FinTechs and the Market for Financial Analysis *

JILLIAN GRENNAN
Fuqua School of Business
Duke University

RONI MICHAELY
Johnson@Cornell Tech;
Cornell University

December 20, 2017

ABSTRACT

Does technology remedy the information inundation problem investors face or make it worst? By changing how investors discover information, FinTechs that streamline and synthesize all available information could serve to level the playing field if it is easier to find the best investment advice. Or, if investors rely on aggregated signals only, FinTechs could change the incentives of those producing information, and thereby, its quality. By gathering novel data on FinTechs, financial analysis online, and investors' internet clicks, we demonstrate FinTechs are counterproductive to the underlying goal of market efficiency. Supporting this conclusion, we show that in response to FinTech entry: (i) investors rely more on aggregate signals; (ii) traditional information-producers such as sell-side analysts reduce the quality of their reports; (iii) this change in information production stems from both an intensive and extensive margin; and (iv) price informativeness does not improve for equities that FinTechs focus on.

JEL classification: G1, G2, O3, D14, G11, G14, G23, O35

Keywords: FinTechs (Financial Technology firms), Sell-side equity research analysts, Financial bloggers, Innovation, Investment recommendations, Earnings estimates, News aggregators, Social media, Information and market efficiency, Datamining, Data signal providers

*Authors: Grennan, Duke University (e-mail: jillian.grennan@duke.edu); Michaely, Cornell University (e-mail: rm34@cornell.edu). We thank Itay Goldstein and Marina Niessner for helpful comments as well as seminar participants at the Federal Reserve Bank of Philadelphia and Duke University. We thank William Song for excellent research assistance. Some of the data used in this study comes from TipRanks, a firm in which Roni Michaely has an equity interest and serves on the Board of Directors.

"With 80% of the data in the world created in the last two years, judgment matters more than ever. Technology is a complement to sound judgment and knowledge, not a substitute."

— Joyce Chang, Global Head of Research, J.P. Morgan, September 2017

I. Introduction

Technology is changing how information is produced and discovered in financial markets. Yet the sheer quantity of information available is making it more difficult for investors to extract what matters most. To combat information overload and restore effective decision-making when one has too much information, financial technology firms or "FinTechs"[1] have begun to streamline and synthesize this abundance of information. Yet by aggregating information, FinTechs have the potential to distort the incentives of those that produce information, and thereby, the quality of information that is being extracted from the underlying data. In this paper, we consider the phenomenon of information overload, how FinTechs are changing it, and the possibility that some unintended consequence of this information aggregation may actually reduce the information content in prices.

Evaluating the implications of excess in our information-based world is important because an ideal financial market is one in which prices fully reflect available information, and thereby, the prices provide accurate signals for investors to allocate capital (Fama (1970); Bond, Edmans, Goldstein (2012)). While efficient outcomes depend on the accuracy of information, gains to informational efficiency since 1960 have been modest (Bai, Philippon, and Savov (2016)). FinTechs have the potential to significantly disrupt this status quo and create positive change. By revolutionizing how information is produced and delivered, FinTechs could help investors make better decisions by allowing them to sieve through the noise. Yet FinTechs function as "aggregators" may be counterproductive. Our analysis considers the potential for both positive and negative disruption.

As a motivating example of how aggregation could have negative effects, consider what happened when the movie ticketing firm Fandango integrated the score from Rotten Tomatoes into its platform. Ticket sales plummeted for movies receiving low scores because consumers relied on the aggregated score. Yet some of the movies received critical acclaim from prominent reviewers,

[1]FinTech covers digital innovations and technology-enabled business model innovations in the financial sector (Philippon (2016)). We focus on a single segment of innovations among many in the FinTech space. In particular, we examine capital market FinTechs, which are also often referred to as market intelligence FinTechs.

but consumers did not realize this because Rotten Tomatoes aggregates reviews from bloggers and YouTubers too. This aggregation feature lowers the informativeness of the overall score. It also incentivizes prominent reviewers when treated as the equals of YouTubers to generate less accurate reviews moving forward (New York Times (2017)). As a second example of the aggregation criticism, consider the election forecasting errors for Clinton-Trump. Slate (2016) argues when people pay more attention to Nate Silver's FiveThirtyEight forecast than to individual polls, the pollsters cater to a particular candidate and bias the poll in favor of that candidate.

We examine if this same logic applies to FinTechs that aggregate information about equities. The business model of FinTechs in the market for financial analysis is typically not to uncover novel private information about a stock but to aggregate and synthesize existent data and analysis. Given the sheer quantity of information out there, this streamlining feature is potentially very useful. FinTechs often take one of two approaches to synthesizing data. One approach is to use computer algorithms to condense the vastly expanded set of financial information into a buy or sell signal. As an example consider FirstAccess; they "turn big data into smart data, by separating the signal from the noise and delivering simple, reliable investment recommendations." Another approach is to evaluate and rank existent information and let the user pick his preferred information based on the ranking. As an example consider TipRanks; they provide a platform that allows investors to see a ranking of the historical performance of anyone who could be considered a financial expert (i.e., analysts, bloggers, corporate insiders, hedge fund managers, etc.).

Aggregation is important because it is unclear if this feature of FinTechs business model decreases or increases the attention paid to those that produce private information. The FinTechs that only provide a signal clearly divert attention, but those that rank information could increase attention, especially on those that provide the best information. In this sense, FinTechs could level the playing field and complement the information production process. In doing so, FinTechs could also potentially improve market efficiency by enhancing price informativeness. Alternatively, if FinTechs divert attention and investors rely more on aggregates rather than any individual piece of information, this may, in turn, generate changes in the production of financial information. Specifically, producers of financial information such as sell-side analysts may respond by decreasing their reporting quality (i.e., accuracy and bias). Thus, the entry of FinTechs could reduce the disciplinary forces that analysts endure for providing inaccurate analysis.

This idea that FinTechs could either level the playing field in financial markets or alter private information production is akin to theoretical arguments from the literature on optimal information disclosure (e.g., see reviews by Bond, Edmans, Goldstein (2012) and Goldstein and Yang (2017)). While the disclosure literature predicts that FinTechs will crowd-out the production of private information and decrease market efficiency, it is unclear if FinTechs necessitate the same logic. There are several reasons to believe FinTechs are their own distinct phenomenon (Philippon (2016)). For example, FinTechs introduce the possibility that trustworthy non-traditional information producers could be identified and that potential investors who were previously excluded from the market could participate and bring their own private information to the market when they join. In this sense, even if the overall quality of some information-producers declines because of the crowd-out effects predicted by the disclosure literature, this may not matter if FinTechs through their inclusion of non-traditional information and attractiveness to non-traditional investors enhance measures of market efficiency such as price informativeness.

To begin to understand how to view FinTechs and the disruptions they could bring to the market for financial analysis, data is necessary. One of the purposes of this paper is to gather a large, comprehensive database on FinTechs, alternative sources of financial advice, and investors' internet use that allows us to explore these issues. To gather data on FinTechs we hand-collect information from Crunchbase, FinTech award lists, and current and historical versions of the FinTechs' websites. Commentary suggests 2007 is the year when cloud computing and processing power improved enough to make data intelligence start-ups viable (Friedman (2016)). Consistent with this timeline, for the 290 FinTechs we observe, the mean founding year is 2008. The most common capabilities for these FinTechs are aggregating financial news (83% do this), datamining for investment signals (57% do this), evaluating and ranking existing financial advice (27% do this), crowdsourcing financial advice (16% do this), and aggregating financial experts' opinions (11% do this). 72% of the FinTechs target retail investors and 60% target professional investors with some targeting both. Our analysis suggest these firms are able to raise capital, which is consistent with VC reports indicating that about \$4 trillion has been invested in capital market FinTechs with most of it going to market intelligence FinTechs like the ones we study.

Given that the FinTechs in our data typically aggregate and streamline pre-existing financial analysis rather than producing their own content, we next turn to the non-traditional financial analysis that they often include in their aggregation. The most common source of non-traditional analysis is financial blogs. As such, we begin by examining the attributes of around 500 financial blogs. Since we are interested in the blogs that people actually read, we first matched a larger sample of blogs to internet click data from ComScore, which provides the most comprehensive data of this type available for Americans. To be included in our sample of 500 financial blogs, at least one internet user tracked by ComScore had to visit the financial blog between 2010 and 2016. Our data indicate that these investors want equity recommendations and they want that information quickly. While 90% of financial blogs that investors visit do not make buy or sell recommendations, investors strongly prefer to visit financial blogs that make equity recommendations. The financial blogs with stock recommendations rank 40 percentiles above blogs without stock recommendations in terms of pages visits and dwell time.

Next, to understand if the stock recommendations internet users look at are any good, we examine 1.3 million pieces of non-traditional investment advice analyzed by FinTechs between 2010 and 2016. This data comes from TipRanks, a FinTech operating in this space. First, it shows the typical blog post can be several paragraphs long and similar in nature to an equity analysts' report. But when we match this data with the ComScore data, we learn that investors rarely consume the whole report. In a given month, the average investor views 16 financial blog posts in only 6.6 minutes. The low dwell time per page may stem from the noisy nature of the analysis provided by bloggers. Our analysis of the 1.3 million blog posts across the 20 financial blogs that make buy-sell recommendations shows that 90% of the time, the market-adjusted returns to bloggers' recommendations were negative at an investment horizon of 6 or 12 months.

The richness of our data also allows us to explore the implications of FinTech entry in three different ways. First, at the investor level, we analyze internet click data to determine if FinTechs divert attention from original-content financial analysis. Second, at the data-producer level, we analyze the quality of analysts' research reports where FinTechs concentrate. And third, at the market level, we analyze how price informativeness and market reaction to analysts' recommendations change with FinTech entry.

At the investor level, we find that investors are 57 percentage points less likely to visit an original-content website when they also visit a FinTech website. Among those who do visit an original content website, they cut their page views in half and spend one-third less time on those websites. This finding is robust to a myriad of controls including income, education, race, age, and time. While the rotating nature of the ComScore sample does not allow for comparison of the same user over time, we create cell-blocks of users over time and follow them. For example, we create a cell-block for caucasian internet users between 30 and 39 years old with income over $100,000 and a college degree, and then we follow that cell-block over time. This within-estimator evidence suggests FinTechs are a true substitute for the reading of traditional financial analysis rather than a segmentation of user-type.

Next, we examine if FinTechs impact the production of financial information. To operationalize this test, we examine two important attributes of analyst reports: (1) aggregate earnings forecast accuracy and (2) aggregate optimism bias. We use an instrumental variable (IV) approach to test for analysts' responses. We proxy for FinTech concentration using the quantity of non-traditional sources that also provide financial analysis on that stock. The IV approach helps to address the challenge that a concentration of FinTechs is not randomly assigned. The ideal instrument is one that varies the amount of non-traditional information on an equity without impacting an analysts' reporting quality.

We find an ideal instrument using the insight that bloggers follow what's popular, often only adding their own commentary to what is already stale information. While popularity is not random, variation in popularity can come from things that are quasi-random. For example, using a randomized experiment, Umar (2017) showed that internet users are more likely to click on an article when the title is short even when the information content is the same. We build on this short title idea. Using data from Ravenpack and a text-based algorithm, we identify when newspapers publish articles about a particular equity with short titles. Unsurprisingly, weak instrument tests suggest this short-title instrument is associated with increased blogging, which is consistent with bloggers writing about firms that are popular. What is key is that this instrument also plausibly satisfies the exclusion restriction because title length is assigned by an editor in a way that is independent of analysts' aggregate accuracy or bias, especially for quantitative forecasts such as EPS. Moreover, the instrument does not require that short headlines do not also attract analysts' attention, rather we just need the short headlines to increase the amount of noise out there about an equity so FinTechs have a reason to concentrate on these equities.

Using the IV strategy, we find analysts respond to FinTech entry by reducing their reporting quality. We find higher aggregate absolute forecast errors and more optimistic bias for analysts where FinTechs concentrate. A one standard deviation increase in FinTech concentration is associated with a 0.30 standard deviation increase in aggregate optimistic bias and a 0.22 standard deviation increase in aggregate absolute forecast error. Respectively, these represent a 24% increase in aggregate optimistic bias and a 17% increase in aggregate absolute forecast error.

To understand the mechanism driving information-producers response, we examine the heterogeneity in our data. We find evidence to suggest that analysts' changes in reporting quality stem from both an intensive and an extensive margin. On the intensive margin, FinTechs are likely to alter existing incentives such as conflicts of interest. On the extensive margin, FinTechs are likely to change an analyst's outside job options, which could lead to compositional changes in the talent of analysts. We find analysts' reports for affiliated stocks (ones where their employer has served as an underwriter for that stocks' initial public offering, a seasoned equity offering or as an advisor on an M&A deal) are more biased in response to FinTech concentration than are their reports for non-affiliated stocks. Similarly, we find independent analysts exhibit a different response than non-independent analysts. And finally, we observe the bias is stronger among less experienced, unranked analysts suggesting at least part of the story stems from changes in the composition of those who choose to be research analysts.

Finally, at the market level, we analyze how price informativeness and market reaction to analysts' recommendations vary with FinTech concentration. We measure price informativeness using price nonsynchronicity (Roll (1988); Durnev et al. (2003)). We observe a small decrease in price informativeness in the stocks where FinTechs concentrate. Even when we focus on the highest quality alternative information, we find price informativeness does not improve (the point estimate is negative but insignificant). Lastly, we examine if the market reacts differently to analysts' recommendations where FinTechs concentrate. We find weaker market reactions where FinTechs concentrate. This can be interpreted in multiple ways, however, when considered in combination with the results on price informativeness, the data supports the story that FinTechs by aggregating financial information crowd-out private information production. On the whole, there is no evidence that FinTechs improve market efficiency.

Overall, our research contributes to the literature on FinTechs pioneered by Philippon (2016) and Yermack (2017). By examining the economic implications of these FinTechs, our research helps to show where FinTechs parallel and diverge from related phenomenon. These FinTechs are comparable to internet news aggregators in that they both divert attention from original content (Chiou and Tucker (2015); Calzada and Gil (2016)). The non-traditional data sources that are being analyzed by these FinTechs rarely provide good investment advice which supports prior research (Tumarkin and Whitelaw (2001); Antweiler and Frank (2004) Das and Chen (2007); Cookson, Niessner (2017)). While crowd-sourced advice can be useful (Jame et al. (2015); Da and Huang (2017)), we show this is an uncommon feature among FinTechs. In fact, our evidence suggests, in aggregate FinTechs crowd-out private information production by decreasing the quality of the information provided by the standard providers (i.e., analysts). This helps to show the parallels between FinTechs and above-optimal information disclosure (Bond, Edmans, Goldstein (2012); Goldstein and Yang (2017)) and to support theoretical predictions related to big data and market efficiency (Dugast and Foucault (2017)). Finally, our research contributes to the literature on analysts by helping to explain the forces shaping their recommendations (Hong and Kubik (2003); Barber, Lehavy, and Trueman (2007); Fang and Yasuda (2009); Merkley, Michaley, and Pacelli (2017)). One of the important implications from our paper is that FinTechs should not be viewed as traditional competition for analysts in the same way as increasing the supply of analysts is viewed but rather as disruptors that are facilitating a change in the way analysts disseminate information

(Gleason and Lee (2003); Berger, Ham, and Kaplan (2016)).

II. Hypotheses Development

In this section we present a stylized model that considers how investors access financial analysis online now that FinTechs aggregate and synthesize such analysis. The goal of the stylized model is to motivate some of our empirical hypotheses. In particular, we consider the ways in which the introduction of FinTechs may make the analysis provided by incumbents in this market (i.e., sell-side research analyst) more or less salient. First, we characterize how investors find online financial analysis. Next, we consider how FinTechs may alter the process by which investors discover financial analysis. This reveals that FinTechs may act as compliments or substitutes for the reading of incumbents' financial reports. We conclude this section by linking the intuition for FinTechs as complements or substitutes to the broader literature on disclosure in financial markets.

A. Investors Response to FinTech Entry

To characterize an individual investor's preferences for online financial analysis, let investor i have one unit of time t that he can allocate between reading financial analysis online and other activities. Every piece of financial analysis that he reads online has characteristics c, which are indexed as $d = 1, \ldots, D$. Characteristics such as the supplier of the information, the sentiment of the analysis, the stock covered, and/or the recommended investment horizon would be typical. The characteristic space is the set $C = \prod_{d=1,\ldots,D} C_d$. We denote investor i's reading of financial advice with characteristics c at time t by $A_{i,c}^t$ and if he reads many pieces of financial analysis, we denote his overall consumption of financial analysis as A_i^t.

Given that reading financial analysis online is one of many ways an investor can spend his time, let L_i^t denote his overall consumption of other leisure activities, L_i^t. Next, we assume the investor's utility from consuming the bundle $\left(A_i^t, L_i^t\right)$ follows a Cobb-Douglas utility function:

$$U_{it} = \left[\prod_{c \epsilon C} \left(A_{i,c}^t\right)^{a_{i,c}}\right]^{\tau^t \tau^i} L^{1-\tau^t \tau^i} \tag{1}$$

We use the Cobb-Douglas functional form to obtain the impliaction that a utility maximizing individual will consume a constant share of online financial advice with a particular set of characteristics, so long as the cost of finding online financial advice with different characteristics does not change.

To characterize how an investor finds online financial analysis, we consider two cases: without access to FinTech websites that aggregate and synthesize financial advice and with access to such websites. In both cases, we assume the investor does not directly know what analysis is available to read. FinTech websites will serve to speed up the discovery of financial analysis with a particular characteristic set. First, consider the case without FinTech websites that aggregate and synthesize financial analysis. To discover financial analysis, the investor must use an internet search engine or search after they've navigated to a financial blog. For each piece of financial analysis with characteristics c, the investor translates 1 unit of time into reading a quantity of financial analysis, denoted π_c. Let this process for finding online financial analysis have constant returns to scale so that allocating more or less time results in a proportional increase or decrease in the quantity of financial analysis found.

Second, consider the case with FinTech websites that aggregate financial analysis. These FinTech websites change the amount of financial analysis an investor can consume per unit of time in two ways. First, FinTech websites change the quantity of analysis with a given set of characteristics that can be found. Second, FinTech websites have a format that includes partial financial analysis such as buy or sell recommendations from a large variety of sources just from visiting the website. In fact, many FinTech websites that aggregate financial analysis have lists or rankings of top stocks based on their aggregation algorithm as well as links to various sources of financial analysis discussing those stocks. Appendix A provides examples of the website interface for different types of FinTechs. Overall, the different interfaces introduce three ways an investor can use his time to find and read financial analysis: (1) finding and reading original-content financial analysis from traditional search (π_c^{search}), (2) clicking-through to original-content financial analysis from a FinTech website ($\pi_c^{click-through}$), and (3) partially reading financial analysis via the FinTech website itself ($\pi_c^{fintech}$).

Given that the FinTech websites speed up the discovery of financial analysis, by definition they are more productive than traditional search alone for a given unit of time. That is, $\pi_c^{search} < \pi_c^{search} + \pi_c^{click-through} + \pi_c^{fintech}$. Even though the use of FinTech websites always weakly increases total financial analysis consumed per unit of time, it may actually decrease the quantity of financial analysis that is consumed directly from the original sources. This is the case when investors rely on $\pi_c^{fintech}$ rather than $\pi_c^{click-through}$. This explains how FinTech websites may make the original content of financial analysis more or less salient.

Formally, the FinTech websites serve as substitutes for the reading of original-content financial analysis when they produce fewer click-through than with search. That is, $\pi_c^{fintech} + \pi_c^{search} < \pi_c$. Conversely, the FinTech websites serve as complements for the reading of original-content financial analysis when they generate more click-throughs than search alone. That is, $\pi_c^{fintech} + \pi_c^{search} \geq \pi_c^{search}$. And utility maximization suggests an investors demand for financial analysis of type c with FinTech websites will be:

$$A_{i,c}^t = \tau^t \tau^i \alpha_{i,c} \left(\pi_c^{search} + \pi_c^{click-through} + \pi_c^{fintech} \right) \tag{2}$$

So far, we assume a characteristic c of an investor's search is if the financial analysis comes from a traditional or non-traditional source. In reality, when FinTechs aggregate and synthesize financial analysis, they make independent yet valuable analysis from non-traditional sources such as that from financial bloggers easier to discover. We explore how excluding the source of the analysis from the characteristic set C may decrease (increase) the reading of original-content financial analysis from sell-side research analysts even if FinTechs serve as complements (substitutes) for the reading of original analysis overall.

Consider the case where FinTechs in their effort to display the best analysis routinely display analysis from financial bloggers rather than financial analysts more prominently on their website. Let β denote the portion of the FinTech's website that displays financial analysis from non-traditional sources such as bloggers. Just as before the FinTech's website will always weakly increase total financial analysis consumed per unit of time but may in fact decrease the total financial analysis consumed from traditional sources. That is, $\pi_c^{search} < \pi_c^{search} + \beta \left(\pi_c^{click-through} + \pi_c^{fintech} \right) +$

$(1-\beta)\left(\pi_c^{click-through}+\pi_c^{fintech}\right)$.

Even if FinTechs serve as complements to the reading of original-content financial analysis they may still be directing readership away from incumbents in this market. In this way, the FinTechs serve as complements to non-traditional sources of financial analysis and substitutes for traditional sources of financial analysis. More formally, the FinTech websites serve as substitutes for the reading of original-content financial analysis from traditional sources when they produce fewer click-throughs than with search. That is, $(1-\beta)\left(\pi_c^{click-through}\right) < \pi_c^{search}$. Conversely, the FinTech websites serve as complements for the reading of original-content financial analysis from traditional searches when they generate more click-throughs than search alone. That is, $(1-\beta)\left(\pi_c^{click-through}\right) \geq \pi_c^{search}$.

B. Information-Producers Response to FinTech Entry

Having characterized how FinTechs could change the way financial analysis is discovered online, we now turn to the economic incentives for information producers that FinTech entry could change. In particular, we consider the role played by competition, outside employment options, and conflicts of interest. If FinTechs divert attention from original-content financial analysis for incumbents (substitutes view), then they produce anti-competitive effects that reduce effort, encourage catering to conflicts of interest, and incentivize alternative employment options. If FinTechs focus attention on incumbent's original-content financial analysis (complements view), then they produce competitive effects that increase effort, reduce catering to conflicts of interest, and discourage seeking alternative employment.

First, consider the role of competition. Theory suggests competition makes it more difficult for financial analysts to suppress unfavorable information (Gentzkow and Shapiro (2006)). Hence, competition incentivizes analysts to produce less biased, more accurate financial reports. Empirically, there is support for this view when the number of analysts covering a stock increases (Hong and Kacperczyk (2010)). While FinTechs do not increase the supply of analysts covering a stock, they have the potential to make the supply of non-traditional sources of financial analysis such as financial blogs more salient. Doing so, would allow for the independence channel of competition

(Gentzkow and Shapiro (2008)) to manifest. Namely, with a greater amount of financial analysis, there is a greater likelihood of drawing at least one supplier of financial analysis such as an independent blogger whose preferences cannot be bought or suppressed by the firm under study. In isolation, theory suggests this will discipline analysts.

But as we saw above, when FinTechs make non-traditional analysis easier to discover, this competition may influence analysts but not in a disciplinary way. This occurs when FinTechs through their placement of non-traditional financial analysis serve as a substitute for the readership of traditional analysis. A reduction in readership for the same quality of report may encourage an analyst to reduce his effort to provide quality financial analysis. It may also lead to a change in the composition of those who choose to be sell-side analysts (Merkley, Michaley, and Pacelli (2017)). It could encourage some more experienced analysts to leave their institutions given their reduced position of prominence and prestige. In this case, the whole pool of analysts may be younger and less qualified compared to previous generations. In other cases, it may depend on the relationships the analyst has if they choose to stay. For example, the unaffiliated analysts may leave and the ones who stay are the affiliated analysts.

Perhaps the most prominent change when FinTechs substitute for traditional readership is the incentives discouraging analysts from catering to conflicts of interest are lower. Potential conflicts of interest come from the analysts employer from its investment banking and/or brokerage business. When investment banking is an important source of revenue for the analysts employer, then the analyst may face pressure to inflate his recommendation. This pressure is due to the fact that the firm would like to sell investment banking services to accompany that the analyst tracks. The company, in turn, would like the analyst to support its stock with a favorable opinion. Similarly, analysts face conflicts from their employers' brokerage businesses. Here, the pressure on analysts originates not from the companies that they follow but from within their employing firms. Brokerage business generates a large portion of most securities firm's revenues, and analyst compensation schemes may be related to trading commissions. Thus, analysts have incentives to increase trading volumes which are more likely to increase with bullish recommendations as institutional investors often face short sale constraints.

C. Market Response to FinTech Entry

The overall desirability of FinTechs that aggregate and synthesize financial information depends on how they change the underlying information production process and ultimately, market quality. Many of the nuances to the argument about FinTechs are similar to those from the literature on information disclosure in financial markets (Bond, Edmans, Goldstein (2012); Goldstein and Yang (2017)). As such, we borrow from this literature to develop testable predictions for the aggregate impact of FinTechs. While FinTechs do not provide new information, they do have the potential to change the type of information being consumed from an individual piece of information to a more aggregated information signal. Information disclosure research shows that the type of information being disclosed is the key to determining whether disclosure is desirable or not.

Early research on information disclosure showed that it can increase the precision of public information, and thereby increase liquidity and market efficiency and decrease the cost of capital for firms and return volatility (e.g., Verrecchia (1982)). However, once the acquisition of private information is endogenized, disclosure can lead to the crowding out of private information (Diamond (1985)). Thus, disclosure could decrease market efficiency and increase return volatility. Other negative consequences highlighted in the literature include the reduction of trading opportunities (Kurlat and Veldkamp (2015)) and the promotion of destabilizing beauty-contest incentives where investors all want to do the same thing (Morris and Shin (2002)). In such a case, the greater precision of public information from disclosure leads investors to put too much weight on the information.

While there are similarities to the information disclosure literature, FinTechs are distinct phenomenon. For example, Dugast and Foucault (2017) argue the key feature of our current data abundance is that it decreases the cost of access to information and as such can reduce asset price informativeness. FinTechs through their aggregation of data clearly decrease the cost of access to information. Another unique dimension of FinTechs is the possibility that trustworthy information producers can be identified. In this sense, even if the overall quality of some information-producers declines, this may not matter if FinTechs through their inclusion of alternative sources of information, actually enhance price informativeness by filtering out untrustworthy data.

D. Empirical Predictions

The discussions above lead to several empirical implications. First, when FinTechs are substitutes (complements) for the readership of original-content financial analysis, an investor will spend less (more) time at an original-content website after visiting a FinTech website. Second, those in the business of private information production such as analysts will respond by decreasing (increasing) their reporting quality (i.e., accuracy and bias) when FinTechs serve as substitutes (complements). Third, analysts' changes in reporting quality will vary as a function of their existing conflicts of interest and outside options. And fourth, even if information production changes, the type of aggregation services that FinTechs provide will be desirable if common measures of market quality such as price informativeness improve.

III. Data and Descriptive Statistics

In this section we present descriptive statistics for the type of financial analysis investors have access to online now that there are FinTech websites that aggregate and synthesize such analysis. We begin by summarizing the financial blog data. The best characterization of the financial blog data is that it is ubiquitous, free, and very noisy. Hence, the need for FinTechs that sift through the noise to detect potentially useful analysis. We end this section by detailing the business plans of the FinTechs that are operating in this market and the ways in which they attempt to elevate the prominence of more accurate, higher quality financial analysis.

Table I summarizes data from financial blog websites that make buy and sell recommendations. On some of the blogs, buy or sell recommendations are made explicit at the end of the blog post, whereas other bloggers explain there thesis without summarizing their overall recommendation. For those that do not provide a summary, natural language processing (NLP) techniques are used to extract the recommendation. Overall, our data includes blog posts from 20 different financial blogs where bloggers make buy or sell recommendations on stocks. The data on financial blog posts comes from TipRanks, a FinTech firm operating in this market. Columns (2) through (6) characterize the internet traffic at the financial blogs. Specifically, Columns (2) and (3) rank the financial blog websites relative to all other websites in terms of page views and minutes spent on the website. Internet traffic data comes from comScore and is based on a nationally representative sample of about 50,000 U.S. internet users per month who have given comScore explicit permission to confidentially capture their detailed browsing behavior at the website level. User sessions are recorded with date and time stamps as well as clickstream data to show within an internet users session the number of pages viewed on a particular website. The sample of internet users changes on a monthly basis. Each month, we calculate the total number of pages views and seconds spent on each website. We, then, calculate the relative percentile for the financial blog websites among all websites. Percentiles allow for comparison over time as the total number of websites on the internet fluctuates.

The most popular financial blogs based on page views and minutes spent on the website are Market Watch, Motley Fool, The Street, Seeking Alpha, and Investor Place. Columns (4) through (6) of Table I present statistics about the typical users visit to the website. For example, among users that visit Market Watch, they visit the website 5 times per month and view 3 pages per visit spending a total of 4 minutes on the website per visit. Columns (7) through (10) show what the internet users are likely to encounter in terms of number of bloggers, blog posts, and stocks covered when they visit the financial blogs. There is no consistent format across blogs nor does there appear to be a correlation between format and popularity. For example, Seeking Alpha has over 10,000 unique bloggers whereas on average across the other financial blogs there are less than 300 bloggers per website. Despite significant variation in the number of unique bloggers, the number of stocks covered is more consistent across blogs. On average, each financial blog covers approximately 2000 stocks. Finally, Columns (11) through (13) present evidence on the average market-adjusted returns for stocks recommendations made on the blogs for a 1-month, 6-month, and 12-month period, respectively. The columns demonstrate how difficult it is to find useful financial advice among the blog posts. Almost all financial blogs earn negative market-adjusted returns, on average, over time. Moreover, the performance appears to be worst, on average, over longer horizons.

Table II provides more descriptive statistics about the sample of financial blog posts. In particular, we are interested in characterizing the way in which financial bloggers provide analysis similar in nature to that of sell-side analysts. Our sample includes 1,315,898 blog posts between 2010 and 2017. About 35% of blog posts provide a buy or sell recommendation on a stock. One-fifth of those blog posts have bearish recommendations while four-fifths are bullish. Among all blog posts there are 14,754 unique bloggers that cover 6,722 stocks. Among those that make buy or sell recommendations, there are 10,488 unique bloggers covering 6,385 stocks. Finally, among those that make at least 25 recommendations, there are 1,585 unique bloggers covering 6,210 unique stocks. These bloggers that are making multiple buy and sell recommendations across a variety of different stocks are those that are most similar to financial analysts. In term of the stocks covered in blogs posts, we observe 196 posts per stock and 12 posts per stock per quarter. We observe 73 recommendations per stock and 5 buy or sell recommendations per stock per quarter. Among those bloggers that make at least 25 recommendations, we see that they post to 1.4 blogs, on average, and have a total of 268 posts. These bloggers write a new blog post approximately every 16 days. The mean (median) number of stocks they cover is 94 (43). Similar to the performance at the blog-level, the performance of the bloggers with at least 25 recommendations (i.e., those that are most similar to equity analysts) demonstrate significant noise. The average financial blogger earns negative market-adjusted returns over time.

Table III characterizes the online market for financial analysis more broadly by describing a large sample of financial blogs and FinTechs in the market for financial analysis. To generate a comprehensive list of financial blogs and FinTechs, we use three different techniques. First, we search for relevant business descriptions on Crunchbase, a public database of company information about early-stage startups to Fortune 500 firms. Second, we search the internet for "Best of" FinTech and financial blog lists. Third, we use Google search to identify potentially relevant firms and blogs. Based on our initial list of FinTechs and blogs, we then examine each website to gather additional information and confirm that the website is in fact a financial blog or FinTech. To reduce survivorship bias, we use Wayback Machine to examine earlier versions of the website if the firm or blog stopped operating. For the financial blogs, the additional information we gather includes if the bloggers made equity recommendations or not and the general theme of the website. For the

FinTech websites, the additional information we gather includes business plan attributes such as what the firm does and its intended user.

Panel A of Table III describes our sample of financial blogs and Panel B describes our sample of FinTechs. To be part of the final sample of financial blogs or FinTechs, at least one internet user from the comScore sample of nationally representative U.S. households must visit the website between 2010 and 2017. The statistics in Panel A reveal that the vast majority of financial blogs (448 or 92.5% of our sample) do not make stock recommendations. A popular example of such a financial blog is `zerohedge.com`, which provides commentary on information that its contributors believe will "move the markets" or "break your trades." Rather than blog about specific stocks, these financial bloggers write about financial markets and investments. Internet users, however, prefer the financial blogs with specific stock recommendations. The mean (median) percentile for page views at financial blogs with stock recommendations is 75.5 (78.8) as compared to 38.7 (37.4) at those without recommendations. Similarly, the mean page views (8.3 vs. 2.4) and minutes per visit (3.3 vs. 2.3) are higher at the websites with recommendations.

Panel B of Table III describes our sample of FinTechs and their business plans. We observe 290 FinTechs operating in the market for financial analysis. We categorize the business operations of these FinTechs into: (1) those that aggregate data from financial experts (e.g., sell-side research analysts and/or bloggers), (2) those that aggregate financial news, (3) those that crowdsource financial advice, (4) those that datamine financial analysis and news for investment signals, and (5) those that rank and evaluate existing financial advice. These categories are not mutually exclusive. To be included in our sample, a FinTech's capabilities must include at least one of these functions. The most common capabilities are aggregating financial news (83% do this), datamining for investment signals (57% do this), and evaluating and ranking existent financial advice (27% do this). Overall, the business plan analysis shows that these firms aggregate and streamline pre-existing financial analysis rather than produce their own original-content.

Column (2) of Panel B shows the mean founding year of FinTechs in our sample is 2008. Column (3) and (4) reveal that 72% of FinTechs target retail investors and 60% target professional investors with some targeting both. Among the different business functions, FinTechs that crowdsource financial advice primarily target retail investors (89%) while those that datamine primarily target professional investors (70%). Column (5) shows that one-in-five FinTechs focus only on a specific type of stock such as consumer goods rather than try to cover all stocks. Columns (6) through (8) demonstrate that many of these FinTechs are credible businesses in the eyes of the investment community. With the average FinTech in the market for financial analysis raising \$10.4 million from 4.8 investors and employing 74 workers.

IV. Empirical Strategy

A. Investors Response to FinTech Entry

To test the hypothesis that FinTech entry in the market for financial analysis changes how an investor discovers financial information, we examine internet traffic data to detect changes in what investors read online. Specifically, we estimate the following equation:

$$OriginalAnalysis_{it} = \alpha + \beta FinTechVisit_{it} + \theta X_{it} + f_i + \delta_t + \epsilon_{it} \tag{3}$$

where $OriginaAnalysis_{it}$ represents a visit to the website containing original-content financial analysis in month t for household i, $FinTechVisit_{it}$ indicates if the household visited a FinTech website in that month, X_{it} is a vector of observables (income, race, age, education, census region, internet connection speed, and number of children), f_i is a household fixed effect, δ_t is a quarter fixed effect, and ϵ_{it} is the unobservable error component. We also consider variations on the definition of $OriginalAnalysis_{it}$ including the number of pages viewed on an original-content website as well as the time spent on the original-content website.

B. Information-Producers Response to FinTech Entry

To test the hypothesis that information producers respond to FinTech entry, we study changes in the optimism bias and accuracy of analysts' earnings forecasts as a function of FinTech concentration in the stocks they cover. We proxy for FinTech concerntation using the frequency of financial blog posts in a given quarter about a stock that the analyst covers. Given that FinTechs aggregate and streamline such financial analysis, their presence directly corresponds to the frequency of financial blog coverage. To provide a credible point estimate and mitigate the influence of factors endogenous to the data generating process for analysts' reporting quality, we use an instrumental variable approach.

Specifically, we use a linguistic-based instrument to generate variation in the concentration of alternative sources covering an equity. We use a text-based algorithm to identify if newspaper headlines use short title lengths to increase readership. Using a randomized experiment, Umar (2017) showed that internet users are more likely to click on an article when the title is short even if the information conveyed by the title was constant. The relevance condition is that bloggers focus on what is popular so are more likely to blog after reading a popular article. The exclusion restriction for IV identification requires the attention-grabbing financial news headlines only alter an analysts' aggregate accuracy or bias via the effect of additional blogging about an equity increasing the concentration of FinTechs in that equity. Given that the restriction relates quantities I cannot observe together, I cannot test it. Rather arguments must support the plausibility of satisfying the restriction. In this case, the main argument is financial news headlines are quasi-random since they are selected at the discretion of the editor. Moreover, even if the headline makes it so that the analyst also reads the article, being current on the latest developments for an equity they cover is exactly what analysts are supposed to do, so this should not change their reporting quality, especially for hard information such as the quantities associated with EPS.

The instrumental variable specification is as follows:

$$ReportQuality_{it} = \alpha + \beta FinTechConc_{it} + \theta X_{it} + f_i + \delta_t + \epsilon_{it} \tag{4}$$

where $ReportQuality_{it}$ represents characterizes the analysts' report quality in terms of optimism bias and accuracy in quarter t for equity i, $FinTechConc_{it}$ proxies for FinTech concentration using the quantity of alternative data sources such as financial blog posts that discuss equity i in quarter t, X_{it} is a vector of observables (analyst coverage, firm size, daily return volatility, mean monthly return, log market-to-book ration, volatility of ROE, profitability, and an indicator for if the stock is a member of the S&P 500), f_i is an firm fixed effect, δ_t is a quarter fixed effect, and ϵ_{it} is the unobservable error component.

C. Market Response to FinTech Entry

To test the hypothesis that the type of aggregation services that FinTechs provide is desirable, we examine a common measure of market quality that of price informativeness. We measure price informativeness using price nonsynchronicity (Roll (1988); Durnev et al. (2003)). It is computed on the basis of the correlation between the stock's return and the return of the corresponding industry and of the market. The idea is that if a firm's stock return is strongly correlated with the market and industry returns, then the firm's stock price is less likely to convey firm-specific information.

The regression specification is as follows:

$$Info_{it} = \alpha + \beta FinTechConc_{it} + \theta X_{it} + f_i + \delta_t + \epsilon_{it} \tag{5}$$

where $Info_{it}$ represents price nonsynchronity in quarter t for equity i, $FinTechConc_{it}$ proxies for FinTech concentration using the quantity of alternative data sources such as financial blog posts that discuss equity i in quarter t, X_{it} is a vector of observables (analyst coverage, firm size, daily return volatility, mean monthly return, log market-to-book ration, volatility of ROE, profitability, and an indicator for if the stock is a member of the S&P 500), f_i is an firm fixed effect, δ_t is a quarter fixed effect, and ϵ_{it} is the unobservable error component.

V. Results

A. Investors Response to FinTech Entry

Table IV examines how investors discover financial analysis online. Column (1) describes who reads financial analysis online. Unsurprisingly, the data indicate older investors with higher income and a college degree are more likely to read financial analysis online. Columns (2) through (4) examine if FinTechs serve as substitutes or complements for the reading of original-content financial analysis. Column (2) reveals investors are 57 percentage points less likely to visit an original-content website if they visit a FinTech website. This correlation is highly statistically significant. The variation explained by the regression is 55% which suggests these variables meaningfully explain investors visits to websites with original-content financial analysis. Overall, this first piece of evidence suggest that FinTechs serve as substitutes. That is, people read the snippets of analysis on the FinTech websites rather than clicking-through to the original content.

Column (3) and (4) further support the view that FinTechs serve as substitutes. Column (3) shows that the page views at original-content websites are reduced by 55% and Column (4) shows that the time spent at the original-content website is reduced by 33% when an investor visits a FinTech website. These results are statistically significant and these findings are robust to controls for income, race, age, education, census region, internet connection speed, number of children. Further, monthly fixed effects help to rule out changes in preference over time as driving these results. Taken together, the internet traffic data suggests that investors use FinTechs as a substitute for traditional financial analysis.

While the rotating nature of the ComScore sample does not allow for comparison of the same user over time periods greater than one month, we create cell-blocks of users over time and follow them. For example, we create a cell-block for white internet users between 30 and 39 years old with income over $100,000 and a college degree, and then we follow that cell-block over time. Our within-estimator evidence also suggests FinTechs are a true substitute for the reading of traditional financial analysis rather than a segmentation of user-type. However, it is important to note that these are simply correlations.

B. Information-Producers Response to FinTech Entry

Next, we turn to the financial analyst sample to understand if analysts respond to the changes stemming from FinTechs. Table V summarizes our sample of data on financial analysts. It displays the analyst coverage, firm size, daily return volatility, mean monthly return, the log of market-to-book, volatility of ROE, profitability, and inclusion in the S&P 500. We note these are the exact same controls used by Hong and Kacperczyk (2010). We explore both cross-sectional and within-equity variation. If financial blog coverage is fairly persistent over time, then focusing on within-equity variation (i.e., with firm fixed-effects) may be too restrictive.

We begin by running OLS regressions and present these results in Table VI. We first present the results with just time fixed effects in Columns (1), (3), (5), and (7), while we present the results with time and firm-fixed effects in Columns (2), (4), (6), and (8). The results show a small positive partial correlation between financial blog coverage and analysts' aggregate optimism bias and aggregate absolute forecast error. Note this is absolute forecast error, so a bigger value means the analysts report is less accurate. In general, we see a small positive increase in bias and reduction in accuracy both in the cross-section and within-equity over time.

Next, we estimate IV regressions and present those results in Table VII. We find that an increase in financial blog coverage for a stock is associated with higher aggregate absolute forecast errors and more optimistic bias for analysts that cover that firm. A one standard deviation increase in the number of blog posts written about a firm are associated with a 0.30 standard deviation increase in aggregate optimistic bias and a 0.22 standard deviation increase in aggregate absolute forecast error. Respectively, these represent a 24% increase in aggregate optimistic bias and a 17% increase in aggregate absolute forecast error.

The statistical evidence for the deterioration in reporting quality is significant at the 99th percentile. The F-statistic from the first stage of the instrumental variable regression is 204.2, which exceeds the requisite 10 to ensure minimal bias of the point estimate. The instrumental variable specification includes controls for analyst coverage, firm size, daily return volatility, mean monthly returns, market-to-book, volatility of ROE, profitability, membership in the S&P 500, momentum, institutional ownership as well as firm and industry-by-time fixed effects. These controls help to account for other industry dynamics that may cause analysts reporting quality to deteriorate.

In Table VIII, we focus on blog entries by the financial bloggers FinTechs firms identify as having the highest quality recommendations. We find that the effect on quality of analyst reports is strongest when we focus on blog entries by the financial bloggers FinTechs firms identify as having the highest quality recommendations. Specifically, we find a 0.37 and 0.26 standard deviation increase in aggregate optimistic bias and aggregate absolute forecast error, respectively. These represent a 30% increase in aggregate optimistic bias and a 20% increase in aggregate absolute forecast error. The results are similar when we bloggers ranked as having high quality recommendations in the short-term (investments under one year) and in the long-term (investments over a year). Overall, the instrumental variable results provide a consistent message that increases in coverage by financial bloggers adversely affect the overall quality of analyst reports.

In Table IX, we change our analysis to the analyst-equity-quarter level. This allows to explore what characteristics of analysts and their employers may be associated with more or less response to FinTech concentration. Column (1) repeats the previous analysis where the dependent variable is optimism bias. As with before, we see an increase in optimism bias where FinTechs concentrate at this more disaggregated analyst-level. The disaggregation allows us to include many additional controls for analyst and brokerage characteristics. Specifically, we control for analyst experience, their experience covering that equity, the number of equities they cover, the number of industries they cover, their average forecast frequency, forecast horizon, and days since last forecast, whether the stock they are covering is an affiliated stocks (ones where their employer has served as an underwriter for that stocks' initial public offering, a seasoned equity offering or as an advisor on an M&A deal), the brokerage size, and if the brokerage is independent.

Next, we analyze sub-samples of the data to understand which economic incentives are influencing analysts to change. Specifically, Columns (2) and (3) focus on affiliated vs. non-affiliated stocks. We see the analysts reports for affiliated stocks exhibit 0.13 standard deviation higher increase in optimism bias in response to FinTech concentration. This supports the notion that investment banking conflicts of interest may be inducing analysts to cater to those clients when FinTechs divert attention from their research. It could also support the notion that the composition of analysts is changing in response to FinTechs. This would be the case if some analysts, especially the unaffiliated analysts leave the profession, and the ones who stay are the affiliated analysts whose forecast are more biased and less accurate.

As an alternative cut on the data, Columns (4) and (5) focus on independent and non-independent brokerage houses. Again, we see the analysts' reports for from non-independent brokerage houses exhibit more increase in optimism bias (0.08 standard deviations) in response to FinTech concentration. This is consistent with a story of analysts catering to their conflicts of interest after attention has been diverted from their research. Finally, in Columns (6) and (7) we elaborate on the change in composition of the analyst workforce argument. Our evidence indicates that inexperienced analysts show more increase in optimism bias (0.17 standard deviations) in response to FinTech concentration. Overall, our evidence suggests both catering to conflicts of interest and changes in the attractiveness of outside employment opportunities are driving analysts response.

C. Market Response to FinTech Entry

In Table X we analyze how price informativeness changes with FinTech entry. While the results about information production by analysts are intriguing in their own right, it is still possible that analysts' information is replaced by more precise information aggregation from bloggers and alternative data sources, so the overall impact of FinTechs and information overload is good since it increases price informativeness. We find price informativeness decreases where FinTechs concentrate. We observe beween a 5 and 8 percentage point decrease in price informativeness in the stocks where FinTechs concentrate, depending on the set of control variables used. The price informativeness results continue to hold when we examine the highest quality financial bloggers suggesting that FinTechs are not improving market efficiency, although the point estimate is smaller.

As an alternative test of price informativeness, in Table XI we examine if the market reacts differently to analysts' recommendations where FinTechs concentrate. We find weaker market reactions where FinTechs concentrate. This holds for upgrades to buy recommendations but not downgrades to sell reccommendations. Specifically, we see a 49 b.p. difference on the day of the announcement and this grows to a 67 b.p. difference with time. Overall this result is consistent with the hypothesis that FinTechs crowd out the production of private information rather than leveling the playing field.

D. Robustness

To understand the robustness and generalizability of our findings, we conduct several additional tests. First, we consider alternative definitions for measuring changes in the quality of information being produced. We examine median response rather than mean response. Then, we examine more nuanced measures of accuracy and bias developed in the accounting literature. Our results do not change when we use these alternative definitions. We also change our level of analysis from the equity level to the analyst level. Doing this allows to include many more controls for analysts such as their brokerage house, if they are an all-star, the number of equities they cover, the frequency with which they update their estimates, etc. This change in level of analysis, does not alter our findings. To assess potential bias introduced by our instrument, we limit the set of newspapers that we examine. Specifically, we exclude New York newspapers that analysts may draw soft information from.

As a second test, we consider other tactics editors use beside title length to enhance readership. The additional tactics we examine include: surprise, questions, curiosity gap, tone, "how to," precision, and audience reference. For example, surprises in headlines work because novelty releases additional endorphins in the brain. Compared to expected pleasant news, unpredicted pleasant news turns on the pleasure centers of the brain even more. Thus, surprises stimulate and grab attention more than other headlines. We find similar results when testing these alternative instruments. Finally, we explore the construct of our measure of FinTech concentration. Rather than look at any non-traditional sources for financial analysis, we limit our sample to high-quality non-traditional sources. We use the FinTech website TipRanks measure of quality to distinguish sources. Overall, the results are consistent with the findings from our primary set of tests.

VI. Conclusion

Our information-based society is often plagued with excess. There are many areas of everyday life in which information overload prevails, but the investment sector may well be where the consequences are the most serious. As new technology promises to remedy the information inundation problem investors face, we do not know if it is actually helping. This paper considers the phenomenon of information overload, how FinTechs are changing it, and the possibility that FinTechs may be counterproductive to the underlying goal of market efficiency.

In particular, this paper gathers a new, comprehensive dataset on FinTechs, financial bloggers, and investors internet use. We use this dataset to document the new options investors have for discovering financial analysis brought about by the entry of FinTechs. The services provided by the FinTechs range from the aggregation of existing financial analysis to the creation of customized buy-sell signals compiled from traditional and non-traditional data sources (financial news, analyst reports, blog posts, tweets, etc.).

Our econometric investigation of this data suggests FinTechs are counterproductive to the underlying goal of market efficiency. In particular, we find investors rely on aggregated signals; they view fewer webpages with original-content financial analysis and spend less time on original-content financial analysis websites. While this may be saving investors time, this also distorts the incentives of those that produce financial information. When FinTechs enter the market for financial analysis, we observe significant decreases in analysts' accuracy and increases in optimism bias for the equities where the FinTechs concentrate. The change in reporting quality stems from an intensive margin in that it is greatest for equities where analysts' conflicts of interest are strongest. The change also stems from an extensive margin in that the new talent pool among analysts is diminishing. Overall, we find price informativeness decreases where FinTechs concentrate, suggesting that analysts information is not being replaced by more precise information aggregated from alternative sources.

REFERENCES

Allen, F., D. Gale, 1994, *Financial Innovation and Risk Sharing*, Cambridge, MA: MIT Press.

Anderson, S., S. Coate. 2005. Market provision of broadcasting: a welfare analysis. *Review of Economic Studies* 72: 947–972.

Andrews, M., M. Van Leeuwen, R. Van Baaren. 2014. *Hidden Persuasion: Psychological Influences in Advertising.* Amsterdam: BIS Publishers.

Antweiler, W., M. Frank. 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59:1259–1294.

Asquith, P., M. Mikhail, A. Au, 2005. Information content of equity analyst reports. *Journal of Financial Economics* 75:245–282.

Bai, J., Philippon, T., A. Savov. 2016. Have financial markets become more informative? *Journal of Financial Economics* 122:625–654.

Bailey, W., H. Li, C. Mao, R. Zhong. 2003. Regulation fair disclosure and earnings information: Market, analysts, and corporate responses. *Journal of Finance* 58:2487–2514.

Barber, B., R. Lehavy, B. Trueman. 2007. Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics* 85:490–517.

Berger, P., C. Ham, and Z. Kaplan. 2016. Do Analysts Say Anything About Earnings Without Revising Their Earnings Forecasts?, Working Paper.

Bond, P., Edmans, A., Goldstein, I., 2012. The real effects of financial markets. *Annual Review of Financial Eocnomics* 4:339–360.

Calzada, J., R. Gil. 2016. What do news aggregators do? Evidence from google News in Spain and Germany. Working Paper, Johns Hopkins University.

Chen, H., P. De, Y. Hu, B. Hwang, 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27:1367–1403.

Chiou, L. and C. Tucker. 2015. Content aggregation by platforms: The case of the news media. Working Paper, NBER No. 21404.

Cohen, L., Frazzini, A., Malloy, C. 2010. Sell-side school ties. *Journal of Finance* 65:1409–1437.

Cookson, A., M. Niessner, 2017, Why Dont We Agree? Evidence from a social network of investors, Working Paper.

Das, S., and M. Chen. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science* 53:1375–1388.

Da, Z., X. Huang. 2017. Harnessing the wisdom of crowds, Working Paper.

Diamond, D. 1985. Optimal Release of Information By Firms, *Journal of Finance* 40, 1071–1094.

Dugast, J., and T. Foucault. 2017. Data Abundance and Asset Price Informativeness. *Journal of Financial Economics*, forthcoming.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does Greater Firm-specific Return Variation Mean More or Less Informed Stock Pricing? Journal of Accounting Research 41: 797–836.

Fama, E. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25:383–417.

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.

Ferreira, M., P. Matos, 2008, The colors of institutions' money: The role of institutional investors around the world, *Journal of Financial Economics* 88, 499–533.

Frankel, R., S.P. Kothari, J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41:29–54.

Friedman, T. 2016. *Thank you for being late.* New York: Farrar Straus Giroux.

Gentzkow, M., J. Shapiro. 2006. Media bias and reputation. *Journal of Political Economy* 114:280–316.

Gentzkow, M., J. Shapiro. 2008. Competition and Truth in the Market for News. *Journal of Economic Perspectives* 22:133–154.

Gentzkow, M., J. Shapiro, M. Sinkinson. 2011. The Effect of Newspaper Entry and Exit on Electoral Politics *American Economic Review* 101:2980–3018.

Gleason C., Lee C. 2003. Analyst forecast revisions and market price discovery. *Accounting Review* 78, 193–225.

Greenstein, S., and F. Zhu. 2012. Is Wikipedia biased? *American Economic Review P&P*, 343–348.

Greenstein, S., J. Lerner, S. Stern (2011). The economics of digitization: An agenda. NSF Report.

Grennan, J., and D. Musto. 2017. Who benefits from bond market modernization? Working paper.

Goldstein, I. and L. Yang, 2017. Information Disclosure in Financial Markets. *Annual Reviews of Financial Economics*, forthcoming.

Hong, H., M. Kacperczyk. 2010. Competition and bias. *Quarterly Journal of Economics* 58:1683–1725.

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

Jame, R., R. Johnston, S. Markov, M. Wolfe. 2015. The Value of Crowdsourced Earnings Forecasts. *Journal of Accounting Research* 54:1077–1110.

Kurlat, P. and Veldkamp, L. 2015. Should We Regulate Financial Information? *Journal of Economic Theory* 158, 697–720.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-ks, *Journal of Finance* 66:35–65.

Lim, T. 2001. Rationality and analysts' forecast bias. *Journal of Finance* 56:369–385.

Ljungqvist, A., F. Marston, L. Starks, K. Wei, H. Yan. 2007. Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics* 85:420–456.

Ljungqvist, A., F. Marston, and W. Wilhelm. 2006. Competing for securities underwriting mandates: Banking relationships and analyst recommendations. *Journal of Finance* 61:301–340.

The New York Times. 2017. Attacked by Rotten Tomatoes. https://nyti.ms/2xcAtwv. September 7.

Malmendier, U. D. Shanthikumar. 2014. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 27:1287–1322.

Merkley, K., R. Michaely, J. Pacelli. 2017. Does the scope of sell-side analyst industry matter? An examination of bias, accuracy and information content of analyst reports. *Journal of Finance*, forthcoming.

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

Michaley, R., K. Womack. 2005. Market efficiency and biases in brokerage recommendations. In Advances in behavioral finance II, ed. R. H. Thaler, 389419. New York: Russell Sage Foundation.

Morris, S., Shin, H. S., 2002. The social value of public information. *American Economic Review* 92:1521–1534.

Mullainathan, S., A. Shleifer. 2005. The market for news. *American Economic Review* 95:1031–1053.

Philippon, T. 2015. Has the U.S. finance industry become less efficient? *American Economic Review* 105:1408–1438.

Philippon, T. 2016. The fintech opportunity. Working Paper, NBER No. 22476.

Roll, R., 1988. "R^2" *Journal of Finance* 43:541–566.

Schumpeter, J. 1934. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle.* Oxford University Press: London, UK.

Slate. 2016. Which Election Forecast Should You Trust? August 2.

Tetlock, P. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62:1139–1168.

Tumarkin, R., R. Whitelaw. 2001. News or noise? Internet postings and stock prices. *Financial Analysts Journal* 57:41–51.

Umar, T., 2017. Complexity aversion when seeking alpha, Working Paper, University of Chicago.

Verrecchia, R. 1982. Information Acquisition in a Noisy Rational Expectations Economy. *Econometrica* 50, 1415–1430.

Yermack, D. 2017. Corporate governance and blockchains. *Review of Finance*, forthcoming.

Zingales, L. 2015. Does finance benefit society? *Journal of Finance* 70:1327–1363.

Table I. Financial Blogs with Stock Recommendations

This table presents summary statistics for our sample of financial blog sites that make stock recommendations over the sample period from 2010-2017. Columns (2) through (6) characterize the mean internet traffic at the blogs. Columns (7) through (10) characterize the content investors would encounter when they visit these blog sites. Columns (11) through (13) report market-adjusted returns based on the recommendations made on the blog for 1-month, 6-months, and 12-months, respectively. For a detailed description of each variable, see Appendix A.

Blog Site (1)	Page View Percentile (2)	Minutes on Site Percentile (3)	Monthly Visits per User (4)	Page Views per Visit (5)	Minutes per Visit (6)	Pct. of Tot. Blog Posts (7)	Pct. of Posts with Rec. (8)	Num. of Unique Bloggers (9)	Num. of Stocks Covered (10)	Market-adjusted 1-month Returns (11)	Market-adjusted 6-month Returns (12)	Market-adjusted 12-month Returns (13)
MarketWatch	99.7	99.9	4.8	2.7	3.8	1.6%	11%	953	2,110	-0.5%	-3.0%	-4.6%
MotleyFool	99.3	99.5	2.0	1.9	2.4	17.1%	30%	1,204	4,424	0.2%	-2.4%	-5.4%
TheStreet	99.2	99.4	2.5	3.3	4.8	18.5%	18%	664	5,172	0.7%	-4.3%	-8.6%
SeekingAlpha	97.0	97.2	2.8	2.4	4.3	32.9%	43%	10,442	6,501	0.3%	-2.5%	-5.7%
Zacks	96.9	96.6	1.8	2.9	2.8	10.6%	23%	111	4,369	0.2%	-3.7%	-6.9%
InvestorPlace	95.0	95.0	2.2	3.3	3.2	7.8%	83%	207	5,040	0.0%	-4.0%	-6.7%
MoneyMorning	92.4	93.1	1.3	1.2	1.6	0.1%	35%	46	482	0.4%	-3.3%	-6.4%
StreetAuthority	89.0	89.6	1.7	2.3	2.3	0.3%	67%	82	1,219	0.1%	-2.4%	-5.4%
GuruFocus	85.8	82.1	1.8	5.0	3.1	3.2%	34%	796	4,168	0.0%	-2.1%	-3.4%
Kapitall	81.6	60.6	1.3	5.8	2.0	0.3%	51%	40	1,764	0.1%	-2.6%	-5.4%
MarketRealist	79.7	70.4	1.4	2.7	2.5	0.7%	20%	47	287	0.1%	0.8%	0.9%
Amigo Bulls	78.0	60.5	1.1	4.1	3.4	0.1%	47%	46	157	0.0%	-0.1%	-0.2%
MoneyShow	73.1	71.1	1.9	6.3	3.5	0.3%	47%	352	1,231	0.0%	-4.0%	-8.3%
Investing	68.1	69.3	4.7	7.5	3.8	1.4%	28%	480	3,105	0.1%	-1.8%	-3.8%
Who Trades	67.8	54.7	1.1	1.9	1.1	0.3%	28%	67	856	0.7%	1.3%	1.2%
TopStockAnalysts	66.8	65.5	1.4	1.8	1.7	0.4%	34%	175	1,647	0.2%	-3.3%	-7.6%
SmarterAnalyst	65.1	47.0	1.6	2.2	1.5	0.1%	97%	75	546	0.6%	-1.7%	-4.8%
ProfitableTrading	57.8	58.2	1.3	1.6	1.8	0.0%	95%	24	405	0.5%	0.0%	-3.7%
SumZero	38.9	29.8	1.5	2.6	2.1	0.0%	98%	1	90	1.5%	10.6%	14.6%
WSObserver	34.4	34.7	1.5	1.5	1.3	4.3%	27%	18	3,335	0.1%	-3.9%	-3.8%

Table II. Characterizing Financial Blog Posts

This table presents summary statistics for our sample of financial blog sites that make stock recommendations over the sample period from 2010-2017. This table provides descriptive statistics about bloggers posts, their recommendations, the stocks they cover, the number of sites the bloggers post to, the days between posts, and the market-adjusted returns associated with their recommendations. For a detailed description of each variable, see the definitions in Appendix A.

Year	Freq.
2010	46,360
2011	110,606
2012	144,868
2013	180,293
2014	257,444
2015	291,201
2016	196,637
2017	88,489
Total	1,315,898

Sentiment	Freq.
Bearish	81,063
Neutral	851,708
Bullish	383,127
Total	1,315,898

Among all blog posts	Freq.
Unique bloggers	14,754
Unique stocks	6,722

Among posts with non-neutral recs	Freq.
Unique bloggers	10,488
Unique stocks	6,385

Among bloggers with at least 25 recs	Freq.
Unique bloggers	1,585
Unique stocks	6,210

Stocks covered in blog posts	Mean
Blog posts per stock	196
Blog posts per stock per quarter	12
Recs per stock	73
Recs per stock per quarter	5

Among all bloggers	Mean	Median
Number of sites bloggers post to	1.1	1.0
Number of posts per blogger	89.2	4.0
Days between blog posts	65.8	23.3
Number of stocks covered	24.8	3.0

Among bloggers with at least 25 recs	Mean	Median
Number of sites bloggers post to	1.4	1.0
Number of posts per blogger	267.8	70.0
Days between blog posts	16.3	10.2
Number of stocks covered	94.6	43.0

Performance with at least 25 recs	Mean	Median
Market-adjusted 1-month Return	0.2%	0.0%
Market-adjusted 3-month Return	-1.5%	-0.2%
Market-adjusted 6-month Return	-3.1%	-0.8%
Market-adjusted 12-month Return	-6.1%	-2.0%

Table III. Financial Blogs and FinTechs

This table presents summary statistics for a broader sample of financial blogs and FinTechs in the market for financial analysis. Panel A describes our sample of financial blogs and Panel B describes our sample of FinTechs. To be part of the sample of financial blogs or FinTechs, at least one internet user from the comScore sample of nationally representative U.S. households must visit the website between 2010 and 2017. Columns (1) through (4) of Panel A describe the mean internet traffic at all financial blogs with stock recommendations and columns (5) through (8) for those without stock recommendations. Panel B describes the business operations of the FinTechs and their progress as a business. For a detailed description of each variable, see the definitions in Appendix A.

	Financial blogs with stock recs.				Financial blogs without stock recs.			
	Mean	Std. Dev.	Median	Max	Mean	Std. Dev.	Median	Max
Panel A. Characteristics of financial blogs	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Page view percentile	75.46	20.91	78.81	99.73	38.73	22.85	37.39	99.92
Minutes on site percentile	72.16	23.12	78.40	99.88	36.50	23.03	33.43	99.92
Monthly visits per user	2.0	1.2	1.5	6.5	1.4	1.4	1.1	17.9
Page views per visit	8.3	22.4	2.7	109.3	2.4	2.0	1.9	17.1
Minutes per visit	3.3	3.4	2.5	19.3	2.3	2.9	1.5	36.0
Observations	36				448			

	Obs.	Mean Year Founded	Targets Retail Investors	Targets Prof. Investors	Covers Specific Stocks	Mean Num. of Investors	Mean Funding ($mil)	Mean Num. of Workers
Panel B. Characteristics of FinTechs	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All FinTechs	290	2008	72%	60%	19%	4.8	10.4	73.8
FinTechs that aggregate financial experts	31	2008	84%	48%	10%	4.0	11.9	17.5
FinTechs that aggregate financial news	242	2008	69%	66%	18%	5.3	11.9	85.7
FinTechs that crowdsource financial advice	45	2011	87%	47%	13%	6.2	18.1	13.9
FinTechs that datamine for financial signals	166	2007	63%	70%	17%	5.2	12.0	114.9
FinTechs that rank financial advice	77	2007	84%	60%	22%	3.5	7.5	53.1

Table IV. Are FinTechs and Financial Analysis Substitutes or Complements?

This table presents OLS estimates of investors discovery of original-content financial analysis when using FinTech websites. Column (1) examines the binary choice to visit a financial blog or FinTech website. Column (2) through (4) examine how visiting a FinTech website changes readership of financial blogs. The dependent variable in Column (2) is an indicator variable for visiting a financial blog. The dependent variable in Column (3) is the percent of page views at the financial blog and Column (4) is the percent of the internet users time spent at the financial blog. Coefficients on demographic variables should be interpreted as relative to the excluded category. For income, less than 50k is excluded. For race, other is excluded. For age, 18-29 is excluded. For education, high school degree or less is excluded. Additional control variables include census region, internet connection speed, and number of children. The data comes from comScore and tracks the internet usage of a set of households reflective of the U.S. population. For a detailed description of each variable, see the definitions in Appendix A.

	Dependent Variable =			
	Internet user visits a financial blog or FinTech website	Internet user visits a financial blog	Log of internet users page views at financial blogs	Log of internet users time spent at financial blogs
Determinants of financial blog use	(1)	(2)	(3)	(4)
FinTech Use				
Visits a FinTech website		-57.4***	-0.55***	-0.33***
		(0.10)	(0.00)	(0.00)
Income				
50-100k	1.15***	0.79***	0.05***	0.05***
	(0.03)	(0.05)	(0.00)	(0.00)
100k+	3.06***	2.04***	0.12***	0.10***
	(0.04)	(0.06)	(0.00)	(0.00)
Race				
White	6.83***	1.27***	0.20***	0.21***
	(0.04)	(0.06)	(0.00)	(0.00)
Black	1.89***	0.32***	0.03***	0.05***
	(0.05)	(0.08)	(0.00)	(0.01)
Asian	9.24***	3.01***	0.25***	0.23***
	(0.09)	(0.10)	(0.01)	(0.01)
Age of Head of Household				
30-39	0.28***	0.19**	0.04***	0.05***
	(0.05)	(0.08)	(0.00)	(0.00)
40-49	0.70***	0.43***	0.05***	0.06***
	(0.04)	(0.07)	(0.00)	(0.00)
50-59	1.74***	0.76***	0.11***	0.13***
	(0.05)	(0.07)	(0.00)	(0.00)
60+	3.66***	1.95***	0.23***	0.27***
	(0.05)	(0.08)	(0.00)	(0.00)
Education				
College degree	4.93***	0.88***	0.08***	0.07***
	(0.05)	(0.06)	(0.00)	(0.00)
Graduate degree	2.27***	0.42*	0.06***	0.06***
	(0.12)	(0.22)	(0.01)	(0.01)
Additional Control Variables	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Adjusted R-squared	2.4%	55.3%	6.5%	6.7%
Observations	7,241,817	1,090,746	1,090,746	1,090,746

Table V. Summary Statistics for Financial Analysts

This table presents summary statistics for the financial analyst sample. The sample is drawn from Zacks Investment Research. For a detailed description of each variable, see the definitions in Appendix A.

	Mean (1)	Median (2)	Std. Dev. (3)	Obs. (4)
Mean Bias (As % of the Absolute Value of Consensus EPS)	37.6%	5.6%	80.9%	91,871
Median Bias	36.5%	3.4%	85.2%	91,871
Mean Accuracy	61.9%	25.0%	77.6%	91,871
Median Accuracy	58.4%	20.2%	81.2%	91,871
Mean Bias (As % of the Previous Quarter's Stock Price)	1.6%	0.5%	3.6%	91,871
Median Bias	1.1%	0.2%	3.7%	91,871
Mean Accuracy	3.1%	1.4%	4.6%	91,871
Median Accuracy	2.2%	0.8%	4.6%	91,871
Analyst Coverage	7.0	5.3	5.6	91,871
Forecast Dispersion	0.7	0.5	0.7	91,871
Firm Size	13.9	13.8	1.8	91,871
Daily Return Volatility	40.5%	34.2%	23.9%	91,871
Mean Monthly Return	1.5%	1.4%	7.1%	91,871
Log Market-to-Book	0.81	0.74	0.44	91,871
Volatility of ROE	23.6%	0.2%	102.7%	91,871
Profitability	1.85%	2.27%	4.43%	91,871
Member of S&P 500	15.2%	0.0%	35.9%	91,871
Institutional Ownership	60.8%	67.8%	30.1%	91,871
Hedge Fund Ownership	9.7%	6.3%	10.1%	52,395

Table VI. OLS Regression of Consensus Analyst Bias and Accuracy

This table presents OLS estimates at the equity-quarter level for analysts' responses when financial bloggers concentrate in the equities they cover. In Columns (1) through (4), the dependent variable is analyst bias, defined as a consensus forecast bias of all analysts tracking stock i in quarter t. Forecast bias is the difference between the forecast of analyst j in quarter t and the actual EPS, expressed as a percentage of the consensus EPS. The consensus is obtained either as a mean as in Columns (1) and (2) or median as in Columns (3) and (4). In Columns (5) through (8), the dependent variable is analyst accuracy, defined as a consensus absolute forecast error of all analysts tracking stock i in quarter t. Below the coefficient estimates are robust standard errors clustered at the equity-level. ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

	Bias (As % of EPS)				Accuracy (As % of EPS)			
	Mean		Median		Mean		Median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial Blog Coverage	0.05***	0.01***	0.05***	0.01***	0.06***	0.01***	0.06***	0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Analyst Coverage	0.01*	0.09***	0.01*	0.08***	-0.02**	0.02*	-0.01*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm Size	-0.23***	-0.14***	-0.22***	-0.14***	-0.22***	-0.47***	-0.22***	-0.45***
	(0.01)	(0.03)	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.03)
Daily Return Volatility	0.23***	0.06***	0.23***	0.06***	0.27***	0.04***	0.27***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Mean Monthly Return	-0.03***	-0.02***	-0.03***	-0.02***	-0.00**	0.00***	-0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log Market-to-Book	0.12***	0.03*	0.12***	0.02*	0.07***	-0.00	0.08***	-0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Volatility of ROE	0.01*	-0.01	0.01	-0.02*	0.01*	-0.01	0.01*	-0.02
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Profitability	-0.42***	-0.15***	-0.40***	-0.14***	-0.40***	-0.17***	-0.38***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Member of S&P 500	0.02***	0.01	0.02***	0.01	0.01	0.02*	0.01**	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	36.2%	72.6%	33.6%	68.3%	38.7%	75.3%	36.4%	71.3%
Observations	91,871	91,871	91,871	91,871	91,871	91,871	91,871	91,871

Table VII. IV Regression of Consensus Analyst Bias and Accuracy

This table presents instrumental variable (IV) estimates at the equity-quarter level for analysts' responses when financial bloggers concentrate in the equities they cover. In Columns (1) through (4), the dependent variable is analyst bias, defined as a consensus forecast bias of all analysts tracking stock i in quarter t. Forecast bias is the difference between the forecast of analyst j in quarter t and the actual EPS, expressed as a percentage of the consensus EPS. The consensus is obtained either as a mean as in Columns (1) and (2) or median as in Columns (3) and (4). In Columns (5) through (8), the dependent variable is analyst accuracy, defined as a consensus absolute forecast error of all analysts tracking stock i in quarter t. Below the coefficient estimates are robust standard errors clustered at the equity-level. ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

	Bias (As % of EPS)				Accuracy (As % of EPS)			
	Mean		Median		Mean		Median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial Blog Coverage	0.52***	0.30***	0.50***	0.29***	0.64***	0.22***	0.62***	0.24***
	(0.10)	(0.07)	(0.10)	(0.07)	(0.10)	(0.06)	(0.10)	(0.06)
Analyst Coverage	-0.08***	-0.02	-0.08***	-0.02	-0.15***	-0.05**	-0.14***	-0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Firm Size	-0.36***	-0.14***	-0.35***	-0.14***	-0.38***	-0.47***	-0.38***	-0.45***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Daily Return Volatility	0.15***	0.04***	0.15***	0.04***	0.18***	0.03***	0.17***	0.03***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Mean Monthly Return	-0.03***	-0.02***	-0.03***	-0.02***	-0.00	0.00***	-0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log Market-to-Book	0.10***	0.00	0.09***	0.00	0.05***	-0.02	0.05***	-0.03*
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Volatility of ROE	0.00	-0.02*	0.00	-0.03**	0.00	-0.02	0.00	-0.02*
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Profitability	-0.42***	-0.15***	-0.40***	-0.15***	-0.40***	-0.17***	-0.39***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Member of S&P 500	-0.04**	-0.01	-0.03**	-0.01	-0.07***	-0.00	-0.06***	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
First Stage F-Stat	182.5	222.5	182.5	222.5	182.5	222.5	182.5	222.5
T-Stat on Instrument	13.51	13.51	13.51	13.51	13.51	13.51	13.51	13.51
Adjusted R^2	36.2%	72.6%	33.6%	68.3%	38.7%	75.3%	36.4%	71.3%
Observations	91,871	91,871	91,871	91,871	91,871	91,871	91,871	91,871

Table VIII. Analysts Reaction to High Quality Bloggers

This table presents instrumental variable (IV) estimates at the equity-quarter level for analysts' responses when high quality financial bloggers concentrate in the equities they cover. In Columns (1) and (2), the dependent variable is analyst bias, defined as the mean consensus forecast bias of all analysts tracking stock i in quarter t, expressed as a percentage of the consensus EPS. In Columns (3) and (4), the dependent variable is analyst bias, defined as the mean consensus forecast bias of all analysts tracking stock i in quarter t, expressed as a percentage of the previous quarter's stock price. The exact specification is: $ReportQuality_{it} = \alpha + \beta QualityBlogCoverage_{it} + \theta X_{it} + f_i + \delta_t + \epsilon_{it}$. The primary independent variable of interest is $QualityBlogCoverage_{it}$ which measures the quantity of financial blog posts identified by FinTechs as high quality in quarter t that discuss equity i. In Columns (1) and (3), quality is defined by short-term investment performance (i.e., less than six months) and in Columns (2) and (4) quality is defined by long-term investment performance (i.e., one year or more). We instrument for $QualityBlogCoverage_{it}$ using $PsychTrick_{it}$ which indicates the percent of newspaper headlines that covered equity i in quarter t that relied on psychological tricks for increasing attention. Below the coefficient estimates are robust standard errors clustered at the equity-level. ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

	Mean Bias (As % of EPS)		Mean Accuracy (As % of EPS)	
	Short-Term Positive Return Bloggers	Long-Term Positive Return Bloggers	Short-Term Positive Return Bloggers	Long-Term Positive Return Bloggers
	(1)	(2)	(3)	(4)
Financial Blog Coverage	0.33***	0.37***	0.24***	0.26***
	(0.07)	(0.08)	(0.07)	(0.08)
Analyst Coverage	-0.02	-0.00	-0.05**	-0.04*
	(0.03)	(0.03)	(0.03)	(0.02)
Firm Size	-0.11***	-0.13***	-0.45***	-0.46***
	(0.03)	(0.03)	(0.03)	(0.03)
Daily Return Volatility	0.04***	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Mean Monthly Return	-0.03***	-0.02***	0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Log Market-to-Book	0.00	0.00	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Volatility of ROE	-0.03*	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Profitability	-0.15***	-0.16***	-0.17***	-0.17***
	(0.01)	(0.01)	(0.01)	(0.01)
Member of S&P 500	-0.01	-0.02	0.00	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)
Time Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes
First Stage F-Stat	204.2	178.6	204.2	178.6
T-Stat on Instrument	13.76	11.97	13.76	11.97
Adjusted R^2	72.6%	72.6%	75.3%	75.3%
Observations	91,871	91,871	91,871	91,871

Table IX. What Economic Forces Drive Analysts' Responses to FinTechs?

This table presents instrumental variable (IV) estimates at the analyst-equity-quarter level. Column (1) repeats the previous anlaysis where the dependent variable is forecast bias at this more disaggregated analyst-level of the data. The remaining columns repeat the analysis for various subsamples of the data: Columns (2) and (3) focus on affiliated and non-affiliated stocks, Columns (4) and (5) focus on independent and non-independent brokerage houses, and Columns (6) and (7) focus on inexperienced and experienced analysts, respecitvely. Additional equity-level controls include firm size, daily return volatility, mean monthly returns, market-to-book, volatility of ROE, profitability, membership in the S&P 500, momentum, institutional ownership. Below the coefficient estimates are robust standard errors clustered at the equity-level. ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

	Dependent Variable = Mean Bias (As a % of EPS)						
	All	Affiliated Stock	Not Aff. Stock	Indep. Broker	Not Indep.	Inexp. Analyst	Exp. Analyst
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Financial Blog Coverage	0.32***	0.42***	0.29***	0.24***	0.32***	0.38***	0.21***
	(0.03)	(0.11)	(0.03)	(0.08)	(0.04)	(0.04)	(0.05)
Analyst Coverage	-0.06***	-0.08***	-0.05***	-0.04	-0.06***	-0.08***	-0.02
	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)
General Experience	-0.00**	-0.00	-0.00**	-0.00	-0.00	-0.00	-0.00**
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Firm Experience	-0.00***	-0.00	-0.00***	-0.00	-0.00***	0.002	-0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Firms Covered	0.00***	-0.00	0.00***	0.01**	0.00**	0.01***	0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Industries Covered	-0.00**	0.006	-0.00*	-0.01**	-0.00*	-0.02***	0.00**
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Forecast Frequency	0.00	-0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Forecast Horizon	0.00***	0.01***	0.00	0.00*	0.00**	0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Days Since Last Forecast	0.01***	0.00*	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Affiliated with Firm	0.00	N.A.	N.A.	0.02	0.00	0.01**	-0.0**
	(0.00)	N.A.	N.A.	(0.06)	(0.00)	(0.00)	(0.01)
Brokerage Size	0.00***	0.00*	0.00***	0.01***	0.00***	0.00***	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Independent Brokerage	0.01***	0.09*	0.00***	N.A.	N.A.	0.01***	0.02***
	(0.00)	(0.05)	(0.00)	N.A.	N.A.	(0.00)	(0.01)
Additional Equity-level Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage F-Stat	820.8	98.6	730.9	78.9	733.5	516.6	234.1
T-Stat on Instrument	29.88	11.66	27.31	9.30	27.90	23.60	16.97
Adjusted R^2	68.2%	76.2%	67.0%	75.1%	68.1%	68.5%	70.0%
Observations	296,743	45,357	251,264	40,505	255,964	189,550	106,927

Table X. Change in Price Informativeness With FinTechs

This table presents estimates of the change in price informativeness for stocks where FinTechs concentrate. In Columns (1) through (4), the dependent variable is price nonsynchronicity. In Columns (1) and (2), the independent variable of interest is FinTech concentration. In Columns (3) and (4), the independent variable of interest measures the quantity of financial blog posts identified by FinTechs as high quality in quarter t that discuss equity i. Additional equity-level controls include firm size, daily return volatility, mean monthly returns, market-to-book, volatility of ROE, profitability, membership in the S&P 500, momentum, institutional ownership. The ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

	Dep. Var. = Price Informativeness			
	(1)	(2)	(3)	(4)
FinTech Concentration	-7.86%***	-4.49%**		
	(0.02)	(0.02)		
High Quality Bloggers			-3.80%**	-1.86%
			(0.02)	(0.02)
Time Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes
Additional Controls	No	Yes	No	Yes
Adjusted R-squared	9.8%	11.0%	9.4%	10.1%
Observations	79,675	79,675	79,675	79,675

Table XI. Market Reaction to Analyst Recommendations With FinTechs

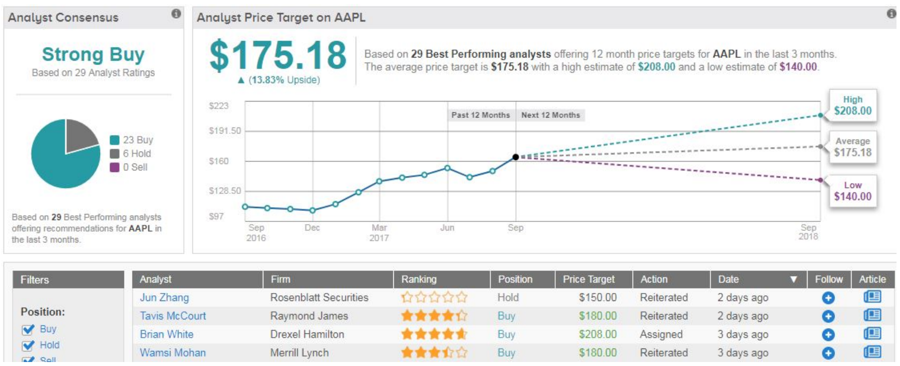
This table presents cumulative abnormal returns following recommendation revisions. Each revision is characterized as an upgrade to a buy or a dowgrade to a sell by comparing the revsied recommendation with the previous active recommendation for the stock by the revising analyst. Within upgrades and downgrades, we further classify them into revisions for equities with above mean concentration by FinTechs and equities with below mean concentration by FinTechs. Abnormal returns are in excess of benchmark portfolios matched on size, book-to-market, and momentum. ***, ** and * indicate p-values of 1%, 5%, and 10%, respectively. For a detailed description of each variable, see the definitions in Appendix A.

Panel A. Upgrades to Buy Recommendation	Obs.	[0,1]	[0,5]	[0,21]
High FinTech Concenation	7,291	1.42%	1.50%	1.63%
Low FinTech Concentation	27,901	1.91%	2.35%	2.31%
High-Low		-0.49%***	-0.84%***	-0.67%***
T-stat		(7.12)	(7.37)	(7.12)

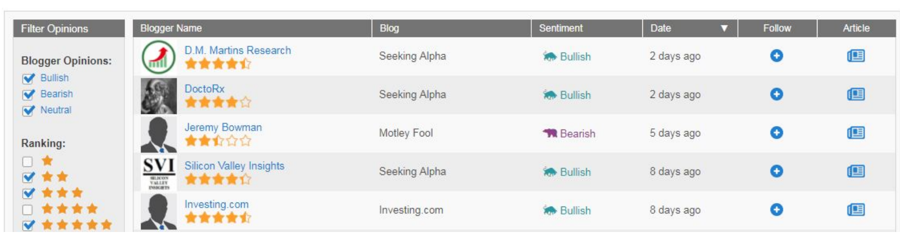
Panel B. Downgrades to Sell Recommendation	Obs.	[0,1]	[0,5]	[0,21]
High FinTech Concenation	1,218	-3.10%	-3.76%	-3.75%
Low FinTech Concentation	3,654	-3.07%	-3.78%	-3.54%
High-Low		-0.03%	0.02%	-0.21%
T-stat		(0.57)	(0.04)	(0.57)

Appendix A. FinTechs as Complements or Substitutes

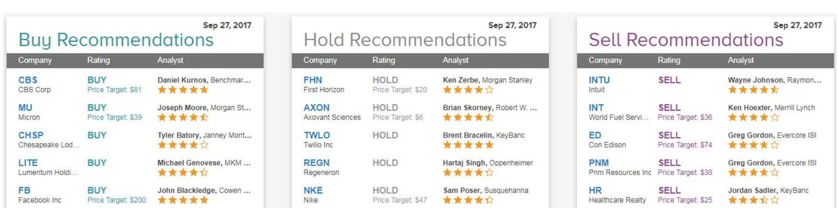
FinTechs are a complement financial analysts when investors learn the best analysts and click-through to their research. The graphic below shows a website where analysts are ranked using a star rating and investors can click through to read the original article.



FinTechs are a substitute for financial analysts when investors learn the best analysis is from bloggers so skip analysts' research. The graphic below shows a website where bloggers are ranked using a star rating in a manner that leads to easy comparison with analysts.



FinTechs are also a substitute for financial analysts when investors rely only on the signal and forgo reading original-content financial analysis altogether. The graphic below shows a website where clicking-through isn't an option rather only the extracted signal is displayed.



Appendix B. Variable Definitions

We use data from IBES, Zacks, CRSP, Compustat, and Thomson Reuters to construct our financial analyst sample. To construct our various measures of accuracy and bias, we use diluted, U.S. currency quarterly earnings per share (EPS) forecasts from 1 to 8 quarters out as well as diluted, U.S. currency annual EPS forecasts from 1 to 2 years out. The remaining EPS forecasts that are greater than 2 years out or more than 8 quarters out represent less than 2% of the universe of forecasts and are not well populated to evaluate the consensus; hence, this is our reason for excluded them. We include in our set of forecasts those that are original forecasts, announced confirmations of previous forecasts, and revised forecasts. Each variable is winsorized at the 1st and 99th percentile to mitigate the influence of extreme observations. Definitions are as follows:

Mean (Median) Bias As a Percentage of the Absolute Value of Consensus EPS is the difference between the analyst's forecast and the actual EPS divided by the absolute value of the consensus EPS for equity i in quarter t. Because our analysis is conducted at the equity level, we further aggregate forecast biases and consider the consensus bias expressed as the mean (median) bias among all analysts covering a particular equity.

Mean (Median) Accuracy As a Percentage of the Absolute Value of Consensus EPS is the absolute value of the signed forecast error (i.e., the difference between the analyst's forecast and the actual EPS) divided by the absolute value of the consensus EPS for equity i in quarter t. Because our analysis is conducted at the equity level, we further aggregate forecast biases and consider the consensus bias expressed as the mean (median) forecast error among all analysts covering a particular equity.

Mean (Median) Bias As a Percentage of the Previous Quarter's Stock Price is the difference between the analyst's forecast and the actual EPS divided by the closing price for equity i in quarter $t-1$. To match the definition of bias used in Hong and Kacperczyk (2010), we use EPS from Compustat rather than IBES. Because our analysis is conducted at the equity level, we further aggregate forecast biases and consider the consensus bias expressed as the mean (median) bias among all analysts covering a particular equity.

Mean (Median) Accuracy As a Percentage of the Previous Quarter's Stock Priceis the absolute value of the signed forecast error (i.e., the difference between the analyst's forecast and the actual EPS) divided by the closing price for equity i in quarter $t-1$. To match the definition of accuracy used in Hong and Kacperczyk (2010), we use EPS from Compustat rather than IBES. Because our analysis is conducted at the equity level, we further aggregate forecast biases and consider the consensus bias expressed as the mean (median) forecast error among all analysts covering a particular equity.

Analyst Coverage is the number of analysts covering stock i in quarter t. ($NUMEST$)

Forecast Dispersion is the standard deviation of all analyst forecasts covering stock i in quarter t. ($VALUE$)

Firm Size is the logarithm of stock i's market capitalization at the end of quarter t. ($\log(PRCC\_F \times CSHO)$)

Daily Return Volatility is the annualized variance of daily raw returns of stock i in quarter t. ($\sigma_{RET} \times \sqrt{252}$).

Mean Monthly Return is the average monthly return on stock i in quarter t. ($R\bar{E}T$)

Log Market-to-book $= log(\frac{PRCC\_F \times CSHO + DLC + DLTT + PSTKL - TXDITC}{AT})$

Return on Equity (ROE) $= \frac{NI}{SEQ_{t-1}}$

Volatility of ROE comes from estimating an AR(1) model for each equity's ROE using a rolling, 10-year series of the company's valid annual ROEs. The variance of the residuals from this regression is the volatility of ROE.

Profitability $= \frac{OIBDP}{AT}$

Member of S&P 500 is an indicator variable that takes the value of one if stock i is included in the S&P 500 index in quarter t.

Institutional Ownership data comes from Thomson-Reuters via 13F SEC filings. Ownership percentages are based on the number of shares outstanding and correspond to calendar dates.

Hedge Fund Ownership data comes from Factset and we use the classification technique created by Ferreira and Matos (2008). (IO_CAT6)

Affiliated Analyst is an indicator variable for if an analyst works at a brokerage house with a pre-existing relationship with the firm through business underwriting an IPO, SEO, or as an advisor on an M&A deal.

All-star Analyst is an indicator variable for if an analyst was listed in the October issue of Institutional Investor as an all-star analyst.

Brokerage Size is the number of analysts at the brokerage firm.

Brokerage Prestige is an indicator variable that takes the value of one if the brokerage firm is listed that year as one of Institutional Investor Magazine's top brokerage houses.

Firm Experience is the number of years analyst j covered stock i.

General Experience is the number of years since the analyst first appeared in the Zacks database.

Number of Firms Covered is the total number of unique stocks covered by the analyst during the year.

Number of Industries Covered is the total number of unique 2-digit SIC industries covered by the analyst during the year.

Days Since Last Forecast is the average number of days elapsed since the most recent forecast for that same stock by i by analyst j in a given quarter t.

Forecast Horizon is the average number of days between the estimate date and the reference date, which is the fiscal period end date, in a given quarter t for a stock i covered by analyst j.

Forecast Frequency is the number of forecasts for stock i issued by analyst j during the previous year.

Price nonsynchronicity is estimated as $1 - R^2$, where R^2 is the R-square from the following regression: $r_{ijt} = \beta_{i,0} + \beta_{i,m} * r_{m,t} + \beta_{i,j} * r_{j,t} + \epsilon_{i,t}$ where $r_{i,j,t}$ is the return of firm i in industry j at time t, $r_{m,t}$ is the CRSP value-weighted market return at time t, and $r_{j,t}$ is the return of 3-digit SIC industry j at time t. If there are fewer than 30 daily price observations in a quarter the observation is set as missing.

To construct our dataset of FinTech firms and financial blogs, we use data provided to us by TipRanks. We supplement this data with data from Crunchbase, ComScore, and internet searches.

Definitions are as follows:

Year Founded is pulled from Crunchbase. If it is not available on Crunchbase, founding date is pulled from the FinTech's website. If the founding date is not on Crunchbase or the FinTech's website, then the first year in which Wayback Machine made a copy of the website is used as the founding year.

Targets Retail Investors is an indicator variable equal to one if the FinTech's business plan suggests the product is meant for retail investors.

Targets Professional Investors is an indicator variable equal to one if the FinTech's business plan suggests that the product is meant for insitutional investors.