Do Financial Analysts Herd?

May 9, 2023

Abstract

Financial analysts may have strategic incentives to herd or to anti-herd when issuing forecasts of firms' earnings. This paper develops and implements a new test to examine whether such incentives exist and to identify the form of strategic behavior. We use the equilibrium property of the finite-player forecasting game of Kim and Shim (2019) that forecast dispersion decreases (resp. increases) as the number of forecasters increases if and only if there is strategic complementarity (resp. substitutability) in their forecasts. Using the I/B/E/S database, we find strong evidence that supports strategic herding incentive of financial analysts through a plausible natural experiment setting of brokerage house mergers. We show further that this finding is robust to different forecast horizons and is more pronounced for firms with low initial coverage.

JEL classification: D83, E37, G17

Keywords: financial analysts, earnings forecasting, finite-player forecasting game, strategic interaction, herding

1 Introduction

Financial analysts aim to accurately estimate a company's earnings on a stock over the coming years. Holding the accuracy fixed, analysts might prefer to herd toward the consensus forecast—the average of all the forecasts from analysts tracking a particular stock—to avoid a reputation loss when wrong. In some other cases, analysts might prefer to deviate from the consensus to stand out and appear talented.[1] Regardless of the underlying reason, analysts may have extraneous strategic incentives to herd or to anti-herd. On the contrary, it may be that analysts only care about the accuracy of their own forecasts without any strategic considerations.

This paper proposes a new approach to test for non-information-driven strategic behavior of financial analysts and, if such strategic incentives exist, to identify the form of their forecasting behavior. In doing so, we first consider a model of finite-player forecasting game (Kim and Shim, 2019).[2] We then use the results of our model to empirically investigate whether analysts exhibit herding, anti-herding, or non-strategic behavior in issuing earnings forecasts.

In our forecasting game, each analyst receives private and public signals about a firm's expected earnings, and chooses an optimal forecast. Each analyst cares both about being correct and about his distance to the average forecast.[3] The payoff structure allows for strategic complementarity, substitutability, or independence in forecasts, which respectively represents the analysts' intrinsic preference for herding, anti-herding, or non-strategic behavior. Importantly, the finiteness of the number of analysts adds another strategic consideration when analysts are issuing forecasts. That is, with a finite number of analysts, each analyst's forecast exerts a non-negligible effect on the average forecast in comparison to a large (competitive) forecasting game that is extensively considered in the literature.

In our analysis, this finiteness of the number of analysts is the key mechanism for identifying the nature of strategic behavior. In the model, the analysts' underlying preference for herding/anti-herding uniquely pins down the relationship between the number of analysts and the forecast dispersion, measured by the variation in the equilibrium forecasts across analysts. Specifically, as the number of analysts

[1] Croushore (1997) points out that professional forecasters may herd to avoid unfavorable publicity when wrong, while others might make bold forecasts to stand out. Ottaviani and Sørensen (2006) show that in the context of a winner-take-all contest, forecasters have incentives to differentiate their predictions from those of others. Also see Prendergast and Stole (1996) who show that agents without an established reputation exaggerate their differences with others to appear talented.

[2] Our game is a version of an aggregate game whose general definition with a linear aggregate is provided by Martimort and Stole (2012).

[3] For ease of exposition, we use male pronouns for the analyst.

increases, the forecast dispersion decreases (resp. increases) if and only if the analysts' intrinsic desire is to herd (resp. anti-herd).

The intuition is as follows. As the number of analysts increases, any analyst's forecast, thus his private information, has less of an influence on the average forecast; so all analysts strategically put less weight on private information and more weight on public information (which is a relatively better predictor of the average forecast) if the analysts tend to herd, generating a lower disagreement among analysts. The opposite is true if the analysts tend to anti-herd, and if there is no such strategic incentive, then the forecast dispersion is not affected by the number of analysts. Thus, by examining the relationship between the number of analysts and dispersion in the data, we can infer the analysts' (anti-)herding incentive that is not directly observed.

We then implement an empirical analysis to test predictions of our model. The most simplest way is to regress a measure of forecast dispersion on the number of analysts. However, interpretation of the estimated coefficients from the regression is not straightforward due to concerns of endogeneity: Some omitted factors may drive both coverage and forecast dispersion. To circumvent this issue, we follow Hong and Kacperczyk (2010) and exploit an exogenous decrease in analyst coverage caused by brokerage house mergers to build a plausible causal relation with forecast dispersion. From a brokerage house merger, at least one analyst will be redundant if both houses have an analyst covering a specific stock. This implies that any change in coverage should be merger-related not due to other firm characteristics. We employ 15 brokerage merger house events from 1984 to 2005. A total of 1299 firms in our sample are identified as treated firms that are covered by both houses prior to merger event.

In this setup, we evaluate a difference-in-difference estimator to gauge the change in dispersion of treated firms relative to a benchmark group of similar characteristics. Overall, we find an increase in dispersion in response to an exogenous loss in coverage due to brokerage house mergers. For example, our baseline estimation of one-year ahead forecast shows an 8% more increase in dispersion of treated firms compared to control groups after merger events. A similar increasing pattern is observed for other forecasting horizons as well. Moreover, the effect is more pronounced for treated firms with low level of coverage prior to merger. We perform a battery of tests to confirm the validity of natural experiment setting and robustness of our results. Taken together, we uncover the hidden strategic incentive of herding of financial analysts from the data.

Financial analysts' (anti-)herding behavior has been extensively explored especially empirically, which offers mixed results.[4] For example, Hong, Kubik, and Solomon (2000); Gallo, Granger, and Jeon (2002); Lamont (2002); and Clements and Tse (2005) document evidence of herding behavior consistent with the framework of Scharfstein and Stein (1990); Trueman (1994); and others. On the contrary, Zitzewitz (2001); Bernhardt, Campello, and Kutsoati (2006); and Chen and Jiang (2006) conclude that analysts tend to anti-herd. However, as Welch (2000) has pointed out, these papers do not adequately distinguish between the strategic incentive and information channel in testing the hypothesis, which might potentially lead to unreliable estimates. Our work hence adds to the literature by resolving the issue emphasized by Welch (2000).

Key contributions of our paper to the literature can be summarized as follows. First, to our best knowledge, we are the first to jointly consider both incomplete information and strategic incentives in a unified framework to derive testable implications on herding motives. Most of the previous papers that test the herding hypothesis are classified into three groups. The first group is the pure empirical works (among many others, see Hong, Kubik, and Solomon, 2000; Hong and Kubik, 2003; Hong and Kacperczyk, 2010; and Clements, 1999). The second group is based on micro-founded predictions that considered models with only incomplete information, not strategic considerations (for example, see Trueman, 1994; Jegadeesh and Kim, 2010; and Clements, 2018). The last group is the pure theoretical paper, which does not provide an empirical test (Ottaviani and Sørensen, 2006). However, there has been no attempt to explicitly consider the strategic motives, which is also an important feature in the market for analysts.

Second, we are the first to exploit the finite property of an aggregate game to obtain a novel identification strategy to draw inferences about analysts' tendencies to herd. While there have been a few of papers, Hong and Kacperczyk (2010) for example, that exploited the exogenous changes in numbers of analysts for obtaining validity of the empirical analysis, our work is different from them by providing a structural interpretation on the link between numbers of forecasters and forecast dispersion. Lastly, to our best knowledge, this is the first study that exploits a natural experimental setting to empirically investigate how forecast dispersion would be affected by analyst coverage. Even though the relation between coverage and dispersion has been documented in previous literature, most studies hinge on a simple regression framework without accounting for potential endogeneity bias.[5] In this

[4]Clements (2018) provides a nice summary of the recent literature on macro forecasting that supports either herding or anti-herding.

[5]See, for example, Diether, Malloy, and Scherbina (2002) and Pouget, Sauvagnat, and Villeneuve (2017) though the

paper, we employ a natural experimental design which allows us to establish a direct causal relationship and demonstrate that an exogenous drop in analyst coverage yields an increase in forecast dispersion among analysts.

The remainder of the paper is organized as follows. Section 2 reviews a theoretical framework that provides testable implications for the data. Section 3 describes our data, presents our empirical findings, and considers their robustness. Section 4 concludes the paper.

2 The model and predictions

In this section, we review briefly the finite-player forecasting game of Kim and Shim (2019) but in the specific context of earnings forecasting by financial analysts, and use its equilibrium properties to derive testable predictions for the form of strategic interaction (if any) among analysts.

2.1 Finite-player forecasting game Consider a simple economy in which there are one risky firm and n financial analysts, each of whom is indexed by i and issues a forecast of the firm's earnings, $\theta \in \mathbb{R}$. We assume that nature draws θ from an improper uniform distribution over the real line. Agents receive noisy signals that are informative about the firm's earnings. That is, each agent i observes a public signal $p = \theta + (\alpha_p)^{-1/2}\varepsilon$ and a private signal $x_i = \theta + (\alpha_x)^{-1/2}\varepsilon_i$. The ε and ε_i are, respectively, common and idiosyncratic noises that are independent of each other as well as of θ, and both follow $N(0, 1)$. We let α_p and α_x denote the precision of public and private signals, respectively. Private signal is the information that a single forecaster knows about the firm's earnings, which might include his own expectation on the profitability of the firm based on his own techniques or inside information so that other forecasters cannot know. On the contrary, public signal includes the information that is publicly disclosed by the firm so that it is a common knowledge across the analysts. It might include capital structure, ownership, and its past earnings/market share etc.

After observing his signals, each analyst i releases a forecast of θ, which we denote as $a_i \in \mathbb{R}$, and receives a payoff u_i, which is given by $u_i(a_i, A_n, \theta) = -\frac{1}{2}\left((1-r)(a_i-\theta) + r(a_i - A_n)\right)^2$ or, equivalently,

$$u_i(a_i, A_n, \theta) = -\frac{1}{2}\left(a_i - (1-r)\theta - rA_n\right)^2, \tag{2.1}$$

main focus is not on the relation between coverage and dispersion.

where $A_n \equiv \frac{1}{n}\sum_{i=1}^{n} a_i$ denotes the average forecast across the population and the parameter $r \in (-1, 1)$ gives the weight that the analyst puts on the average forecast relative to the fundamentals.[6]

While the payoff specification is quite stylized, it is general enough to encompass a variety of situations.[7] When $r = 0$, each analyst cares only about being correct, generating a fundamental motive to be close to the true θ; so there is no strategic interaction across analysts. When $r \neq 0$, each analyst cares both about being correct and about the distance of his forecast to the average forecast A_n, which entails two channels of strategic motives. The first motive, which we call the *strategic motive*, arises from the analysts' intrinsic preferences for (anti-)herding—i.e, whether analysts' forecasts are strategic complements ($r > 0$, herding) or strategic substitutes ($r < 0$, anti-herding). The second motive, which we call the *market-power motive*, arises from the analysts' ability to strategically influence the average forecast by changing his forecast, due to the finiteness of the number of analysts ($n < \infty$).

In this game, the equilibrium forecast of agent i is uniquely characterized as follows.[8]

$$a_i(x_i, p) = \lambda_n x_i + (1 - \lambda_n)p, \ \forall i \in \{1, \ldots, n\}, \tag{2.2}$$

where $\lambda_n = \frac{\alpha_x}{\alpha_x + \frac{1}{1-\gamma}\alpha_p}$ and $\gamma \equiv \frac{r(n-1)}{n-r}$. The coefficient λ_n measures how the agents allocate their use of private information relative to public information in equilibrium. This equilibrium weight λ_n reflects a combination of both the strategic and market-power motives, the degrees of which are together captured by the parameter γ.

Lemma 2 of Kim and Shim (2019) establishes the following result, which we restate for the convenience of the reader.

Result 1. *For any given α_x and α_p and for any given n such that $2 \leq n < \infty$, $\frac{\partial \lambda_n}{\partial n} < 0$ when $r > 0$, $\frac{\partial \lambda_n}{\partial n} = 0$ when $r = 0$, and $\frac{\partial \lambda_n}{\partial n} > 0$ when $r < 0$.*

Proof. The proof is immediate: $\frac{\partial \lambda_n}{\partial n} = -r\frac{\alpha_x \alpha_p}{n^2(1-r)}\left(\alpha_x + \frac{n-r}{n(1-r)}\alpha_p\right)^{-2} \lesseqgtr 0$ iff $r \gtreqless 0$ □

That is, as the number of analysts increases, the analysts put less (resp. more) weight on private information when their forecasts are strategic complements (resp. strategic substitutes). The intuition

[6] For tractability, we assume that analyst's preferences are quadratic to ensure linearity in the best responses. The equilibrium is unique if and only if $r < 1$.

[7] The forecasting game described here is an example of an aggregate game in which each agent's payoff is a function of his own strategy and some aggregator of the strategy profile of all agents. Aggregate games are studied in Acemoglu and Jensen (2013), Cornes and Hartley (2012), Martimort and Stole (2012), among many others.

[8] The detailed proof can be found in Kim and Shim (2019).

comes from the fact that with a finite number of analysts the average forecast of the population contains the analysts' private noises, which disappear as n goes to infinity. Accordingly, as more analysts participate in issuing forecasts, any analyst's private information has less of an influence on the average forecast; so all analysts strategically put less weight on private information when their intrinsic desire is to herd ($r > 0$), whereas the opposite happens when the analysts' intrinsic desire is to be distinctive from the herd ($r < 0$). When agents do not care about what others do ($r = 0$), then the number of analysts has no effect on λ_n.

The equilibrium forecast in equation (2.2) can be rewritten as $a_i = \theta + \lambda_n(\alpha_x)^{-1/2}\varepsilon_i + (1 - \lambda_n)(\alpha_p)^{-1/2}\varepsilon$. Then the equilibrium level of forecast dispersion for any given realizations of θ and p is given by

$$Var(a_i|\theta, p) = \left(\lambda_n \, (\alpha_x)^{-1/2}\right)^2. \tag{2.3}$$

This measure of forecast dispersion depends directly on the weight λ_n, which is defined in terms of r and n in addition to signal precisions.

2.2 Discussion of the model First of all, the study of a finite-player model is pertinent due to the following reason. The preference parameter r that measures the underlying behavior of analysts and the weight λ_n that measures the allocation of information signals are generally not observable to researchers. The model with a finite number of agents enables us to explore the relationship between n and $Var(a_i|\theta, p)$, which can be observed in the data. We can then infer from the data whether analysts exhibit herding or anti-herding behavior by estimating empirical patterns of forecast dispersion in relation to the number of analysts.

Second, our model assumes that all analysts release their forecasts simultaneously. One might consider a situation in which analysts provide their forecasts sequentially. If the analysts have the flexibility to optimally choose when to disclose their forecasts, any analyst might have an incentive to delay his announcement so that he can have access to more information and condition his forecast on any previously released forecast. Hence, if all analysts are symmetric in terms of preferences, all forecasts will be issued at the same time in equilibrium. As a result, the equilibrium in the case of sequential forecasting would be substantively equivalent to the equilibrium of simultaneous forecasting.[9]

[9] Trueman (1994) analyzes the case where the order in which the analysts disclose their forecasts is determined exogenously. In such case, the paper finds that analysts tend to behave according to their non-information related incentive to herd.

Lastly, while we focus on the static model, one natural extension is to consider a dynamic model. For example, we may assume that the fundamental variable θ_t follows AR (1) process and analysts observe noisy private and public signals in each period together with θ_{t-1}. Under some conditions, we can show that the analysis of the static model is exactly preserved in this dynamic version of forecasting game. In particular, the expression of forecast dispersion that is essentially equivalent to equation (2.3) can also be derived for the dynamic model, thus we focus on the static model for simplicity of analysis.

2.3 Testable implications

To derive testable implications about strategic interaction in analysts' forecasts, we focus on how the dispersion of forecasts in equation (2.3) changes in response to a change in the number of analysts issuing those forecasts. The following predictions lay the basis for our empirical tests in Section 3.

Prediction. *Suppose that the degree of the strategic motive, r, does not depend on the number of analysts issuing forecasts, and that r is the same across all analysts and across different forecast horizons. For any given value of α_x and α_p, as the number of analysts increases, the following results hold:*

1. *The forecast dispersion decreases iff $r > 0$, does not change iff $r = 0$, and increases iff $r < 0$.*
2. *The above relationship is preserved across different forecast horizons.*
3. *The magnitude of the effect of an additional analyst on the forecast dispersion becomes smaller if $r \neq 0$, whereas there is no such size effect if $r = 0$.*

Proof. Prediction 1: An observation of equation (2.3) yields $\frac{\partial Var(a_i|\theta,p)}{\partial n} \propto \frac{\partial \lambda_n}{\partial n}$. Then the proof follows from Result 1. Prediction 2 is a direct implication of Prediction 1. Prediction 3: Following from the proof of Result 1, $\partial\left(|\frac{\partial \lambda_n}{\partial n}|\right)/\partial n < 0$ when $r \neq 0$, and is zero otherwise. □

The intuition behind Prediction 1 is as follows. Public information is a relatively better predictor of the average forecast than private information. While any analyst's forecast, thus his private information, exerts a non-negligible effect on the average forecast, it becomes less influential as the number of analysts increases. So as n increases, the analysts whose preference is for herding ($r > 0$) rely less on private information, generating a lower disagreement among analysts. On the other hand, when the analysts want to deviate from the herd ($r < 0$), they find it optimal to use more private information, which leads to a higher disagreement among analysts. Finally when the agents do not care about the herd

($r = 0$), there is also no finite-player strategic consideration in place, and so the forecast dispersion is independent from the number of analysts.

Prediction 1 provides the key channel for identifying the underlying (anti-)herding behavior of financial analysts. We can exploit the relationship between the number of analysts issuing earnings forecasts of a firm and the forecast dispersion observed in the data to infer such strategic interaction, if there is any.

How about Prediction 2? Financial analysts issue earnings forecasts of companies at different forecast horizons. Intuitively, as is noted by Patton and Timmermann (2010), it is more difficult to forecast long-run earnings than short- or medium-run earnings and differences among analysts' information signals tend to matter more at short forecast horizons where signals are stronger. We can capture this feature, a varying length of the forecast horizon, by changing precisions of signals in our model. Hence, given the assumption that the underlying degree of the strategic motive, r, does not depend on the forecast horizon, varying forecast horizons should not change Prediction 1.

The last prediction, Prediction 3, was not introduced in Kim and Shim (2019) and hence is unique to this paper. This prediction arises from the feature of our model in which there are two strategic effects, one due to the finiteness of the number of analysts and the other due to the analysts' preference for (anti-)herding. If those two forces are at play for financial analysts, then the marginal effect of an additional analyst on forecast dispersion should be larger when fewer analysts are issuing forecasts. Again, this would not be observed if there is indeed no strategic motive ($r = 0$): As analysts only care about being correct, addition or removal of one additional analyst should not be associated with the degree of the effect of changes in numbers of analysts on the forecast dispersion.

3 Empirical analysis

In this section, we first introduce our identification strategy of empirical analysis. We then describe our data and sample. Lastly, we provide estimation results.

3.1 Identification strategy

In the previous section, we derive testable hypotheses on financial analysts' herding behavior from simple finite-player forecasting game framework. To formally test these predictions, one may simply regress a measure of forecast dispersion on the number of analysts. However, the estimated coefficients from OLS regression is likely to be biased due to potential endogeneity concerns of analyst coverage. For example, the negative relation between analyst coverage and forecast dispersion may reflect the fact that analysts are attracted to firms with less uncertain earnings, which may suggest a possibility of reverse causality. Therefore, to build a direct causal relation between the two, we need to exploit a proper event that exogenously changes the analyst coverage while not affecting firms' earnings.

To circumvent this issue, we follow Hong and Kacperczyk (2010) and use mergers of brokerage houses as exogenous source of variation in analyst coverage.[10] Because brokerage houses engaging in merger events fire redundant analysts, a drop in analyst coverage is expected while it is unlikely to affect firms' business or future earnings. As the termination of coverage is not determined by the analyst, this setting properly isolates any effect from the drop in analyst coverage. Hong and Kacperczyk (2010) evaluate this natural experiment setting to find a discernible decrease in analyst coverage which results in an increase in optimism bias after the merger events. To be consistent with the herding hypothesis, we should observe an increase in forecast dispersion after brokerage merger events.

Table 3.1 shows the list and dates of 15 merger events identified from Hong and Kacperczyk (2010). They identify mergers among brokerage houses from the SDC Mergers and Acquisition database.[11] In each merger event, the names of bidder (top row) and target (bottom row) brokerage houses are indicated. For example, merger number 4 shows the merger between Morgan Stanley and Dean Witter Reynolds which occurred on May 31, 1997.

To empirically test our hypothesis of forecast dispersion, we implement a difference-in-difference strategy around the merger dates. This method estimates the difference in the variable of interest across the event window between the treated and control groups. The treatment group consists of stocks that were covered by both brokerage houses (bidder and target) before the merger, and the control group includes all the remaining stocks. To mitigate the concern that the results can be partially driven by differences in firm characteristics, we follow Hong and Kacperczyk (2010) that each stock in the treatment group is matched with its own benchmark portfolio obtained using the stock in the control group. From this matching procedure, we can effectively eliminate potential heterogeneity between two groups, thereby focusing on the true causal effect of brokerage mergers on forecast dispersion.

The benchmark portfolio is constructed using firm size, book-to-market ratio, return momentum, and analyst coverage. We first sort stocks into three portfolios according to their market capitalizations.

[10]See also Kelly and Ljungqvist (2012), Derrien and Kecskés (2013), Irani and Oesch (2013), and Chen, Kelly, and Wu (2020) for application of brokerage mergers.

[11]See Hong and Kacperczyk (2010) for further details.

Table 3.1: List of brokerage merger event.

Number	Brokerage house	Merger date
1	Merrill Lynch Becker Paribas	9/10/1984
2	Wheat First Securities Butcher & Co., Inc.	10/31/1988
3	Paine Webber Kidder Peabody	12/31/1994
4	Morgan Stanley Dean Witter Reynolds	05/31/1997
5	Smith Barney (Travelers) Salomon Brothers	11/28/1997
6	EVEREN Capital Principal Financial Securities	1/9/1998
7	DA Davidson & Co. Jensen Securities	2/17/1998
8	Dain Rauscher Wessels Arnold & Henderson	4/6/1998
9	First Union EVEREN Capital	10/1/1999
10	Paine Webber JC Bradford	6/12/2000
11	Credit Suisse First Boston Donaldson Lufkin Jenrette	10/15/2000
12	UBS Warburg Dillon Read Paine Webber	12/10/2000
13	Chase Manhattan JP Morgan	12/31/2000
14	Fahnestock Josephthal Lyon & Ross	9/18/2001
15	Janney Montgomery Scott Parker/Hunter	3/22/2005

Notes: This table reports the list of brokerage merger event of Hong and Kacperczyk (2010). The names and dates of the merging brokerage houses are included. For each merger event, the brokerage house in the top row is the acquiring house and the brokerage house in the bottom row is the target house.

Next, firms within each size tercile are grouped into three book-to-market portfolios. We further sort stocks in each of the portfolios sorted on size and book-to-market into tercile portfolios according to their past returns. Lastly, we sort stocks in each of the 27 portfolios into tercile portfolios according to their analyst coverage. As a result, we obtain 81 benchmark portfolios. Therefore, the matched portfolios have similar features with treated stocks in terms of the four dimensions. Using benchmark portfolios as a control group, the partial effect of a change in dispersion due to merger is calculated as follows,

$$DID^i = (Disp^i_{T,Post} - Disp^i_{T,Pre}) - (Disp^i_{C,Post} - Disp^i_{C,Pre}), \tag{3.1}$$

where the first term in parenthesis is the difference in forecast dispersion of firm i in the treatment sample before and after the merger event and the second term is the change in the average dispersion of the benchmark portfolios that are matched to firm i during this period. This requires us to set an estimation window around the merger events. In choosing a proper window to obtain forecast dispersion in pre- and post-event period, we use a two-year window, with one year of data selected for each period since most analysts typically issue at least one forecast within a twelve-month window. Finally, we take the average of difference-in-difference estimators (DID) across firms and merger events to gauge the average effect.

3.2 Data and sample We draw financial analysts' earnings forecasts data from the Thomson Reuters' Institutional Brokers Estimate System (I/B/E/S) database. The database provides analysts' historical earnings estimates for more than 20 forecast measures, including earnings per share. In particular, we utilize the I/B/E/S Detail History Unadjusted file.[12] We extract firm-level data from the Center for Research on Security Prices (CRSP) files and the Compustat database. Then, we merge these database to construct our sample. The sample firms are basically all public firms listed on the stock market.

3.3 Measure of analyst forecast dispersion Our empirical proxy of analyst forecast dispersion is constructed following Diether, Malloy, and Scherbina (2002), which is defined as the standard deviation scaled by the absolute value of mean of current-fiscal-year earnings estimates across analysts. Since our difference-in-difference estimation exploits a twelve-month window for each pre- and post-merger

[12] We use the unadjusted data to avoid the rounding error reported in Diether, Malloy, and Scherbina (2002). We use the cumulative adjustment factors from the CRSP to adjust the forecast for stock splits.

period, we may observe multiple forecast announcements for each analyst during the event window. Therefore, we keep only the observation that is closest from the merger date. As a result, there are two observations remaining for each analyst (one for each period) covering a stock.

We also consider other firm characteristics that may differ across treated and control groups. We consider the firm size (Size), the book-to-market ratio (BM), the past twelve month stock return momentum (Momentum), the profitability (Profit), the past twelve month stock return volatility (Sigma), and the return-on-equity (ROE) volatility (ROEVol). Size is the logarithm of a firm's market capitalization calculated as the number of shares outstanding times stock price. BM is defined following Fama and French (1993). Momentum is the past one year cumulative stock returns, and Profit is defined as operating income over book value of assets. Sigma is the variance of daily returns of a stock over past one year, and ROEVol is the past ten year annual ROEs where ROE is firm's return on equity measured as the ratio of earnings to the book value of equity.

In Table 3.2, we document descriptive statistics for the treated firms and all other firms. To mitigate the effect of extreme values, we winsorize all variables except coverage at the 1% and 99% levels. On average, we see that treated firms are larger and covered by more analysts than other firms in the sample. Since the treated firms are involved with large brokerage house mergers and large houses tend to cover large firms, it is natural to observe a notable difference in size. The treated firms also show smaller dispersion in analyst forecasts. This negative association between coverage and dispersion between two groups implies the analysts' herding. Table 3.2 also shows that treated firms have lower book-to-market ratio, larger return momentum, lower volatility measures (both return and ROE), and larger profitability.

As shown in Table 3.2, we see a notable difference in the variation of analyst coverage between treated and control firms. However, this difference is mitigated through our matching process described in Section 3.1. To begin with, we first report the range of cutoff values for analyst coverage in each size tercile instead of reporting all 54 cutoff values as coverage is generally correlated with firm size. For the first size tercile, the first cutoff value ranges from 2 to 3, and the second cutoff value ranges from 3 to 5. For the second (third) size tercile, the first cutoff value ranges from 3 to 7 (6 to 18) while the second cutoff value ranges from 5 to 10 (11 to 26). As treated firms are generally larger and covered by more analysts on average, these firms are mostly sorted into the third size tercile. Then, we investigate the average difference of coverage between treated and matched control firms in each size tercile. It turns

Table 3.2: Descriptive statistics.

Group	Variable	N	Mean	Std. dev.	Q1	Median	Q3
Treated firms	Coverage	1299	24.731	10.785	17	23	32
	Dispersion	1299	0.229	0.667	0.047	0.082	0.171
	Size	1299	8.281	1.449	7.220	8.319	9.294
	BM	1299	0.505	0.380	0.217	0.404	0.710
	Momentum	1299	0.335	0.786	-0.055	0.176	0.432
	Profit	1299	0.373	0.291	0.237	0.334	0.450
	Sigma	1299	0.104	0.064	0.060	0.088	0.128
	ROEVol	1299	0.398	1.544	0.054	0.095	0.184
Control firms	Coverage	29898	9.581	8.089	4	7	13
	Dispersion	29898	0.376	0.957	0.050	0.107	0.257
	Size	29898	6.111	1.655	4.901	5.968	7.154
	BM	29898	0.584	0.462	0.262	0.468	0.776
	Momentum	29898	0.309	0.814	-0.146	0.133	0.485
	Profit	29898	0.275	0.346	0.170	0.280	0.393
	Sigma	29898	0.140	0.078	0.085	0.123	0.175
	ROEVol	29898	0.633	1.838	0.070	0.145	0.369

Notes: This table summary statistics for treated and control firms. The treated firms are all stocks covered by two merging houses around the event date. The control firms are all other firms remaining in the database. Coverage denotes the number of analysts covering a stock. Dispersion is defined as the standard deviation scaled by the absolute value of mean of current-fiscal-year earnings estimates across analysts. Size is the logarithm of a firm's market capitalization calculated as the number of shares outstanding times stock price. BM is defined following Fama and French (1993). Momentum is the past one year cumulative stock returns, and Profit is defined as operating income over book value of assets. Sigma is the variance of daily returns of a stock over past one year, and ROEVol is the past ten year annual ROEs where ROE is firm's return on equity measured as the ratio of earnings to the book value of equity.

out that the average differences are in a reasonable range from 0.6 to 1.2, implying that treated and matched control groups are generally similar in terms of coverage.

3.4 Main results Prediction 1 implies that we should observe a negative (resp. positive) relationship between the number of analysts and the forecast dispersion if the analysts' underlying strategic behavior is herding (resp. anti-herding) in the data. Since our setting exploits an exogenous drop in coverage, treated firms should exhibit an increase (resp. decrease) in forecast dispersion if analysts herd (resp. anti-herd).

Table 3.3 presents the average difference-in-difference estimator for forecast dispersion. In Panel A, we begin by investigating the change in analyst coverage around the event window to check the validity of natural experiment setting. We find that there is a significant drop in coverage for treated firms compared to benchmark group. For example in the baseline case (column (1)), the estimated coefficient is -0.444 which is highly significant with t-statistic -3.23.[13] The estimate -0.444 is relatively small compared to the number reported in Hong and Kacperczyk (2010). This is partially due to our sample restriction that mechanically drop stocks covered by only one analyst.

More importantly, we find that earnings forecast has become more dispersed as a consequence of merger between brokerage houses. In Panel B, for example of baseline case, we find an increase in dispersion for treated firms by 0.019 which is significant at 1% level. The magnitude is also economically meaningful considering the average dispersion (0.229) of treated firms before merger period. In other words, the effect of brokerage mergers solely contributes to an 8% increase in forecast dispersion. Overall, there is an increase in forecast dispersion in response to exogenous drop in coverage. The result is consistent with herding behavior of analysts.[14]

One natural critique on our argument that it is the herding motives ($r > 0$) behind the negative relationship between the numbers of forecasts and forecast dispersion is that there can be potentially other channels than the herding (or non-herding) motives that can explain our main empirical findings. To consider this possibility, assume that there is no strategic motive ($r = 0$). From equation (2.1) and (2.2), one can easily observe that the optimal forecast would be a function of precision of signals and the realization of the signals. This implies that if there is no strategic consideration, then numbers of

[13]The standard errors are clustered at the merger groupings since the error can be correlated within each merger date.

[14]We find the results robust to matching procedures using quartile or quintile portfolios for coverage. The results are available upon request.

analysts cannot affect the forecast dispersion. This is a natural consequence since the objective of the analyst is to provide the best forecast regardless of accuracy of other analysts' forecast.

Now consider alternative channels: If there is a 'competition'[15] across the analysts, greater numbers of forecasters would affect the analyst to exert more effort to make more accurate forecast, and hence the forecast dispersion would become lower. While this hypothesis is consistent with our finding, the channel through which the analyst exert more effort to receive better private signals (endogenous information acquisition) requires strategic motives in the utility function: If $r = 0$, then there is no way for the changes in competition to affect the endogenous information acquisition. We need a motivation for the agent to change his information as a function of numbers of forecasts and assuming $r = 0$ makes the analyst not to respond to changes in numbers of forecasts.

Table 3.3: Change in stock-level analyst coverage and forecast dispersion: DID estimator.

Variable	(1) Baseline (1 year forecast)	(2) Short-term (1 quarter forecast)	(3) Long-term (2 years forecast)
Panel A: Change in analyst coverage			
Coverage	-0.444	-0.423	-0.531
	(-3.23)	(-3.53)	(-3.80)
Panel B: Change in forecast dispersion			
Dispersion	0.019	0.026	0.024
	(3.19)	(3.04)	(2.29)
Panel C: Change in forecast dispersion: Conditional on pre-merger coverage level			
Dispersion (Low)	0.076	0.041	0.071
	(2.67)	(2.04)	(2.08)
Dispersion (Med)	0.023	0.026	0.030
	(2.53)	(2.63)	(2.99)
Dispersion (High)	0.010	-0.016	0.001
	(0.79)	(-0.46)	(0.08)

Notes: This table shows the result of difference-in-difference test described in equation 3.1. We report the average change in coverage (Panel A) and dispersion (Panel B) around mergers. In Panel C, we classify treated firms into three groups at each merger depending on a firm's pre-merger coverage level and report the average change in dispersion of three groups (denoted in parenthesis). Low includes sample covered by less than or equal to 10 analysts, Med includes sample covered by more than 10 and less than or equal to 20 analysts, and High includes sample covered by more than 20 analysts. We consider three forecasting horizons (1 year ahead forecast in column (1), 1 quarter ahead forecast in column (2), and 2 years ahead forecast in column (3)). The t-statistics in parentheses are robust to clustering at the merger groupings.

We next test Prediction 2. The I/B/E/S data also contain various earnings estimates in terms of different forecast horizons, from current-fiscal-quarter to two-years-ahead.[16] We repeat the difference-in-difference using the estimates of different forecast horizons, and the findings are reported in columns (2) and (3) of Table 3.3. Overall, we find a similar increase in dispersion for both short-term estimates

[15]This can also be a form of reward for a better forecast or penalty for a worse forecast.

[16]Following van Binsbergen, Han, and Lopez-Lira (2020), we do not consider estimates beyond two years because analysts' forecasts for longer horizons have significantly fewer observations.

(0.026) and two-years-ahead estimates (0.024) which are also highly significant at conventional levels. In sum, the analysts' strategic motive is not altered by varying forecast horizons.

In Panel C of Table 3.3, we test the Prediction 3. Our hypothesis indicates that the effect should be more pronounced at low coverage level. To investigate, we divide treated firms into three groups at each merger depending on a firm's pre-merger coverage level similar to Hong and Kacperczyk (2010). Specifically, *Low* group contains observations covered by less than or equal to 10 analysts, *Med* group contains those covered by more than 10 and less than or equal to 20 analysts, and *High* group contains more than 20 analysts. We then estimate the change in forecast dispersion around merger event and report the average change in dispersion of three groups. Consistent with our hypothesis, the effect of exogenous drop in coverage is larger for the low coverage groups. While the estimates of *Low* and *Med* are all positive and significant at conventional levels, we do not find any significant change in dispersion for firms in *High* group. In particular, the estimated magnitude of *Low* is sizable considering the average change shown in Panel B. Overall, we find a herding behaviour of analysts that is more significant when there are fewer analysts around.

The difference-in-difference assumes a parallel trend between treated and control firms that the difference in dispersion between two groups is not an ongoing trend. To test the validity of natural experiment, we rerun our difference-in-difference test described in equation 3.1, $DID^i = (Disp^i_{T,Post} - Disp^i_{T,Pre}) - (Disp^i_{C,Post} - Disp^i_{C,Pre})$, but shifting the event window up to three years before and after the merger. We expect to observe insignificant estimates for difference-in-difference coefficients if the results documented in Table 3.3 is driven by exogenous drop in analyst coverage due to brokerage mergers.

We provide the placebo test result in Table 3.4. Panel A shows the difference-in-difference tests for baseline case of 1 year forecast. In all periods considered, we do not find any economically significant changes in forecast dispersion. We find similar patterns for short-term forecast (Panel B) and long-term forecast (Panel C). The result shows that the adjustment in analyst forecast among the treated firms takes place only around the merger event dates and is not due to some trend either in the pre- or the post-event window. This also addresses the potential concern that some latent characteristics or macro factors might drive the divergence between treated and control groups. In sum, Table 3.4 validates our difference-in-difference setting and does not necessarily imply the disappearance of analyst herding. For example, one might observe a herding behavior among analysts for certain group of stocks one year

Table 3.4: Change in forecast dispersion: Placebo test.

	(1)	(2)	(3)	(4)	(5)	(6)
	Time from event date (in years)					
Variable	-3	-2	-1	+1	+2	+3
Panel A: Baseline (1 year forecast)						
Dispersion	-0.021	-0.003	0.031	0.002	0.004	0.004
	(-1.03)	(-0.05)	(0.92)	(0.04)	(0.47)	(0.16)
Panel B: Short-term (1 quarter forecast)						
Dispersion	0.008	-0.018	0.001	0.007	0.002	-0.004
	(0.91)	(-0.47)	(0.03)	(1.12)	(0.52)	(-0.10)
Panel C: Long-term (2 years forecast)						
Dispersion	0.009	0.007	0.012	0.006	-0.022	0.018
	(0.93)	(0.68)	(1.25)	(0.32)	(-0.33)	(0.96)

Notes: This table shows the result of difference-in-difference test described in equation 3.1 with shifting event window up to three years before and after the merger. We consider three forecasting horizons (1 year ahead forecast (Panel A), 1 quarter ahead forecast (Panel B), and 2 years ahead forecast (Panel C)). The t-statistics in parentheses are robust to clustering at the merger groupings.

after the merger. The placebo test result simply tells that it is unrelated to our proposed treatment effect of brokerage mergers.

3.5 Robustness Test In this section, we implement a difference-in-difference methodology in a regression framework instead of using benchmark portfolios to check the robustness of results. Specifically, we consider the following specification,

$$Disp_i = \alpha + \beta_1 Post_i + \beta_2 Treated_i + \beta_3 Post_i \times Treated_i + \gamma X_i + \epsilon_i, \tag{3.2}$$

where *Post* denotes an indicator variable that is equal to one if the observation is in the post-merger period and zero otherwise, and *Treated* is an indicator variable that is equal to one if a firm is treated and zero otherwise. The treated stocks are all stocks that experience a sudden drop in coverage due to brokerage mergers as identified in the previous section. The coefficient of interest is β_3 which gauges the effect of the merger on forecast dispersion of treated firms compared to control firms.

The specification also requires us to include several control variables to mitigate concerns of ex-ante difference between treated and control groups. We mostly follow Hong and Kacperczyk (2010) to consider six control variables. These include the firm size, the book-to-market ratio, the past-year return momentum, the past-year return volatility, the profitability, the past ten-year ROE (return on equity) volatility. Firm size, book-to-market ratio, and momentum are firm characteristics used to construct benchmark portfolios in our difference-in-difference estimator. We conjecture that large firms tend to attract more coverage and are likely to have less forecast dispersion. We include the logarithm of book-to-market ratio following Fama and French (1993) as the high book-to-market values are likely to be risky and therefore are more dispersed in earnings forecast. Return momentum and profitability measures a firm's past financial performances. A better performance may be associated with less disagreement among analysts. Lastly, two volatility measures are used to capture the overall riskiness of a firm's assets. In addition, we consider fixed effects to further account for unobservable heterogeneity. Our preferred specification includes merger, year, and firm fixed effects that account for unobservable time-invariant factors particular to a merger, year, and firm that may influence the forecast dispersion.

In Table 3.5, we provide the estimation results of equation 3.2. Similar to our main results in Table 3.3, we start with how analyst coverage has changed during the period. In column (1), we find that treated firms lose coverage compared to control firms in the year following a brokerage merger events. Again, this corroborates our natural experiment setting.

Then, we examine the effect of coverage loss on forecast dispersion. We document the baseline results in columns (2) and (3), the short-term forecast results in columns (5) and (6), the long-term forecast results in columns (8) and (9). Columns (2), (5), and (8) show the estimates without firm-level control variables, while columns (3), (6), and (9) show the estimates with firm-level control variables. In all columns except (1), we include merger, year, and firm fixed effects. Overall, the increase in forecast disperion in response to coverage drop seems a robust feature of data. In other words, the coefficient estimates on treatment effect (Post×Treated) are all positive and significant at the 1% level regardless of forecast horizons. The magnitudes become slightly larger than estimates reported in Table 3.3.

The treatment effect is not significantly changed after firm-level variables are controlled for. Nonetheless, the coefficient estimates on control variables are noteworthy. We find that firm size has a negligible effect on forecast dispersion. The positive and significant coefficient on book-to-market ratio is consistent with our conjecture that book-to-market represents a dimension of firm riskiness. The estimates on two performance measures (Momentum and Profit) show a strong negative sign as predicted. However, we find two opposite signs for volatility measures. While the return volatility (Sigma) is positively related to dispersion, the ROE volatility has a negative effect on dispersion.[17]

[17]Even though we use a similar set of control variables, we find that the signs on controls are different to the results

Table 3.5: Change in forecast dispersion: Robustness test.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Coverage	Dispersion (1 year)			Dispersion (1 quarter)			Dispersion (2 years)		
Post	-0.288***	-0.023*	-0.023*	-0.023*	-0.038***	-0.038***	-0.038***	-0.005	-0.005	-0.005
	(-3.25)	(-1.94)	(-1.94)	(-1.94)	(-3.93)	(-3.93)	(-3.93)	(-0.79)	(-0.79)	(-0.79)
Treated	14.803***	-0.033***	-0.036***	-0.037***	-0.028***	-0.033***	-0.033***	-0.018*	-0.021**	-0.021**
	(17.10)	(-4.47)	(-4.51)	(-4.49)	(-3.51)	(-4.13)	(-4.04)	(-2.10)	(-2.18)	(-2.15)
Post x Treated	-0.813***	0.059***	0.059***		0.045***	0.045***		0.041***	0.041***	
	(-2.98)	(4.65)	(4.64)		(3.27)	(3.27)		(3.45)	(3.45)	
Post x Treated x Low				0.076**			0.062***			0.061**
				(2.59)			(3.38)			(2.67)
Post x Treated x Med				0.067***			0.054**			0.045***
				(4.17)			(2.33)			(4.65)
Post x Treated x High				0.053***			0.034			0.032*
				(4.02)			(1.76)			(2.04)
Size			0.007	0.007		0.019	0.019		0.000	0.000
			(0.61)	(0.62)		(1.10)	(1.09)		(0.02)	(0.02)
log(B/M)			0.022*	0.022*		0.039***	0.038***		0.018	0.018
			(2.02)	(2.02)		(3.43)	(3.42)		(1.37)	(1.37)
Momentum			-0.041***	-0.041***		-0.052***	-0.052***		-0.030***	-0.030***
			(-4.68)	(-4.68)		(-4.47)	(-4.48)		(-3.10)	(-3.09)
Profit			-0.180***	-0.180***		-0.160**	-0.160**		-0.165***	-0.165***
			(-3.39)	(-3.39)		(-2.58)	(-2.59)		(-4.24)	(-4.24)
Sigma			0.259***	0.259***		0.311**	0.311**		0.319**	0.319**
			(3.45)	(3.45)		(2.68)	(2.68)		(2.86)	(2.87)
ROEVol			-0.010*	-0.010*		-0.016	-0.016		-0.011**	-0.011**
			(-1.88)	(-1.90)		(-1.72)	(-1.72)		(-2.79)	(-2.79)
Merger FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
R2	0.106	0.433	0.442	0.442	0.392	0.396	0.396	0.481	0.488	0.488
N	62394	62394	62394	62394	57402	57402	57402	59946	59946	59946

Notes: This table shows the result of difference-in-difference test described in equation 3.2. The dependent variable are analyst coverage in column (1), 1 year ahead forecast in column (2), (3), and (4), 1 quarter ahead forecast in column (5), (6), and (7), and 2 years ahead forecast in column (8), (9), and (10). Post denotes an indicator variable that is equal to one if the observation is in the post-merger period and zero otherwise, and Treated is an indicator variable that is equal to one if a firm is treated and zero otherwise. The treated stocks are all stocks that experience a sudden drop in coverage due to brokerage mergers. Low is an indicator variable that is equal to one if the obervation is treated and has a pre-merger coverage level less than or equal to 10 analysts at each merger. Med is an indicator variable that is equal to one if the obervation is treated and has a pre-merger coverage level more than 10 and less than or equal to 20 analysts at each merger. High is an indicator variable that is equal to one if the obervation is treated and has a pre-merger coverage level more than 20 analysts at each merger. Size is the logarithm of a firm's market capitalization calculated as the number of shares outstanding times stock price. log(B/M) is the logarithm of book-to-market ratio following Fama and French (1993). Momentum is the past one year cumulative stock returns, and Profit is defined as operating income over book value of assets. Sigma is the variance of daily returns of a stock over past one year, and ROEVol is the past ten year annual ROEs where ROE is firm's return on equity measured as the ratio of earnings to the book value of equity. We also include merger fixed effects, year fixed effects, and firm fixed effects. The t-statistics in parentheses are robust to clustering at the merger groupings.

To examine Prediction 3, we divide treated firm into three groups as in Table 3.3 at each merger based on a firm's pre-merger coverage level. We then estimate equation 3.2 allowing the treated effect to differ among these three groups by interacting Post × Treated with Low, Med, or High. We report the results in columns (4), (7), and (10). Overall, the point estimates imply that the increase in dispersion is concentrated among firms with low initial coverage.

4 Conclusion

In this paper, we analyze whether financial analysts have strategic herding or anti-herding incentives when issuing forecasts about firms' earnings. To this end, we propose a new approach using the equilibrium predictions from the finite-player forecasting game of Kim and Shim (2019), and examine whether such incentives exist. Our empirical analysis using a plausible natrual experiment setting indicate that financial analysts exhibit strategic herding behavior in their forecasts. This finding is robust to forecast horizons and is more significant when there are fewer analysts around, consistent with predictions developed in our proposed framework.

documented in Table VII of Hong and Kacperczyk (2010).

References

Acemoglu, D., and M. K. Jensen (2013): "Aggregate Comparative Statics," *Games and Economic Behavior*, 81, 27–49.

Bernhardt, D., M. Campello, and E. Kutsoati (2006): "Who Herds?," *Journal of Financial Economics*, 80, 657–675.

Chen, Q., and W. Jiang (2006): "Analysts' Weighting of Private and Public Information," *Review of Financial Studies*, 19(1), 319–355.

Chen, Y., B. Kelly, and W. Wu (2020): "Sophisticated Investors and Market Efficiency: Evidence from a Natural Experiment," *Journal of Financial Economics*, 138(2), 316–341.

Clements, M. B. (1999): "Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter?," *Journal of Accounting and Economics*, 27, 285–303.

Clements, M. B., and S. Y. Tse (2005): "Financial Analyst Characteristics and Herding Behavior in Forecasting," *The Journal of Finance*, 60(1), 307–341.

Clements, M. P. (2018): "Do Macroforecasters Herd?," *Journal of Money, Credit and Banking*, 50(2-3), 265–292.

Cornes, R., and R. Hartley (2012): "Fully Aggregative Games," *Economics Letters*, 116(3), 631–633.

Croushore, D. (1997): "The Livingston Survey: Still Useful After All These Years," *Business Review - Federal Reserve Bank of Philadelphia*, 2, 1–12.

Derrien, F., and A. Kecskés (2013): "The Real Effects of Financial Shocks: Evidence from Exogenous Changes in Analyst Coverage," *Journal of Finance*, 68(4), 1407–1440.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002): "Differences of Opinion and the Cross Section of Stock Returns," *Journal of Finance*, 57(5), 2113–2141.

Fama, E. F., and K. R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1), 3–56.

Gallo, G. M., C. W. Granger, and Y. Jeon (2002): "Copycats and Common Swings: The Impact of the Use of Forecasts in Information Sets," *IMF Staff Papers*, 49(1), 4–21.

Hong, H., and M. Kacperczyk (2010): "Competition and Bias," *The Quarterly Journal of Economics*, 125(4), 1683–1725.

Hong, H., and J. D. Kubik (2003): "Analyzing the Analysts: Career Concerns and Biased Earnings Forecast," *Journal of Finance*, 58(1), 313–351.

Hong, H., J. D. Kubik, and A. Solomon (2000): "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31(1), 121–144.

Irani, R. M., and D. Oesch (2013): "Monitoring and Corporate Disclosure: Evidence from a Natural Experiment," *Journal of Financial Economics*, 109(2), 398–418.

Jegadeesh, N., and W. Kim (2010): "Do Analysts Herd? An Analysis of Recommendations and Market Reactions," *Review of Financial Studies*, 23(2), 901–937.

Kelly, B., and A. Ljungqvist (2012): "Testing Asymmetric-Information Asset Pricing Models," *Review of Financial Studies*, 25(5), 1366–1413.

Kim, J. Y., and M. Shim (2019): "Forecast Dispersion in Finite-Player Forecasting Games," *The B.E. Journal of Theoretical Economics*, 19(1).

Lamont, O. A. (2002): "Macroeconomic Forecasts and Microeconomic Forecasters," *Journal of Economic Behavior & Organization*, 48(3), 265–280.

Martimort, D., and L. Stole (2012): "Representing Equilibrium Aggregate in Aggregate Games with Applications to Common Agency," *Games and Economic Behavior*, 76, 753–772.

Ottaviani, M., and P. N. Sørensen (2006): "The Strategy of Professional Forecasting," *Journal of Financial Economics*, 81(2), 441–466.

Patton, A. J., and A. Timmermann (2010): "Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion," *Journal of Monetary Economics*, 57(7), 803–820.

Pouget, S., J. Sauvagnat, and S. Villeneuve (2017): "A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets," *The Review of Financial Studies*, 30(6), 2066–2109.

Prendergast, C., and L. Stole (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104, 1105–34.

Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *The American Economic Review*, 80(3), 465–479.

Trueman, B. (1994): "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7(1), 97–124.

van Binsbergen, J. H., X. Han, and A. Lopez-Lira (2020): "Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases," *NBER Working Paper No. 27843.*

Welch, I. (2000): "Herding Among Security Analysts," *Journal of Financial Economics*, 58(3), 369–396.

Zitzewitz, E. (2001): "Measuring Exaggeration by Analysts and Other Opinion Procedures," *Working Paper.*