career. The second is whether or not she is promoted to a "better" brokerage. We define promotion as moving from a less to more reputable brokerage, where less and more reputable are defined as below and above median value of the Carter-Manaster (1990) and Carter, Dark, and Singh (1998) measure of broker reputation, or moving from a below-median sized to above-median sized brokerage. The third outcome is whether an analyst is demoted, which is defined as a move opposite of a promotion. The fourth outcome is whether the analyst is terminated. We define termination as disappearing from I/B/E/S, though obviously some departures represent voluntary career changes.

We estimate logistic models with each of these outcomes as dependent variables. The primary independent variables are the estimates of influence and tendency to be influenced described in Section 6.1 measured over the most recent prior two-year window. We also control for the accuracy of an analyst's forecasts by including the analyst's median forecast error (absolute value of the difference between forecast and actual earnings, scaled by stock price) for all of her forecasts over the prior year. For each outcome measure, we show two specifications: one with only the first lags of the explanatory variables, and another with the first two lags. Table 9 presents the results from these regressions.

Insert Table 9 here

Columns 1 and 2 of Table 8 shows that the likelihood that she is ranked as an All-Star in a given year increases with her recent level of influence over other analysts, controlling for the accuracy of an analyst's recent forecasts. Columns 3 and 4 show that these analysts are more likely to subsequently go to work for a more prestigious brokerage firm, while columns 7 and 8 show that these analysts are less likely to leave the analyst profession. These results indicate that more influential analysts tend to subsequently experience more positive career outcomes, perhaps suggesting that being influential identifies an analyst as being more skilled. We do not find evidence that any of the career outcomes we examine are related to an analyst's tendency to be influenced. This may explain why analysts facing stronger career concerns (i.e., less experienced analysts) do not herd more on other analysts' forecasts: doing so does not result in better career outcomes.

7. Conclusion

In this paper, we developed and applied a regression approach to disentangling the role of influence from common information shocks in driving stock analyst forecast clustering. In summary, we find that an analyst responds a forecast that is *c* cents above (below) her own most recent forecast by revising her subsequent forecast up (down) by between 0.21c and 0.36c cents. This accounts for 75% to 83% of the overall tendency to move towards other analysts' forecasts.

Using a variety of alternative specifications, including eliminating intervening forecasts and removing periods where common information shocks are likely to affect multiple analysts (i.e. earnings announcement periods), we confirm our basic findings. Further, we find that influence is stronger when analysts can observe each other's forecasts is almost real time and are more likely to have unique private information. Conditioning on a variety of proxies for analyst skill or quality, we observe that more reputable analysts are more influential, and that more optimistic analysts are less influential and more likely to be influenced by others. We do not find a relation between the tendency to be influenced and an analyst's experience, inconsistent with arguments that less experienced analysts tend to herd on the forecasts of other analysts because of career concerns. We also find that influence or the tendency to be influenced exhibits persistence across time. Further, we find evidence that being influential can affect career outcomes, including All-Star rankings and promotions to more prestigious brokerage firms, though we show little effect of the tendency to be influenced on analyst career prospects.

This is one of the first papers to distinguish the influence of analysts on each other's forecasts from common information shocks in driving co-movement in forecasts. Although it is beyond the scope of this paper to determine whether influence among analysts is valuable, quantifying the degree of influence, how it varies across analysts, how influence has changed over time, and how influence affects career outcomes are all important contributions.

Appendix A: Sample Construction Example

This appendix presents a brief example of our sample construction process using actual I/B/E/S earnings forecast data. The series of all analyst forecasts for DGII for the annual earnings that were announced on 11/19/98 are below:

Analyst	Forecast Date	EPS forecast
685	11/13/97	0.95
10367	12/8/97	0.80
685	2/2/98	1.15
10367	2/13/98	1.05
685	5/11/98	1.40
685	7/7/98	1.20
46894	7/8/98	1.36
10691	7/28/98	1.34
46894	8/18/98	1.28
46894	9/11/98	1.13
685	9/28/98	1.10
10691	10/7/98	1.13
10691	11/12/98	1.12

In this series of data, there are six pairs of consecutive forecasts by different analysts for which the next revision of the first analyst in the pair can be observed. These sequences, along with the *Gap*, *Revision*, and *GapClosed* for these six observations are as follows:

R	I	T_R	T_I	Δt	T_R’	Gap	Revision	GapClosed
685	10367	11/13/97	12/8/97	25	2/2/98	-0.15	+0.20	-1.33
10367	685	12/8/97	2/2/98	55	2/13/98	+0.35	+0.25	+0.71
685	10367	2/2/98	2/13/98	11	5/11/98	-0.10	+0.25	-2.50
685	46894	7/7/98	7/8/98	1	9/28/98	+0.16	-0.10	-0.63
46894	10691	7/8/98	7/28/98	20	8/18/98	-0.02	-0.08	+4.00
10691	46894	7/28/98	8/18/98	21	10/7/98	-0.06	-0.21	+3.50

Note that there are five unique (R, I) pairs: (685, 10367), (10367, 685), (685, 46894), (46894, 10691), and (10691, 46894). A dummy variable is created for each of these to use in analyst-pair fixed effects regression specifications.

References

Agrawal, A., M. Chen, and S. Chadha, 2006, Who is afraid of Reg FD? The behavior and performance of sell-side analysts following the SEC's Fair Disclosure rules, *Journal of Business* 79: 2811-2834.

Carter, R., F. Dark, and A. Singh, 1998, Underwriter reputation, initial returns and long run performance of IPO stocks, *Journal of Finance* 53: 285-311.

Carter, R. and S. Manaster, 1990, Initial public offering and underwriter reputation, *Journal of Finance* 45: 1045-1067.

Chava, S. and M. Roberts, 2008, How does financing impact investment? The role of debt covenants, *Journal of Finance* 63: 2085-2121.

Chen, Q. and W. Jiang, 2006, Analysts' weighting of private and public information, *Review of Financial Studies* 19: 319-355.

Clement, M. and S. Tse, 2005, Financial analyst characteristics and herding behavior in forecasting, *Journal of Finance* 60: 307-341.

Easterwood, J. and S. Nutt, 1999, Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism, *Journal of Finance* 54, 1777-1797.

Eleswarapu, V., R. Thompson, and K. Venkataraman, 2004, The impact of Regulation Fair Disclosure: Trading costs and information asymmetry, *Journal of Financial and Quantitative Analysis* 39, 209-225.

Fang, L. and A. Yasuda, 2009, The effectiveness of reputation as a disciplinary mechanism in sell-side research, *Review of Financial Studies* 22, 3735-3777.

Fang, L. and A. Yasuda, 2011, Are stars' opinions worth more? The relation between analyst reputation and recommendation value, working paper, INSEAD.

Findlay, S. and P. Mathew, 2006, An examination of the differential impact of Regulation FD on analysts' forecast accuracy, *Financial Review* 41: 9-31.

Francis, J., D. Nanda, and X. Wang, 2006, Re-examining the effects of regulation fair disclosure using foreign listed firms to control for concurrent shocks, *Journal of Accounting and Economics* 41: 271-292.

Gallo, G., C. Granger, and Y. Jeon, 2000, Copycats and common swings: the impact of the use of forecasts in information sets, working paper, Università degli Studi di Firenze.

Gibbons, R., and K. Murphy, 1992, Optimal incentive contracts in the presence of career concerns: theory and evidence, *Journal of Political Economy* 100: 486-505.

Gintschel, A. and S. Markov, 2004, The effectiveness of Regulation FD, *Journal of Accounting and Economics* 37: 293-314.

Graham, J., 1999, Herding among investment newsletters: theory and evidence, *Journal of Finance* 54: 237-268.

Hong, H., and J. Kubik, 2003, Analyzing the analysts: Career concerns and biased earnings forecasts, *Journal of Finance* 58: 313-351.

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

Kasznik, R. and B. Lev, 1995, To warn or not to warn: Management disclosures in the face of an earnings surprise, Accounting Review 70, 113-134.

Kerr, W., J. Lerner, and A. Schoar, 2011, The consequences of entrepreneurial finance: Evidence from angel financings, *Review of Financial Studies* forthcoming.

Lamont, O., 2002, Macroeconomic forecasts and microeconomic forecasters, 2002, *Journal of Economic Behavior and Organization*, 48: 265-280.

Michaely, R. and K. Womack, 1999, Conflict of interest and the credibility of underwriter analyst recommendations, *Review of Financial Studies* 12: 653-686.

O'Brien, P. and R. Bhushan, 1990, Analyst following and institutional holdings, *Journal of Accounting Research* Supplement: 55-76.

Predergast, C. and L. Stole, 1996, Impetuous youngsters and jaded old-timers: acquiring a reputation for learning, *Journal of Political Economy* 104: 1105-1134.

Richardson, S., S.H. Teoh, and P. Wysocki, 2004, The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives, *Contemporary Accounting Research* 21, 885-924.

Roberts, M. and A. Sufi, 2009, Control rights and capital structure: An empirical investigation, *Journal of Finance* 64: 1657-1695.

Sciaraffia, V., 2013, A tournament of equity analysts: Compensation, performance, and risk-taking behavior, working paper, University of Texas at Austin.

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

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

Zitzewitz, E., 2001, Measuring herding and exaggeration by equity analysts and other opinion sellers, working paper, Stanford GSB.

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

Figure 1
Timing of Analyst Forecast Revisions

This figure presents the empirical distribution of Δt (measured in days) in our sample, where Δt is defined as $t_I - t_R$, with t_I the date of the influencing analyst's forecast and t_R the date of the responding analyst's first forecast in the sequence.

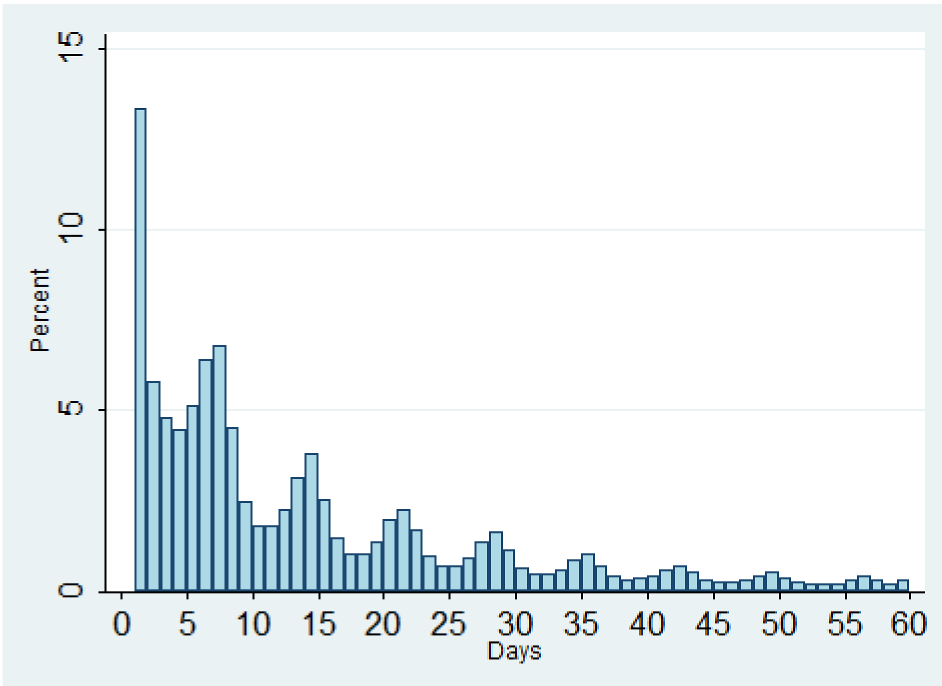
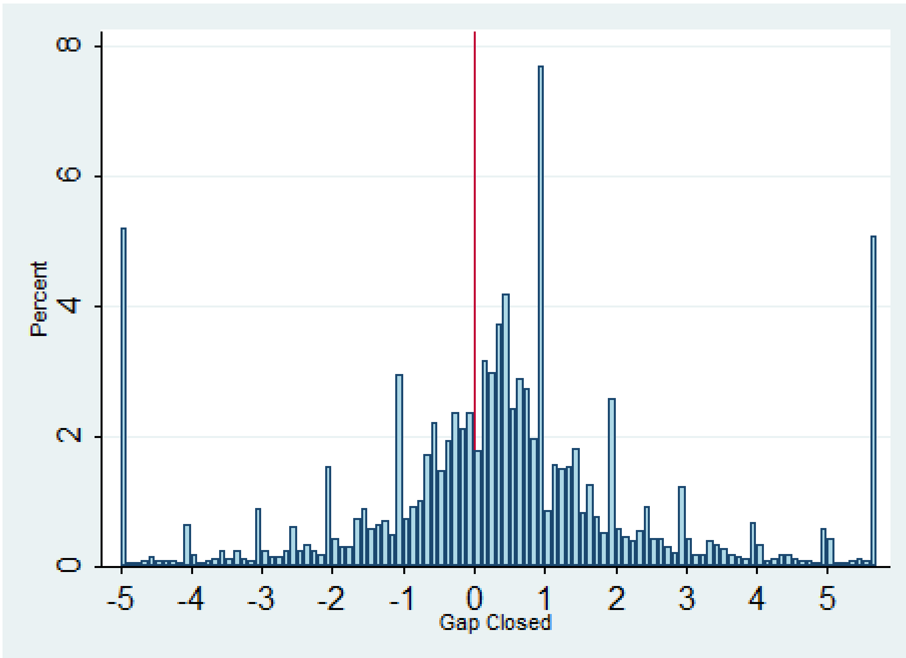


Figure 2

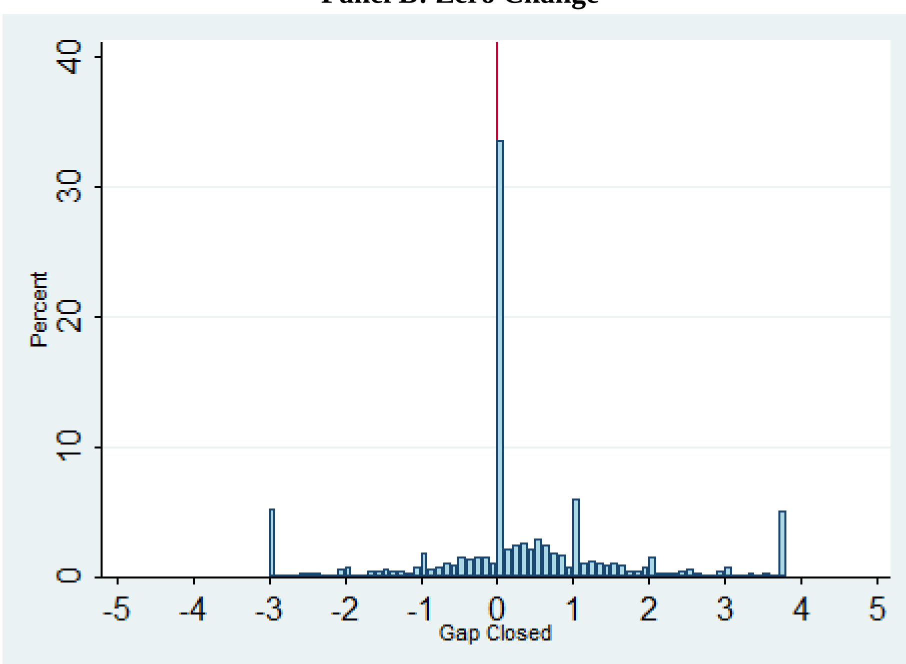
Patterns of Analyst Forecast Revisions

This figure presents the empirical distribution of *GapClosed* in our sample. *GapClosed* is defined as *Revision/Gap*, where $Gap = F_I - F_R$ and $Revision = F_R' - F_R$, with F_I defined as the influencing analyst's forecast, F_R as the responding analyst's first forecast in the sequence, and F_R' as responding analyst's first revision following the influencing analyst's forecast in the sequence. The mass of cases where $GapClosed = 0$ are excluded from the figure to make it easier to read.

Panel A: Next Forecast



Panel B: Zero Change



Panel C: Omitted Observation

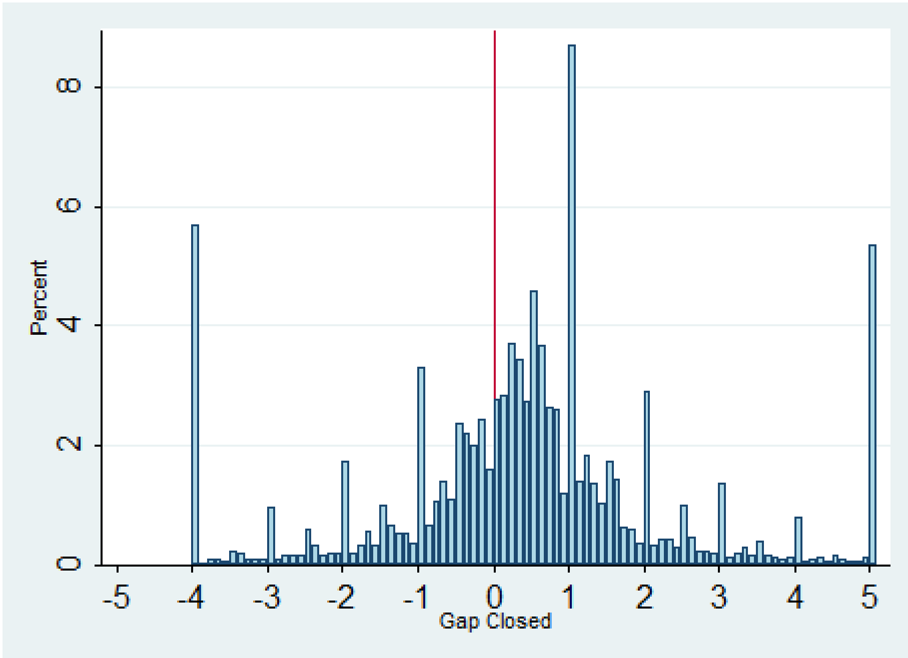
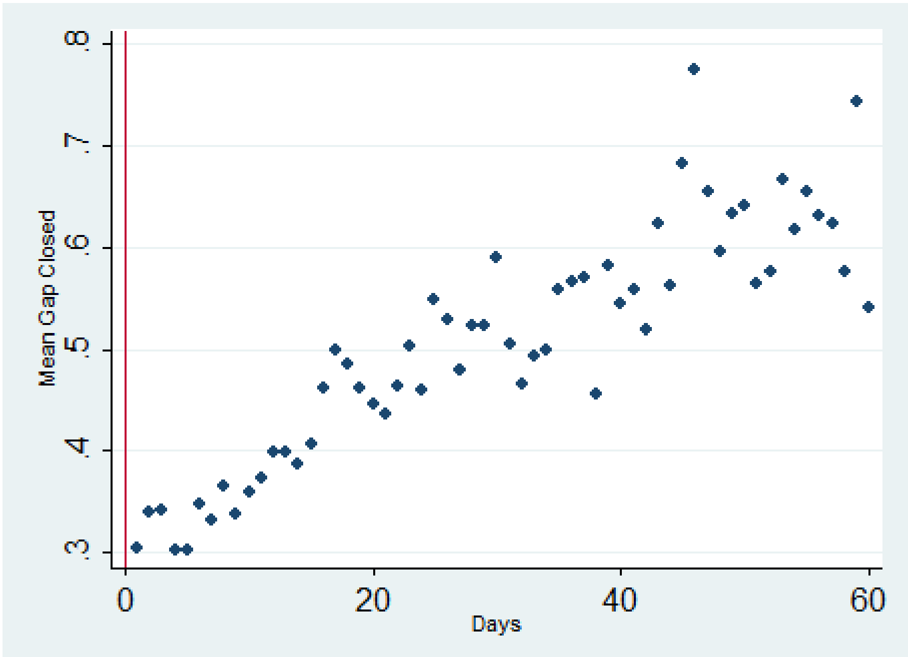


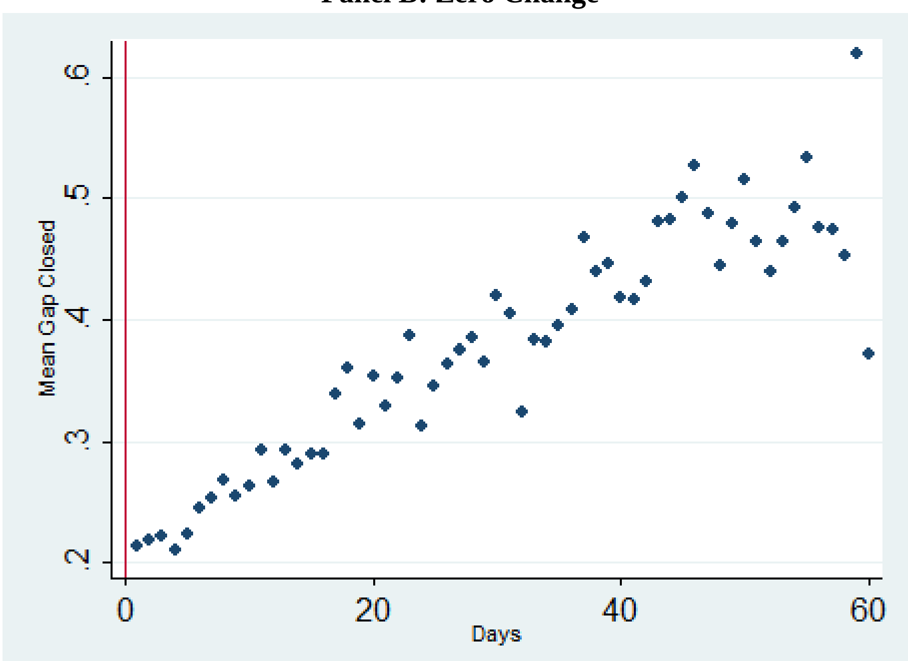
Figure 3
Fraction of the Gap Closed

This figure depicts mean *GapClosed* for each value of Δt (measured in days) between 1 and 60 in our sample.

Panel A: Next Forecast



Panel B: Zero Change



Panel C: Omitted Observation

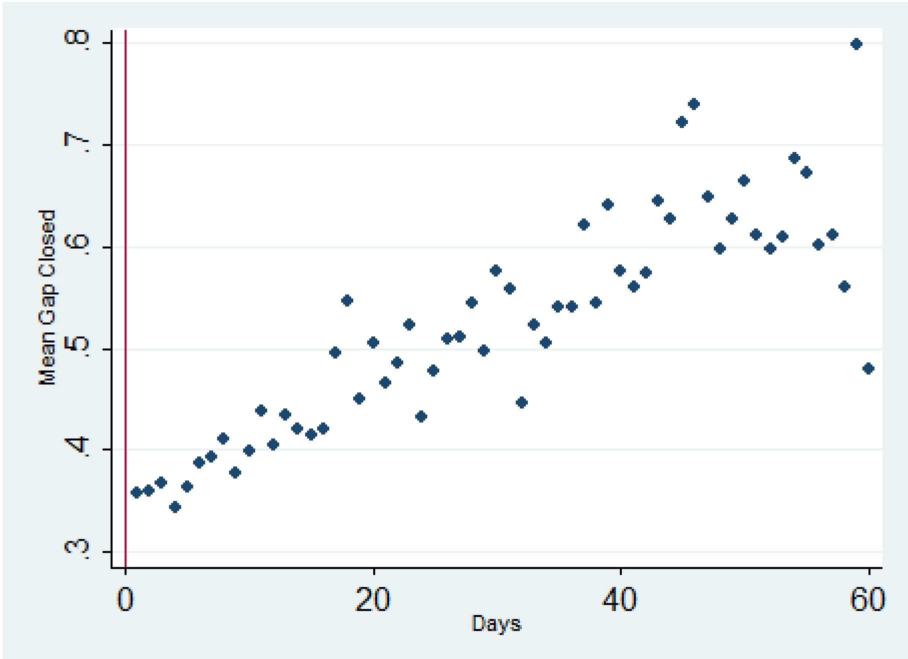
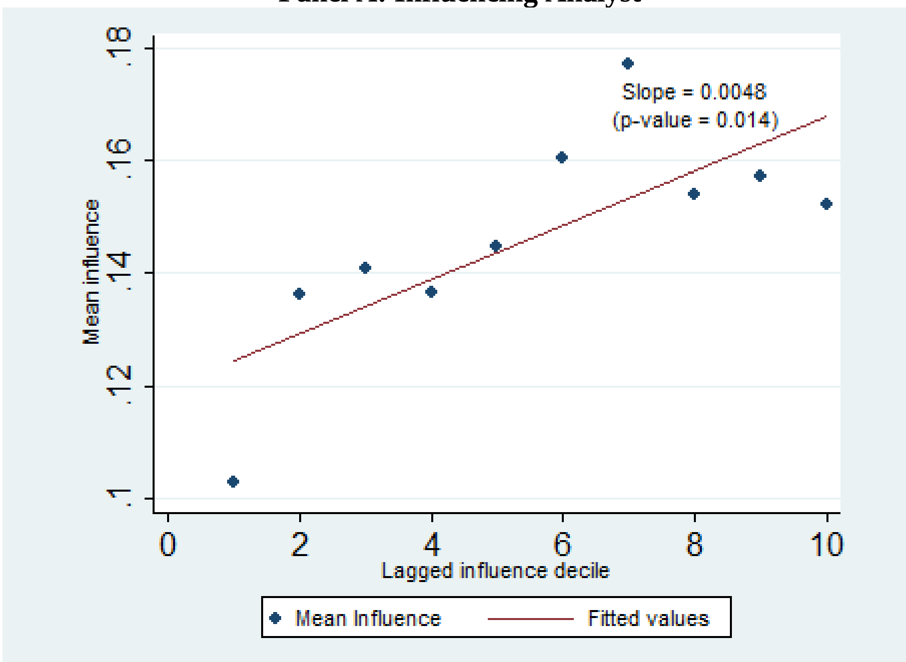


Figure 4:
Persistence in Influence

This figure plots an analyst's estimated influence over a two-year period against the decile of her estimated influence over the previous two-year period (Panel A), and her estimated tendency to be influenced over a two-year period against the decile of her estimated tendency to be influenced over the previous two-year period (Panel B).

Panel A: Influencing Analyst



Panel B: Responding Analyst

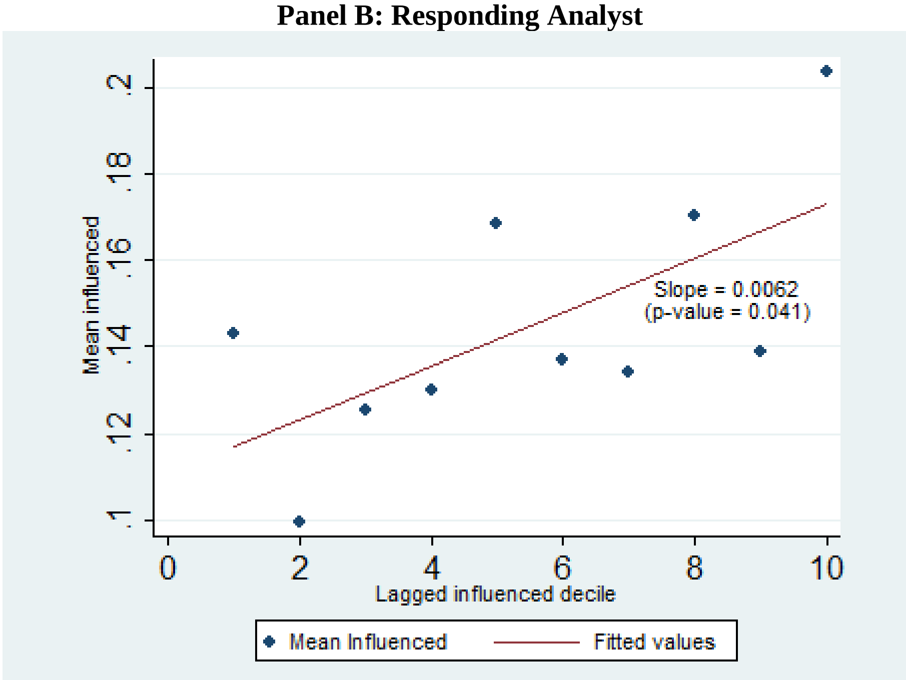


Table 1:
Summary Statistics: I/B/E/S Earnings Forecast Sample

This table provides summary data on I/B/E/S annual earnings forecasts during the sample period 1983-2009. Days to Revision is the number of days between a given analyst's forecasts, as calculated from the I/B/E/S forecast data. 3-day Ret is the 3-day market-adjusted cumulative abnormal return around a forecast, collected from CRSP, with positive revisions (increases in EPS forecasts) and negative revisions (decreases in EPS forecasts) separated. Means (medians) of these variables are reported. The table also reports the percentage of forecasts provided by CM top-ranked brokerage firms (CM Top-ranked), where a brokerage is defined as top-ranked if its Carter-Manaster ranking is greater than 9.000; by analysts ranked by the *Institutional Investor* All-American Research Team poll (II-ranked); by analysts working for the brokerage firm that took the stock public (Banker); and by analysts with more than two years of experience as provided by the I/B/E/S database at the time the forecast is made (Experienced).

	(1983-2009)
N	1,428,647
Days to Revision	97
	(70)
3-day Ret (positive revisions)	1.09%
	(0.54%)
3-day Ret (negative revisions)	-1.71%
	(-0.89%)
CM Top-ranked	26.2%
II-ranked	12.4%
Banker	3.4%
Experienced	25.9%

Table 2
Summary Statistics: Percentage of the Gap Closed

This table provides summary statistics for *GapClosed*. All annual (one-year ahead) earnings forecasts from any analyst in the I/B/E/S dataset are collected. An observation consists of adjacent forecasts by two different analysts for a given firm and fiscal year, where the first subsequent revision by the first analyst in the pair can also be identified. *Gap* is defined as $F_I - F_R$, where F_I is the influencing analyst's forecast and F_R is the responding analyst's first forecast in the sequence. *Revision* is defined as $F_R' - F_R$, where F_R' is the responding analyst's second forecast (i.e., her revision) in the sequence. *GapClosed* is defined as *Revision/Gap*, and is winsorized at the 5% and 95% levels. There is a mass in the distribution of *GapClosed* at zero, as *GapClosed* is set to zero when the responding analyst's revision does not occur before the influencing analyst's next forecast. Statistics for *GapClosed* are reported separately for cases where we do and do not include cases where *GapClosed = 0* separately.

Method	**Next Forecast**	**Zero Change**	**Omitted Observation**
Number of Forecast Pairs	247,032	247,032	169,287
Mean	0.396[a]	0.287[a]	0.438[a]
Median	0.449[a]	0.000[a]	0.500[a]
Standard Deviation	2.325	1.428	1.999
Percent > 0	62.7%[a]	44.5%[a]	65.4%[a]
Percent = 0	1.0%	32.2%	0.1%
Percent < 0	36.3%	23.4%	34.5%

a denotes significance at the 1% level.

Table 3

Analyst Forecasts, Influence, and Common Information Shocks

This table provides regression output from estimation of equations (1) and (2). The dependent variable is *GapClosed*, which is defined in Table 2. The interval Δt is defined as $t_I - t_R$, where t_I is the date of the influencing analyst's forecast in the pair and t_R is the date of the responding analyst's forecast. Δt enters linearly in columns 1, 3, and 5, while columns 2, 4, and 6 include the first three powers of Δt. In Panel A, the sample is restricted to cases where the interval of time is one day and $\Delta t \leq 60$ days, while Panel B uses time-stamped forecasts and the interval of time is one minute. Each regression includes pair fixed effects and the intercept shown is the average of the pair-specific intercepts. p-values are reported in parentheses.

Panel A: Interval = 1 day

Δt	**Next Forecast**		**Zero Change**		**Omitted Observation**	
Intercept	0.316	0.311	0.216	0.212	0.364	0.363
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Δt	0.006	0.007	0.005	0.006	0.005	0.005
	(0.00)	(0.07)	(0.00)	(0.01)	(0.00)	(0.27)
Δt^2		-0.000		-0.000		0.000
		(0.92)		(0.99)		(0.90)
Δt^3		0.000		-0.000		-0.000
		(1.00)		(0.85)		(0.85)
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.0283	0.0283	0.0397	0.0397	0.0380	0.0379
N	247,032	247,032	247,032	247,032	169,287	169,287

Panel B: Interval = 1 minute

Δt	Next Forecast		Zero Change		Omitted Observation	
Intercept	0.293	0.269	0.201	0.181	0.324	0.300
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Δt	0.040	0.419	-0.006	0.239	0.004	0.345
	(0.37)	(0.13)	(0.79)	(0.07)	(0.94)	(0.29)
Δt^2		-0.637		-0.363		-0.552
		(0.11)		(0.06)		(0.25)
Δt^3		0.250		0.130		0.211
		(0.10)		(0.06)		(0.24)
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.0281	0.0281	0.0262	0.0262	0.0319	0.0319
N	150,261	150,261	150,261	150,261	93,766	93,766

Table 4
Alternative Regression Approaches

This table provides regression output from alternative regression methods based on regression equation (2). Column 1 shows estimates from a linear probability model where the dependent variables is equal to one if *GapClosed* is positive and zero if it is negative (the observation is omitted if *GapClosed* is zero). Column 2 shows a regression where the dependent variable is the empirical percentile of *GapClosed* minus 0.4544, divided by 0.1151 to make it comparable with the results from Table 3. Column 3 shows a regression where the dependent variable is *GapClosed*, and only forecast sequences where there are no intervening forecasts between the second and third forecasts in the sequence are included. Column 4 shows a regression where the dependent variable is *GapClosed*, and forecast sequences where either analyst produces a forecast within ±5 days of an earnings announcement date are excluded. All models include analyst-pair fixed effects. p-values are reported in parentheses.

	Probability of an Upward Revision	**Percentile of GapClosed**	**No Intervening Forecasts**	**Not Around EPS Ann**
Intercept	0.589	0.327	0.261	0.307
	(0.00)	(0.00)	(0.00)	(0.00)
Δt	0.004	0.013	0.021	0.005
	(0.00)	(0.00)	(0.14)	(0.26)
Δt^2	-0.000	-0.000	-0.001	0.000
	(0.18)	(0.53)	(0.26)	(0.82)
Δt^3	0.000	0.000	0.000	-0.000
	(0.46)	(0.70)	(0.31)	(0.86)
Adjusted R^2	0.0264	0.0317	0.0124	0.0320
N	244,587	247,032	39,593	209,641

Table 5
Influence over Time

This table provides regression output from estimation of equation (1) for three different sub-periods within our full sample period: 1983-1991, 1992-2000, and 2001-2009. p-values are reported in parentheses.

	1983-1991	**1992-2000**	**2001-2009**
Intercept	0.281	0.346	0.287
	(0.00)	(0.00)	(0.00)
Δt	0.013	0.003	0.006
	(0.14)	(0.64)	(0.38)
Δt^2	-0.000	0.000	0.000
	(0.46)	(0.62)	(0.93)
Δt^3	0.000	-0.000	-0.000
	(0.58)	(0.73)	(0.80)
Adjusted R^2	0.0247	0.0366	0.0484
N	55,795	118,471	72,766

Table 6
Influence within the Fiscal Year

This table provides regression output from estimation of equation (1) for forecasts provided in the early half of the fiscal year (columns (1) and (2)) and those issued in the later half of the fiscal year (columns (3) and (4)). Similar to prior tables, regressions contain ordered pair fixed effects. p-values are reported in parentheses.

	Early Fiscal Year		**Late Fiscal Year**	
	(1)	**(2)**	**(3)**	**(4)**
Intercept	0.303	0.306	0.337	0.317
	(0.00)	(0.00)	(0.00)	(0.00)
Δt	0.006	0.004	0.006	0.010
	(0.100	(0.48)	(0.00)	(0.17)
Δt^2		0.000		-0.000
		(0.72)		(0.70)
Δt^3		-0.000		0.000
		(0.66)		(0.84)
Adjusted R^2	0.0354	0.0354	0.0413	0.0413
N	127,969	127,969	105,686	105,686

Table 7

Influence and Analyst and Firm Characteristics – Characteristic subsamples

This table provides regression output from estimation of equation (1) for different subsamples formed on the basis of analyst and firm characteristics. Each row in the table shows the values of α_i (our estimate of influence) for two subsamples formed by dividing the full sample in two based on the given characteristic, as well as the difference between the two values and the p-value of that difference. Panel A shows results for splits based on the influencing analyst's characteristics. Panel B shows the results for splits based on the responding analyst's characteristics. Analyst characteristics include indicators for whether the analyst was employed by a Carter-Manaster top-ranked brokerage firm (CM top-ranked), where top ranks are values greater than 9.000; whether the analyst was ranked by the *Institutional Investor* All-American Research Team poll in the year prior to the forecast (II-ranked); whether the analyst had an investment-banking relation with the covered firm (Banker); whether more than 70% of an analyst's forecasts exceeded consensus in the given year (Optimist); and whether the analyst has fewer than two years of experience (Inexperienced). p-values for differences are reported in column 4.

Panel A: Influencer characteristics

	Yes	**No**	**Difference**	**p-values**
CM Top-ranked	0.351[a]	0.278[a]	0.073	0.00
II-ranked	0.352[a]	0.289[a]	0.063	0.05
Banker	0.317[a]	0.298[a]	0.019	0.77
Optimist	0.223[a]	0.367[a]	-0.144	0.00
Inexperienced	0.282[a]	0.304[a]	-0.022	0.15

Panel B: Responder characteristics

	Yes	**No**	**Difference**	**p-values**
CM Top-ranked	0.269[a]	0.311[a]	-0.042	0.09
II-ranked	0.267[a]	0.305[a]	-0.038	0.23
Banker	0.314[a]	0.298[a]	0.016	0.80
Optimist	0.346[a]	0.254[a]	0.090	0.00
Inexperienced	0.287[a]	0.302[a]	-0.015	0.56

a, b, and c denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 8
Influence and Analyst and Firm Characteristics – Multivariate Analysis

This table provides regression output from estimation of regression equation (3). The characteristics used in the regression in column 1 are characteristics of influencing analyst in a given sequence (i.e., the characteristics of the analyst whose influence we are estimating). The characteristics used in the regression in column 2 are characteristics of the responding analyst in the sequence (i.e., the characteristics of the analyst whose response we are estimating). Both influencing and responding analyst characteristics are included in column 3. See Table 7 for descriptions of the characteristics. p-values are reported in parentheses.

	Influencer Characteristics	**Responder Characteristics**	**All Characteristics**
CM Top-ranked	0.075		0.077
	(0.09)		(0.08)
II Ranked	0.005		0.008
	(0.92)		(0.88)
Optimist	-0.126		-0.127
	(0.00)		(0.00)
Banker	0.029		0.030
	(0.78)		(0.77)
Inexperienced	-0.023		-0.023
	(0.62)		(0.63)
CM Top-ranked		-0.070	-0.070
		(0.23)	(0.22)
II Ranked		-0.034	-0.032
		(0.58)	(0.60)
Optimist		0.067	0.071
		(0.06)	(0.05)
Banker		0.112	0.112
		(0.31)	(0.31)
Inexperienced		0.019	0.015
		(0.69)	(0.75)
Adjusted R^2	0.0290	0.0285	0.0292
N	247,032	247,032	247,032

Table 9
Probability of Changes in Career Outcomes

This table provides logistic regressions measuring how analysts' propensity to influence or respond to other analysts affects career outcomes. Four measures of career outcomes are examined. In columns 1 and 2, an analyst's likelihood of being ranked to the *Institutional Investor* All-Star Research Team is examined. Columns 3 and 4 measure analyst promotions, while columns 5 and 6 examine demotions. Analyst terminations are examined in columns 7 and 8. Explanatory variables include first and second lags of average analyst propensity to influence, propensity to respond, and median analyst accuracy. p-values are reported in parentheses.

	Ranked		**Promoted**		**Demoted**		**Terminated**	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Intercept	-1.719	-1.708	-1.430	-1.431	-1.596	-1.595	-2.246	-2.116
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Influence_{t-1}$	0.277	0.153	0.097	0.040	0.064	0.044	-0.034	0.087
	(0.00)	(0.11)	(0.09)	(0.66)	(0.30)	(0.65)	(0.65)	(0.47)
$Influence_{t-2}$		0.265		0.083		-0.015		-0.239
		(0.01)		(0.37)		(0.88)		(0.05)
$Respond_{t-1}$	-0.038	-0.194	-0.019	-0.088	0.072	-0.041	0.065	0.261
	(0.54)	(0.03)	(0.75)	(0.35)	(0.25)	(0.68)	(0.41)	(0.03)
$Respond_{t-2}$		0.177		0.085		0.151		-0.271
		(0.07)		(0.37)		(0.14)		(0.03)
$Accuracy_{t-1}$	0.819	-0.097	2.384	1.921	-2.601	-0.319	7.758	11.258
	(0.61)	(0.96)	(0.10)	(0.35)	(0.17)	(0.90)	(0.00)	(0.00)
$Accuracy_{t-2}$		4.390		3.312		-2.125		-0.339
		(0.03)		(0.11)		(0.43)		(0.89)
Pseudo-R^2	0.0018	0.0044	0.0004	0.0011	0.0004	0.0005	0.0028	0.0065
Observations	13,300	9,854	12,609	9,258	12,609	9,258	12,509	9,206