

Market Efficiency in the Age of Machine Learning*

Leonidas G. Barbopoulos[†]

Rui Dai[‡]

Tālis J. Putniņš[§]

Anthony Saunders[¶]

This version
September 15, 2021

*We are grateful to comments and suggestions from Simona Abis (EFA discussant), Yakov Amihud, Gjergji Cici, Ranjan Chakravarty, Jo Danbolt, Angelica Gonzalez, Alistair Haig, Pia Helbing, Pantelis Kazakis, Andrew Marshall, Georgios Panos, Leung Woon Sau, Alexi Savov, Georgios Sermpinis, Charalampos Stasinakis, and David Yermack and participants in the seminar series on NYU Stern School of Business (Finance), Franklin Templeton Investments (Virtual / San Mateo, CA), the World Finance Conference (Virtual / Kristiansand, Norway, August 2021), and the European Finance Association (Virtual / Milan, Italy, August 2021). We are responsible for any remaining errors or omissions.

[†]University of Edinburgh Business School, The University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK. Email: leonidas.barbopoulos@ed.ac.uk. Tel: +44 131 650 8344. Corresponding author.

[‡]Wharton Research Data Services (WRDS), The Wharton School, University of Pennsylvania, 3819 Chestnut Street Philadelphia PA, 19104. Email: rdai@wharton.upenn.edu.

[§]University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia and University of Edinburgh Business School, The University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK. Email: Talis.Putnins@uts.edu.au.

[¶]Department of Finance, Kaufman Management Center, Stern School of Business, New York University, New York, N.Y. 10012. Email: asaunder@stern.nyu.edu.

Market Efficiency in the Age of Machine Learning

Abstract

As machines replace humans in financial markets, how is informational efficiency impacted? We shed light on this issue by exploiting a unique data-set that allows us to identify when machines access important company information (8-K filings) versus when humans access the same information. We find that increased information access by cloud computing services significantly improves informational efficiency and reduces the price drift following information events. We address identification through exogenous power and cloud outages, a quasi-natural experiment, and instrumental variables. We show that machines are better able to handle linguistically complex filings, less susceptible to bias from negative sentiment and less constrained in attention/processing capacity than humans.

Keywords: Market efficiency; Information acquisition; Machine learning; Informed trading; Algorithmic trading.

JEL Classification Codes: G10; G12; G14.

“Thirty years ago the best fund manager was the one with the best intuition... Now those who take a “scientific approach”, using machines, data and AI, can have an edge.”

David Siegel, co-chairman of Two Sigma (*The Economist*, October 5th, 2019)

1. Introduction

As machines replace humans in financial markets, how is informational efficiency impacted? One possibility is that human biases are reduced, information is processed faster and in larger volumes, leading to improved price discovery and higher informational efficiency. Yet another possibility is that as machines replace humans, some of the “soft” information that humans (but not machines) are able to interpret is lost and consequently prices become less informative. While humans make a range of errors in interpreting information, machines are also not infallible, and they too can make different errors. Which of these opposing effects dominates is an empirical question that has broad implications for the functioning of financial markets and ultimately the efficient allocation of resources and risk.¹

We analyze the impact of machines (i.e., robots, algorithms), and in particular cloud computing services², on informational efficiency using a unique data-set in which we can distinguish between when a machine reads information released by a company versus when a human accesses the same information. We use the SEC server log files, which record viewership of corporate events disclosed in Form 8-K filings via the EDGAR system. The Form 8-K filings are the means through which companies notify the market of important unscheduled corporate events including changes to a material agreement, financial information, acquisitions, substantial impairments, and any other events deemed important to shareholders. What distinguishes our paper from other studies of these company filings is that we separate each viewership of an 8-K filing into machine viewers and human viewers based on

¹This question has also attracted significant attention from the business press. In an article appeared in the *Briefing* section of *The Economist* on October 5th 2019, stated that “The rise of financial robotisation is not only changing the speed and makeup of the stockmarket. It also raises questions about the function of markets, the impact of markets on the wider economy, how companies are governed and financial stability.”

²According to PwC, in 2020 cloud computing services have been named among the biggest fintech trends with public cloud (a virtual space on common hardware) poised to becoming the dominant infrastructure model.

the downloading behavior captured by the server logs.³ We further partition viewership at the level of the viewer's computing facility using IP addresses. No prior study to the best of our knowledge has examined material information acquisition at this granular level.

First, we show that cloud computing viewership of 8-K filings has experienced significant growth in recent years.⁴ Specifically, the viewership from cloud computing services, such as Amazon Web Services (AWS), was only 1% of the total machine viewership before 2008, and the same figure increased to 14% in 2008, to 35% in 2012, and to 62% in 2016. Similarly, the total machine viewership has increased from 0.57 million views in 2008 to 8.78 million views in 2016, while total human viewership remains stable with an average of 0.3 million views per year. This rapid growth coincides with the increase in popularity of both quantitative investment strategies, whose trading volume grew from 20% to 36% of institutional volume between 2014 and 2019, and cloud computing services, which offer many unique benefits to machine-based investment strategies, such as anonymity, stability, and minimum on-site maintenance.⁵

To understand the differences between how machines and humans access information, we compare their determinants. Both human and machine viewers pay more attention to timely and complex information. Where humans and machines differ is that humans show a tendency to disproportionately allocate their attention to large firms and value firms, consistent with a need to ration attention due to cognitive and capacity constraints. In contrast, machines more uniformly access different types of information across different segments of the market, showing no significant preference for the type of firms, consistent a greater information processing capacity. Importantly, human viewers have a tendency for increased viewership of 8-K filings that have negative sentiment content, but machines are indifferent to the sentiment content of the 8-K.

A noticeable difference between machines and humans that helps explain our findings is the

³Following the previous literature (Drake et al., 2015; Dechow et al., 2015), we identify information acquisition by a machine when the acquiring entity consistently downloads a large volume of 8-K filings beyond human comprehension within a short period of time.

⁴An article in *The Economist* on December 21, 2020, states that “For most of the past decade more trades have been done at high frequency by complex algorithms than by humans.”

⁵Due to technological advances and market structure developments in recent decades, computer-based trading including algorithmic and high frequency trading, account for the majority of trading in current financial markets. Along similar lines, Hendershott et al. (2011) reported that 75% of trading volume in the US in 2009 was conducted through computer-based trading.

type of information that they choose to focus on. A useful feature of the 8-K filings is that they arrange events into topical categories. We find that while human viewers pay significantly more attention to earnings information (item 2.02), machine viewers pay significantly more attention to corporate events not listed by the SEC but deemed material non-public information by the firm (item 8.01). These results are consistent with the notion that the greater computational capacity of machine-based strategies allow them to digest a broader set of unanticipated but important events that human viewers may neglect due to limited attention or limited cognitive capacity.

To quantify the impact of machines on informational efficiency, we measure the magnitude of absolute price drift following the 8-K publication date over horizons of 10, 20, and 40 days and relate the magnitude of the drift to human and machine viewership. Our key finding is that more viewership by cloud-based machines tends to decrease the price drift, consistent with more efficient price discovery around the information release. In contrast, relatively more human viewership of an 8-K filing is significantly associated with increased price drift following the 8-K publication date across various windows, consistent with an inefficient assimilation of the information. This evidence is robust in all three event windows, with the inclusion of various controls and fixed effects, and remains qualitatively similar with alternative ways to define drifts. The results are also robust to an alternative measure of informational efficiency based on a variance decomposition that separates noise and information. In terms of economic significance, a one-standard-deviation increase in machine viewership from cloud services leads to a 10.49% reduction in price drift over the month after 8-K filing date whereas a one-standard-deviation increase in the human viewership (holding fixed the level of machine viewership) yields a 13.23% aggregate increase in the price drift.

We find that not all machines are beneficial to informational efficiency. We partition machine viewership (using IP addresses) into six categories: cloud computing services, financial institutions, data & media publishers, auditing & law firms, internet services providers, and a diverse group of other entities not related to the investment industry. This allows us to separate machine viewership that is unrelated machine-based investment strategies. The beneficial effects of machines on price discovery stem almost entirely from the cloud computing services, which are most likely to be associated with machine-based investment strategies and capture a class of sophisticated quantitative

investors that require high computational power. Interestingly, machine viewership from financial institutions is not significantly related to post-event price drift, suggesting that the more sophisticated machine-driven investment strategies tend to be deployed via cloud computing servers and not via IP addresses known to belong to financial firms.

To examine causality in these effects, we use two main identification strategies: exogenous cloud and power outages, and a quasi-natural experiment. We also consider evidence from instrumental variables. Our first approach exploits exogenous disruptions to machines from cloud service and major power outage events. Our second approach uses the firm's inclusion in the S&P500 index as a quasi-natural experiment because inclusion events lead to a sudden surge in viewership primarily from human analysts and fund managers who, unlike machines, have limited capacity to follow stocks outside of major indices (Farboodi et al., 2020). Our final approach is a two-stage least-squares (2SLS) estimation in which human viewership is instrumented by market sentiment and macroeconomic news, which disproportionately affect human ability to assimilate information. The results from all of these identification strategies support the notion that cloud computing machines cause an improvement in informational efficiency.

What is it that machines do better than humans to result in improved informational efficiency? We test two potential mechanisms. First, we examine whether the positive effects of cloud computing machines are due to better information processing (machines making informed decisions) or faster information processing, or both. We find that machine viewership is significantly positively related to measures of informed trading following the 8-K publication date, whereas for humans we get the opposite result. Second, our analysis suggests that the cloud-machine viewership of 8-Ks is significantly associated with increased algorithmic trading activity around the information release, based on odd lot and trade size proxies. Therefore, the evidence is consistent with machines having an advantage in both the accuracy of information processing and speed. Machines that access information, use the information in investment decision models, leading to trading by machines, and ultimately this trading is what drives the information to be impounded in prices.

Finally, we turn to the issue of in what settings do humans vs. machines have an advantage? While machines are faster and better able to process large volumes of information than humans,

their advantage is unlikely to be universal, and in some setting humans may still have the upper hand. We find that the advantage of machines increases for information that is difficult for humans to read (linguistically complex filings). We also find that cloud-based machines are more effective in processing information that has negative sentiment content or tone compared to humans, consistent with the notion that machines are less susceptible to bias from emotions. For example, [Tetlock \(2007\)](#) documents that media pessimism leads to inefficient price reactions as human traders show behaviors expected of noise traders and our evidence suggests that machines are not as vulnerable to such bias. We also find that humans benefit when information is repeated, consistent with attention and capacity constraints among humans that result in some humans missing information in the first instance. In contrast, repeat information has no significant effect on machines as they are more likely to process information when it first arrives. Our findings are also consistent with theory that humans specialise in slower decision making that combines sequential information whereas machines have an edge in making quick decisions on isolated information.

Our paper makes two main contributions to the information disclosure literature. Prior research on information acquisition via SEC filings considers human heterogeneity and how it affects the informational efficiency of prices ([Drake et al., 2015](#); [Dyer, 2019](#)). Another strand of literature proposes various machine learning models to show that computationally-intensive quantitative strategies can generate superior performance compared to traditional asset pricing models ([Gu et al., 2020](#)). They do not, however, investigate how information-based quantitative investment strategies impact the price discovery process. We fill this void by examining the impact of machine viewership of information on the price discovery process and find that quantitative investment strategies, especially those implemented through cloud computing services, help improve market efficiency.

Our paper also contributes to the literature on price informativeness by examining the specific mechanism of information acquisition by different types of investors in a casual setup. Prior studies suggest that sophisticated investors improve stock price efficiency ([Akbas et al., 2015](#); [Kokkonen and Suominen, 2015](#); [Cao et al., 2018](#)). Recent literature focuses on investors' capability of handling information and examines their impact on market efficiency. [Chen et al. \(2020b\)](#), for example, document that aggressive trades by hedge funds, based on acquired information from the SEC's

EDGAR server, mitigates the impairment of market efficiency caused by analyst coverage reductions. [Begenau et al. \(2018\)](#) argue that certain investors can process larger data more effectively and therefore help price assets of firms with big data more accurately. Our paper expands this line of research by showing that information acquisition by machine-based investment strategies impacts post-announcement price drift differently from information acquisition by humans.

2. Theoretical Framework

Several branches of the literature help understand how machines could impact informational efficiency, the channels through which those effects occur, and where machines or humans have an advantage. The related theory includes rational expectations models of endogenous information acquisition through to recent models of competing algorithmic investors and traders. Also relevant are studies that analyze the relative efficacy of machine learning methods in assimilating information about stocks.

2.1. How do machines impact informational efficiency?

A key insight from rational expectations models is that when information acquisition and processing is costly, stock prices only partially reflect the available information and hence investor efforts in becoming informed are compensated ([Grossman and Stiglitz, 1980](#); [Kyle, 1985, 1989](#); [Admati and Pfleiderer, 1988](#)).⁶ A direct consequence is that a reduction in the cost of information acquisition and processing will tend to increase the amount of information acquisition and improve stock price informativeness. Advances in computing and automation reduce the marginal costs of information acquisition and processing. Therefore, the first effect we expect as machines become more prevalent in investment settings is more information acquisition, likely accompanied by better informational efficiency of prices.

Recent studies of the role of machines and machine learning in financial markets argue that the cost of analyzing information has declined as machines replace humans. For example, machines have

⁶[Sims \(2003, 2006\)](#) further argues that investors' limited information-processing capacity discourages the information assimilation process into stock prices.

greater economies of scale in processing information (Chen et al., 2020a). Increased computational power and the use of machine learning algorithms can allow investors to outperform traditional empirical asset pricing models (Gu et al., 2020), make use of highly unstructured textual data (Bybee et al., 2020), and front-run slower traders (Van Kervel and Menkveld, 2019).

A wide range of behavioural and information processing effects are also relevant in understanding the effects of replacing humans with machines in financial markets. Machine-based strategies that make investment decisions following rules extrapolated from data may be less susceptible to factors that bias human decision-making including emotions, heuristics, and impulsiveness. They are also likely to be faster than humans, especially for demanding and complex tasks (Sims, 2003, 2006; Bradshaw et al., 2020).⁷ Studies of heterogeneity among humans show that investors have different information-processing models and, in turn, respond to identical public signals differently (Xiong, 2013). As a result, human heterogeneity contributes noise into stock prices, lowering market efficiency. Humans also trade for non-investment rationales.⁸ Presumably due to these reasons, quantitative and machine-driven investment strategies have grown rapidly, accounting for approximately 36% of institutional stock trading activity and nearly 35% of US equity ownership in 2019.⁹

In sum, given the potential for machines and machine learning methods to reduce information acquisition and processing costs (Chen et al., 2020a) and produce faster, less biased trading decisions based on firm-specific information (Sims, 2003, 2006; Peng, 2005; Veldkamp, 2006), our *main hypothesis* is: *Machine viewership of company information (SEC's Form 8-K) increases informational efficiency (relative to human viewership) and decreases the price drift following the 8-K publication.*

2.2. Shortcomings of machines and potential negative effects

Machines, like humans, are not infallible. Practitioners and academics have expressed concerns that computationally intensive black-box algorithms may end up identifying spurious correlations or data

⁷Bradshaw et al. (2020) argue that soft news is relatively more costly to process than hard news and markets are likely to react more strongly to analysts' processing of soft news than of hard news as hard news is more likely to have already been factored into prices.

⁸For example, Barber and Odean (2002) argue that overconfident traders can cause markets to underreact to the informational content of rational traders. Another strand of studies has identified numerous behavioral rationales for over-trading, including entertainment (Dorn and Sengmueller, 2009), sensation seeking (Barber and Odean, 2008; Grinblatt and Keloharju, 2009), gambling (Kumar, 2009; Cookson, 2018), and learning by doing (Linnainmaa, 2011).

⁹Anonymous. "March of the machines. The stock market is now run by computers" *The Economist*, Oct 5th 2019.

biases, and under-performing traditional methods in the long-run.¹⁰ Critics have also argued that machine learning methods produce strategies that appear profitable, but load on difficult-to-arbitrage stocks and therefore often fail when trading costs and other frictions are taken into consideration (Avramov et al., 2020). Moreover, machines are limited in their capacity to process soft information, compared to humans (Bradshaw et al., 2020).

Dugast and Foucault (2018) show that there is a trade-off between using information to make quick and automated decisions versus making slower and more deliberated decisions. The quick and automated decisions (reflecting machine-driven investment) have the advantage of acting on information faster and thereby reaping the profits of acting ahead of slower investors, but at the cost of sometimes making substantial errors that are more severe than those of slower and more deliberated decision-making. The slower and more deliberated decision-making of humans can benefit from the use of soft information and from combining low and high precision signals received at different points in time. It is therefore possible that while machines may speed up price discovery, they also introduce new types of errors into the price discovery process.

Another potential downside, which has been explored in the algorithmic trading literature, such as Weller (2017), is that automated models may have an adverse effect on stock price informativeness, despite their importance for transmitting available information into prices. For example, Baldauf and Mollner (2020) find that faster speeds enable high-frequency traders (HFT) to be more successful at order anticipation, which can have a negative effect on information production due to informed traders having less time to trade before HFTs react. Similarly, Lee and Watts (2020) find that when the SEC's "Tick Size Pilot" increased the tick size and reduced algorithmic trading, pre-announcement stock returns are better able to predict the news of the upcoming earnings release, consistent with improved price discovery.¹¹ In our setting, it may be that machines beat humans in trading on information, reducing the profits of humans and thereby disincentivizing them from costly information acquisition and interpretation.

¹⁰For example, see Mark Hulbert (2020, Jan 5). Use AI for Picking Stocks? Not So Fast. *Wall Street Journal*. Retrieved from <http://online.wsj.com>.

¹¹On the contrary, Chordia and Miao (2020) use comprehensive intraday data and find that low-latency trading improves the long-term informational efficiency of stock prices.

Automated investment models that rely on fast processing of big data may have little or no impact on stocks with large quantities of data that are not machine-readable. For example, [Begenau et al. \(2018\)](#) investigate whether investors more efficiently price firms with big data. They find that big data disproportionately benefits large firms, allowing investors to produce more accurate forecasts and reduce uncertainty. In a similar spirit, [Farboodi et al. \(2020\)](#) find that the informativeness of stock prices varies significantly across groups of stocks, with large and growth stock prices increasingly reflecting information about future earnings.

Given that there are potentially both positive and negative effects of machines on informational efficiency, it is ultimately an empirical question of whether the positive or negative effects dominate and whether it depends on the context.

2.3. When do machines have an advantage over humans?

While there are many reasons why machines have an advantage over humans such as ability to quickly process large volumes of information, machines are unlikely to have an advantage over humans in all settings. The key question is in what settings do machines have an edge, and conversely, when do humans have an advantage over machines?

We begin with the issue of machine-readability. [Hwang and Kim \(2017\)](#) document that issuing financial disclosure documents with low readability causes firms to trade at significant discounts relative to their fundamentals. They argue that firms with difficult-to-read documents could evoke feelings of uncertainty and distrust among investors. In contrast, automated strategies lack such subjective components in their decision making processes. [Cao et al. \(2020\)](#) find that increased machine downloading activity motivates firms to prepare filings according to machines' readership and processing capability. This would further differentiate the impact of readability on the information assimilation speed of automated models from that of human readers. Therefore, we expect the positive effect of automated strategies on stock price efficiency to be pronounced when machines gather and analyze the information in 8-Ks that is difficult for humans to read.

Another strand of the literature suggests that investor sentiment affects stock prices. [Tetlock \(2007\)](#) uses the daily content from a popular *Wall Street Journal* column to proxy for investor

sentiment and finds that high media pessimism predicts short-run downward pressure in stock prices. [García \(2013\)](#) uses the [Loughran and McDonald \(2011\)](#) dictionaries to measure the sentiment of *New York Times* articles, and finds that sentiment predicts stock returns during recessions. While the preceding research demonstrates that human sentiment has an effect on stock prices, the question of whether machine-driven investment strategies are less susceptible to human sentiment remains an unresolved one. We anticipate that the positive impact of machines on stock price efficiency is stronger when the information disclosure contains negative sentiment or takes place on negative sentiment days, which is when human decisions are most likely to be biased by sentiment.

Lastly, recent studies have theoretically investigated whether lower information processing cost and increased information transparency leads to more precise signals. [Dugast and Foucault \(2018\)](#) demonstrates that decreasing costs of corporate disclosure gathering (by machines) can hamper price efficiency by lowering equilibrium demand for more precise fundamental signals. In other words, this model implies that humans have an advantage over machines when performing slower but more in-depth analyses of sequential information pieces. Hence, we expect that machines will not improve stock price efficiency or could even harm it when it comes to the sequential arrival of information that lends itself to a deeper analysis by human experts.

3. Data and Summary Statistics

We first describe our data sources and the variables used in our analysis. We then provide an overview of the cloud computing services that have revolutionised modern financial systems. Lastly, we provide descriptive statistics of the key variables.

3.1. Data sources and identifying machines vs humans

Our sample is made up of all firms in the CRSP and Compustat universe during a 14-year period from January 2003 to December 2016.¹² We collect data on each viewing (referred to as a visit) of each 8-K filing (a notification to investors of an unscheduled material or extraordinary event that is

¹²Up to writing this article, [EDGAR log file data set](#) is available from 2003 to the first half of 2017. We select the log records from January 2003 to December 2016 to allow for 14 whole years of observations.

important to shareholders) from the SEC's EDGAR server, giving us nearly 4 billion visits.¹³ Each observation contains a partially anonymized IP address, time stamp, HTTP status codes (e.g., 200 for successful delivery), and crawler flag.¹⁴ We obtain 8-K filings from WRDS SEC Analytics Suite (approximately 1.2 million 8-K filings). We restrict our sample to the filings that match firms in the CRSP and Compustat universe through the CIK-GVKEY link provided by WRDS SEC-Suite.

To protect the privacy and intellectual property of the filing viewers, the SEC redacts the last octet in their IP addresses. Similar to [Chen et al. \(2020a\)](#), we match IP addresses to organizations using a dataset from MaxMind by the first three octets in the IP address. In our sample, the proportion of the IP addresses that share the same first three IP octets, but come from different organizations, is small (3.46%).¹⁵ When multiple organizations are assigned to a redacted IP address, we use the organizations associated with most IP addresses within the redacted octet for the IP address.

Next, we categorize the organizations of 8-K viewers into seven groups using data from Thomson Reuters Global Ownership and Capital IQ. The former data set contains approximately 1.6 million (financial) institutional investors and investment vehicles, such as mutual funds and hedge funds. The latter contains approximately 14.2 million public and private firms, as well as their subsidiaries. We label institutional investors and funds as "Institutional Investors" based on industry type from both databases, while we group the remaining organizations into "Auditing & Law Firms", "Clouding Services", "Media & Data Vendors", "Educations & Regulators", "Internet Services Providers" and "Others".¹⁶

¹³To be precise, we have 3,966,935,088 visits to 1,203,881 8-K filings issued between January 2003 and December 2016. We remove the visits that fail to retrieve documents from EDGAR server (HTTP code \neq 200), only visit index page (index dummy=1) that come from search engines, and other similar web crawlers (crawler=1).

¹⁴The raw files also include filing-specific information: Central Index Key (CIK) used in SEC's EDGAR server to identify filers, accession numbers that uniquely match a specific SEC filing, and files visited (e.g., exhibits or index file).

¹⁵Each redacted octet comprises 256 distinct IP addresses, ranging from 123.123.123.0 to 123.123.123.255, for example.

¹⁶To facilitate the matching between the MaxMind's organization and entity names and the other two data sets, we use a fuzzy-name matching algorithm developed by WRDS, which involves human validation through Amazon's Mechanical Turk. To separate Internet Service Provider and Cloud Computing from other technology firms (in order to identify unmatched entities), we first rely on Google to obtain the URL and short text results in the top 10 hits. We only search the organizations that have visited the EDGAR server for 8-K at least 100 times in our sample period. For the ones we cannot identify, we rely on Wikipedia and organizations' website that is obtained from the first 10 Google search hits. Our manual search allows us to reduce the visits from IP addresses from entities categorized as "Others". The fine-tuned "Others" category contains entities from IP addresses that remain unidentified after manual searching, as well as the industries outside the remaining six categories. These exclusions account to be less than 3.86% of total 8-K visits in our final sample. The Appendix Table B reports top 10 Organizations of "Cloud Computing",

We use two approaches to separate human viewership from machine viewership. We first follow [Drake et al. \(2015\)](#) to define visits from the IP addresses with any downloads per minute greater than 5, or IP addresses with more than 1,000 downloads per day as machine visits. Then, we follow [Dechow et al. \(2015\)](#) to count the visits from IP addresses with any downloads per minute greater than 25, or with the number of CIK's downloaded per minute being greater than 3, or with more than 500 downloads during the day as machine visits.

For each 8-K filing in our sample, we obtain the report date (Conformed Period of Report), the filing date, the SEC release date, the item type, and the text of the 8-K filing, including the text from any attached exhibits (Item 9.01). Almost all 8-K filings must be reported within *four* business days of the event date, with a very small number of exceptions ([Ben-Rephael et al., 2020](#)). To focus on the 8-K filings that carry new information, we limit the sample to filings that contain events that are released within four business days from the event date.¹⁷

We also use a number of other data sources. We obtain stock price data from CRSP and fundamental data from Compustat. We use I/B/E/S to for analyst coverage, and 13F data from Refinitiv for ownership. We obtain the S&P500 inclusion dates from CRSP, quarterly earnings announcement dates from Compustat, cloud service disruption event data from [Gunawi et al. \(2016\)](#) and power outage event data from [Mukherjee et al. \(2018\)](#).¹⁸ To calculate the informed trading and algorithmic trading measures, we use NYSE Trade and Quote (TAQ) and the SEC's Market Information Data and Analytics System (MIDAS) available through WRDS. All variables are defined in the Appendix Table A.

“Institutional Investors”, “Internet Provider Services”, and “Others”.

¹⁷On August 23, 2004, the SEC redefined qualifying events that trigger 8-K filers and accelerated the 8-K publication date. For more information about this rule see the SEC Financial Reporting Releases Nos. 33-8400 and 34-49424. To ensure data consistency, we match the items defined in the previous version of 8-K filings to the ones in the post-2004 version over our sample period.

¹⁸[Gunawi et al. \(2016\)](#) collects the cloud service outage from 2009 to 2015. We follow their method to obtain complementary cloud service outage events from 2003 through to 2016. Besides, we drop the outage events cloud service related to gaming (e.g., Play Station), business storage (e.g., Salesforce), and social media (e.g., Facebook). We also manually collect the major power outage events in the second half of 2016 from Google to complement [Mukherjee et al. \(2018\)](#), which provides the outage data from January 2000 to July 2016.

3.2. Construction of efficiency measures

The absolute cumulative abnormal return (ACAR) of [Ball and Brown \(1968\)](#) and [Fama et al. \(1969\)](#) is a standard measure of the incorporation of information into prices. This measure constructs post-event price drift net of a predicted returns from a factor model, for example,

$$CAR_{i,t}^{0,T} = \sum_{t=0}^T \left(r_{i,t} - \alpha_i - \sum_{k=1}^k \beta_{i,k} f_{m,t} \right) = \sum_{t=0}^T \varepsilon_{i,t} \quad (1)$$

where $r_{i,t}$ is the raw return of stock i on date t and α_i and $\beta_{i,k}$ are estimated from a regression of k mimicking portfolios. The cumulative abnormal return (CAR) cumulates the abnormal return from announcement date $t = 0$ to T , and the *ACAR* is the absolute value of this number. The $ACAR(0, T)$ measures the announcement information that enters prices through the announcement day to the day T . To measure the drift components of the information integrated into price, we proceed along the lines of [Meulbroek \(1992\)](#) and take the difference between total post-event cumulative abnormal return and those over a short window post to 8-K announcement. Specifically, we use the absolute difference between post-announcement price variation and variation over a short window right after the announcement:

$$DRIFT(2, T) = \left| CAR_{i,t}^{0,T} - CAR_{i,t}^{0,1} \right| \quad (2)$$

with $T > 1$ to capture post announcement drift.

For robustness, we also use an alternative measure of inefficiency in how information is incorporated into prices: the noise return variance from [Brogaard et al. \(2020\)](#). This measure is based on a variance decomposition model that partitions stock return variance into information of various types (e.g., market-wide, stock-specific private information, and stock-specific public information) and noise (temporary price deviations from efficient prices, driven by under or over reaction).

3.3. Summary statistics

Table [\(1\)](#) provides the number of observations, mean, standard deviation, median, and the 25th and 75th percentiles of our key variables. To focus on the immediate viewership right after the 8-K announcement date, we restrict our sample to the visits on the SEC's 8-K filing day and the day following ($t \in \{0, 1\}$). This leaves us with 547,883 8-K filings with all main control variables available.

Among other notable statistics reported in Panel A, machines view more 8-K filings (log viewership of 3.30 and 3.35, based on Dechow et al. (2015) and Drake et al. (2015) methods, respectively), on average, compared to humans (log viewership of 1.35 and 1.05, based on Dechow et al. (2015) and Drake et al. (2015) methods, respectively). The median and standard deviation of both groups indicate that machines are far more active in viewing 8-Ks than humans. Among other statistics reported in Panel A, the average firm size is about \$6.5m while the average book-to-market ratio is 0.75. The average institutional holding is 61%, with 4.1 analysts covering the firm, on average. Panel B reports viewership broken down into six item types that appear in 8-K filings. The most frequently viewed is item 2.02, which is largely related to quarterly and annual earnings announcements. The mean (median) 2-day viewership of a filing with item 2.02 is 69.91 (41) views.

Panel C reports viewership broken down by the viewer type using seven categories defined in Section (3.1). Viewership from “Cloud Computing” and “Internet Provider” are considerably greater than those from the rest organization types. The proportion of machine viewership from Cloud Computing is 98.33%, the highest among all seven categories.¹⁹ The total viewership per organization is heavily skewed. For example, the average number of 8-K views per Cloud Computing organization is 13,780, while the median is only 10 views with a standard deviation of 204,000 across our entire sample period. This indicates the 8-K viewership is highly concentrated – there is a small number of highly active viewers.

4. What Information is Accessed by Machines vs Humans?

First we examine how machine and human access to information has changed through time. Next we look into the cross-section of stocks and information types and ask whether machines have similar patterns of information access as humans?

4.1. Human and machine information access through time

Figure (1) plots annual number of visits in our sample by machine and human viewers, as well as the total number of visits. There appears to be exponential growth in machine viewership of 8-K

¹⁹The “Other” category has the lowest proportion of robot viewers, roughly at 46.02%.

filings since 2007. In contrast, human viewership has remained at relatively low levels compared to machines through time. Therefore, while machines now account for the vast majority of information access, they have increased their dominance compared to humans through time.

Cloud computing services, such as Amazon Web Services, enable largely anonymous, on-demand network access to computing services over the Internet. These services include data backup, disaster recovery, machine learning, and big data analytics. Cloud computing services can protect users' identities, avoid internet service disruption, and access servers from the desired geographic locations. Therefore, cloud computing services are appealing to various financial institutions that otherwise need to maintain an in-house computing center. The technology hurdle to deploying investment strategies through cloud computing and the cost associated with its features mentioned above establish barriers to entry for less sophisticated investors to benefit from cloud computing.²⁰

The trends in viewership of 8-K filings by cloud computing services are consistent with the overall trends of cloud computing development in recent years. As illustrated in Figure (2), cloud computing viewership represented only 1% of the machine viewership before 2008, 14% in 2008, 35% in 2012, and 62% in 2016. Therefore, cloud computing is an important means of information acquisition and processing, which is expected to directly affect market efficiency.

Abis and Veldkamp (2021) show that during the later part of this sample period there was strong growth in the use of artificial intelligence (AI) by financial institutions, based on data from job ads. For example, they show that from 2015 to 2018 the stock of AI labor increases about 13 times.

4.2. Human and machine information access in the cross-section

To understand the drivers of *Machine* and *Human* viewership in the cross-section, we estimate the following regression:

$$V_{i,j,t}^{M,H} = \alpha_0 + \sum_{j=1}^k \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,j,t} \quad (3)$$

²⁰The cost-efficient benefits of cloud computing to many organizations stem primarily from an economy of scale. However, cloud computing might be cost-prohibitive for most retail investors, even when they are technically capable of overcoming the computational hurdles. For example, the monthly cost for a 60% utilization of a cloud machine instance recommended by AWS as a minimum requirement for a deep learning algorithm (8 vCPU, 1 GPU, 61 RAM, and 2T SSD) is nearly \$2,000. The monthly cost is more than \$7,000 for a cloud instance at the next level (32 vCUP, 4 GPU, 244 RAM, and 2T SSD). An AWS cost calculator can be found at <https://calculator.aws>.

where $V_{i,j,t}^{M,H}$ is our measures of the number of 8-K views by machines or humans or both (total visits) on days $t \in \{0, 1\}$. The control variables, $\Gamma_{i,j,t}$, include several firm characteristics (indexed by i), market conditions, and 8-K characteristics (indexed by j), at time t , such as the negative sentiment (*FinNeg*) content in the 8-K, the Fog readability (*FOG*), the word count of text in the 8-K (*WordCount*), the days-to-release that captures the number of days between the event date and the 8-K filing date (*DaysRelease*), the number of items in the 8-K (*#Item*), the firm's book-to-market ratio (*BM*), the firm's market capitalization (*SIZE*), the firm's institutional ownership (*InstOwn*), and the number of analysts following the firm (*Analysts*). All variables are defined in the Appendix Table A. \tilde{f} and $\tilde{\tau}$ are firm and year fixed effects, respectively.

Table (2) reports the results of these regressions. Column (1) shows that 8-K filings with considerable negative sentiment content tend to receive more total visits, as reflected by the positive, though marginally significant, *FinNeg* coefficient (1.026). This tendency is strongest for humans (Columns 3 and 5) and is insignificant for machines (Columns 2 and 4). These results suggest that negative sentiment content tends to attract more attention from human viewers. Column (1) also shows that larger 8-Ks are more likely to have a significantly larger total number of visits, as proxied by the word count in the 8-K (*WordCount*) and the number of items included in the 8-K (*#Item*). High-dimensional and voluminous 8-Ks are likely to reveal more material information, prompting both machine and human readers to examine them more actively (Columns 2-5). Column (1) also shows that timely 8-Ks are likely to receive more attention, as the number of days between the event date and the 8-K publication date (*DayRelease*) is negatively related to total viewership.

Table (2) also shows that larger firms (*SIZE*) have greater human viewership (Columns 3 and 5) but not significantly greater machine viewership (Columns 2 and 4). This result suggests that human readers tend to download the filings of larger and perhaps more established corporations with longer histories than those of smaller firms, consistent with Cao et al. (2020). Columns (3 and 5) also reveal that value firms (*BM*) have greater human viewership but not significantly greater machine viewership (Columns 2 and 4). Similarly, humans pay more attention to the 8-Ks of firms with low institutional ownership (*InstOwn*), which are likely to lack information transparency, yet machines show no preference over those firms. These findings are consistent with prior studies highlighting the

implications of two commonly understood differences between humans and machines, which are due to their capacity and rationality (Abis, 2020).

We partition 8-K filings into topical categories based on the items contained within the 8-K and examine how human and machine viewership varies across the topical categories.²¹ We regress the measures of human or machine viewership on a set of dummy variables for the different 8-K item types and control for the same factors as in Table (2).²²

Table (3) reports the results of those regressions. They show that there is significant heterogeneity in the viewing activity of the different items in the 8-K by machines and humans. For example, column (1) shows that machines do not pay significant attention to item 2.02 that mostly refers to non-GAAP earning disclosures, as reflected in the insignificant item2.02 coefficient. In contrast, humans pay significant attention to item 2.02 (column 6). Humans may have a particular preference for events that convey certain and well-known types of information, but machine viewers do not have such preference. Columns (2) and (7) show that machines, but not humans, pay significantly more attention to item 8.01 (i.e., Other Events). These results are consistent with the notion that machines rely on their ample computational capacity to identify material information that is not standardized in a typical, anticipated 8-K. Human viewers, in contrast, are more likely to target specific information disclosure events rather unspecified ones.

Humans tend to pay less attention to item 5.02, relative to other items in 8-K filings (Column 9). Therefore, human readers appear to be less attentive to unanticipated events related to the departure or election of officers or directors. In contrast, machines do not significantly under allocate attention to item 5.02. Lastly, the total viewership by both machines and humans is higher for voluntary disclosure items 7.01 (i.e., Regulation FD Disclosure) and mandatory disclosure item 1.01 (i.e., Entry into a Material Definitive Agreement) as reported in columns (3, 5, 8, and 10).

In summary, our findings are consistent with the notion that humans and machines scrutinize specific 8-K items differently. Our results suggest that automated models are more attentive to new

²¹We focus on the five most common categories. As suggested by Panel B of Table 1, the rest of the categories altogether are around 13.15% of 8-K categories in our sample.

²²While the main results are based on the measures of machine and human viewership proposed by Ryans (Dechow et al., 2015), in an unreported table we repeat the tests with the DRT measures (Drake et al., 2015) and find similar results.

information disclosed in non-standardized 8-K items.

5. Impact of Machines on Market Efficiency

This section investigates the effects of machine and human viewership of 8-K filings on the underlying firm's stock price efficiency. First, we examine machines in aggregate. Then we partition machines into types and compare their effects. Finally, we examine an alternative efficiency measure. We examine endogeneity issues in the subsequent section.

We hypothesize that machines will contribute to a faster and more efficient reaction of stock prices to new information. Therefore, we expect that the price drift following 8-K releases will be larger when there is a greater prevalence of humans relative to machines reading the filing. Among machines, we expect that the investors using cloud computing services are likely to be among the most sophisticated and therefore will have the largest positive impact on informational efficiency.

5.1. Machines in aggregate

We estimate the following regression:

$$DRIFT(2, T)_{i,t} = \alpha_0 + \sum_{j=1}^2 \beta_j V_{i,t}^{M,H} + \sum_{j=1}^k \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \quad (4)$$

where $DRIFT(2, T)_{i,t}$ is a measure of *inefficiency* defined as the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). $V_{i,t}^{M,H}$ is our measures of the 8-K viewership by machines or humans. The control variables, $\Gamma_{i,j,t}$, include several firm and 8-K filing characteristics, such as firm i 's book-to-market ratio, size, return-on-assets, leverage, standard deviation of monthly return over the year prior to the 8-K, institutional ownership, and analyst coverage. All variables are defined in the Appendix Table A. \tilde{f} and $\tilde{\tau}$ are firm and year fixed effects, respectively.

Table (4) reports the results. Columns (1-3) provide the first evidence that the price drift following the 8-K publication date is mainly associated with human viewership of the 8-K, rather than by aggregated machine viewership. In particular, column (1) shows that the *aggregate* machine viewership (including all categories of machines) of 8-K has no significant effect on the price drift

following the 8-K publication date, as shown by the insignificant *Machine* coefficient. In contrast, column (2) shows that 8-K viewership by humans has a positive and highly significant effect on the price drift. Column (3) simultaneously accounts for viewership by machines and humans and confirms the findings reported in columns (1-2).

In columns (4-6) of Table (4) we extend our analysis by increasing the $DRIFT(2, T)$ window from $DRIFT(2, 10)$ to $DRIFT(2, 20)$, and in column (7-9) to $DRIFT(2, 40)$, with the same control variables. We find consistent and even stronger evidence regarding the impact of viewership by machines and humans on the price drift.²³ We also examine an alternative measure of inefficiency in Section (5.3) and find similar results. The effects of human viewership on the price drift are also economically significant. For example, a one standard deviation increase in human viewership is associated with a 13.23% aggregate increase in the price drift over the post-publication month.²⁴ Therefore, the well-established inefficiencies from overreaction and underreaction to information in financial markets is driven by humans and not machines.

The regressions control for several firm characteristics. Across all columns (1-9) we find that firm size and ROA are negatively related to the price drift post to 8-K filing date suggesting that larger and more profitable firms have more efficient prices. Consistent with Ben-Rephael et al. (2020), columns (1-9) show a negative and highly significant relation between the institutional ownership and the price drift post to 8-K filing date, with the coefficient of *InstOwn* ranging between -0.011 and -0.021 (significant at the 1% confidence level).²⁵

5.2. Heterogeneity among machines

Machines are far from a monolithic group. In fact, a considerable proportion of machine-based viewership is not directly associated with sophisticated investors, nor with investment strategies

²³We also repeat our analysis with an alternative machine viewership, which is defined as the fraction of machine or cloud machine viewership to total visits. The results, reported in the Appendix Table C are qualitatively similar. We also repeat all tests with the machine and human viewership measured by the DRT (Drake et al., 2015) method and the results are qualitatively similar.

²⁴The unreported standard deviation of raw human visits in the first two days is 51.96, and its log is $\ln(1 + 51.96) = 3.97$. Given average number of post-event price drift, $DRIFT(2, 20)$ is 0.09, one standard deviation increase in human visits leads a 13.23% ($0.003 \times 3.97/0.09$) increase in post-event price drift.

²⁵In their analysis of abnormal institutional attention and price discovery during the filing period, Ben-Rephael et al. (2020) find that the price discovery prior to the filing period is 9.7% higher when institutional investors are paying attention, which results in a reduction of 9.7% in the subsequent price discovery during the filing period.

(such as viewership from university, regulators, and auditing firms). While the same can be said of humans, the focus on this paper is in better understanding the role machines play in markets and therefore this section examines whether different types of machines have different effects.

We expect that the investors using cloud computing are likely to be among the most sophisticated and therefore will have the largest positive impact on informational efficiency. Cloud computing enables financial institutions to adopt low latency strategies by providing scalable, sophisticated, and resilient infrastructures. Traditionally, as with start-ups in other industries, many small and mid-sized institutional investors, including hedge funds, turn to cloud computing to save operating expenses and minimize capital investment on hardware.²⁶ More recently, in comparison to on-premise solutions, cloud computing has been shown to increase agility, database efficiency, and access to cutting-edge analytics and artificial intelligence capabilities. Numerous prominent financial organizations, including Vanguard and Wellington, have migrated the majority of their on-premise hardware resources to cloud-based services.²⁷ Additionally, stock exchanges such as Nasdaq, Euronext and Deutsche Börse, as well as brokerage firms such as Robinhood and iSTOX, rely extensively on cloud servers. Investment banks such as Goldman Sachs and data vendors have also begun to offer services to institutional investors through cloud computing.²⁸ However, how institutional investors implement their investment strategies through cloud computing facilities largely remains secretive due to the competitive nature of the financial industry and the confidentiality of cloud computing.

We test the differences between machine types by using the categories that we classified on the basis of the machine’s IP address: (a) cloud computing services, (b) institutional investor, and (c) data & media publishers. We regress the inefficiency measure, $DRIFT(2,T)_{i,t}$, on measures of viewership by each of the machine types. We include the same control variables as before: BM , $SIZE$, ROA , LEV , STD_{RET} , $InstOwn$, and $Analysts$.

Table (5) reports the results from this analysis. The results (columns (1) and (4)) show that

²⁶“Mirae Asset Global Investments Case Study,” <https://aws.amazon.com/solutions/case-studies/mirae-asset/>, accessed August 23, 2021.

²⁷For more details, please see AWS re:Invent Talks of Jeff Dowds (2019) and Coline Mazzola (2018).

²⁸More disclosed cases can be found in “Financial Services at AWS.”, <https://aws.amazon.com/financial-services/> and “Google Cloud for financial services” <https://cloud.google.com/solutions/financial-services> Accessed August 23, 2021.

machine viewership from cloud computing services does indeed have a significant positive association with informational efficiency (negative association with drift). In terms of economic significance, a one standard deviation increase in the machine-based viewership of 8-K from cloud computing services is associated with a 10.49% aggregate decrease in the price drift over the post-filing month.²⁹ Extending the price drift window up to 40 trading days following the 8-K publication date, the machine viewership of 8-K from cloud computing services is no longer statistically significant (column 7) although the point estimate remains the same. These results are consistent with more variability and less statistical power in the longer horizon measure, or that investors using cloud computing services tend to exploit their informational advantage over a short-term horizon.

The estimates in columns (2-3), (5-6), and (8-9) of Table (5) show that machine viewership from financial institutions other than those using cloud computing facilities (*InstInvMachine*) and machine viewership from data & media publishers (*DataVendMachine*), respectively, does not have a significant effect on the price drift.³⁰

5.3. Evidence from alternative measures

To test the robustness of our main results, we re-estimate the regressions from the previous subsections but using a different measure of information (in)efficiency – the *Noise* measure developed by Brogaard et al. (2020). This measure is the standard deviation of estimated pricing errors obtained from a variance decomposition model that separates information and noise. It captures both under-reaction and overreaction to information as two forms of inefficiency that contribute to “noise” in prices. In a series of validation tests, Brogaard et al. (2020) show that the *Noise* measure captures informational inefficiency in prices (Appendix Table E).

Similar to the previous subsections, we regress *Noise* on measures of viewership by each of the machine types and include the same control variables as before. The results support the conclusions

²⁹The unreported standard deviation of raw machine visits from clouding services in the first two days is 110.98, and its log is $\ln(1 + 110.98) = 4.72$. Given average number of post-event price drift, $DRIFT(2, 20)$, is 0.09, one standard deviation increase in machine visits from cloud services leads a 10.49% ($= 0.002 \times 4.72/0.09$) decrease in post-event price drift.

³⁰Once again, in an unreported table in which we repeat our tests with the machine-based viewership defined by DRT (Drake et al., 2015), the results are qualitatively similar.

made based on the drift measure.³¹ Specifically, human viewership has a significant positive association with noise in prices, while cloud computing machine viewership has a significant negative association with noise. These results therefore support the notion that access to information by machines operating from cloud computing services tends to improve the informational efficiency of prices.

6. Addressing Endogeneity

While the evidence so far is consistent with our main hypothesis, to more formally examine the causality between human/machine access to information and informational efficiency, we exploit two identification strategies: (a) exogenous shocks, and (b) a quasi-natural experiment. In addition, we also exploit instrumental variables as a final identification strategy.

6.1. Evidence from exogenous disruptions

Our first approach is based on exogenous cloud computing service outages and major power outages that disproportionately disrupt machines. While cloud computing outages unambiguously reduce the capacity for cloud-based machines to access information, we use major power outage events in the US as a disruption to machine-based viewership more generally. To focus on the events that are likely to affect many investors using automated models, we consider a day as having an outage (an *Outage Day*) if it hits one of major cloud service providers, such as Amazon Web Services, or impacted the electricity supply of at least 500,000 customers and lasted for at least 10 hours. Such outages, by interrupting the connectivity of computing facilities running automated investment models, are expected to reduce the ability for machines to improve price discovery.

We regress $DRIFT(2, T)_{i,t}$ on machine and human viewership and a dummy variable that is one on *outage* days and zero otherwise. We include interactions between cloud computing, machine and human viewership, and the *outage* days dummy variable, and control for our standard list of stock characteristics: *BM*, *SIZE*, *ROA*, *LEV*, *STD_{RET}*, *InstOwn*, and *Analysts*.

Table (6) reports the results. As expected, we find that the coefficients of the interaction between

³¹The results are reported in the Appendix Table E.

CloudMachine or *Machine* viewership and *Cloud Outage Day*, $\hat{\beta}_2$ and $\hat{\beta}_4$, are positive and significant in columns (1) and (2). These estimates suggest that during unexpected cloud outage days the viewership through some clouds is less effective in reducing the price drift post to 8-K publication day, and hence there is a corresponding deterioration in market efficiency (positive association with drift). The relationship remains consistent up to 20 trading days after the 8-K publication day (columns 4 and 5), though it becomes statistically weaker. In contrast, we find no significant effects for the interaction term between human viewership and major cloud outages (insignificant $\hat{\beta}_6$ coefficients in columns (3) and (6)).

Similarly, the coefficient of the interaction between *Machine* viewership and *Power Outage Day* ($\hat{\beta}_4$) is also positive and significant in column (8), though the coefficient of interaction between cloud viewership and power outage is statistically insignificant. Besides, the significant relationship is only present up to 10 trading days after the 8-K publication day. Lastly, no significant effects are found for the interaction term between human viewership and major power outages (insignificant $\hat{\beta}_6$ coefficients in columns (9) and (12)). These results suggest that the power outages are more likely to interrupt the functionality of computing facilities that are not as geographically diversified or not as technologically advanced as major cloud service providers that tend to have backup power supplies.

6.2. S&P500 inclusion as a quasi-natural experiment

Our second approach uses a firm i 's inclusion in the S&P500 index as a quasi-natural experiment (Shleifer, 1986; Wurgler and Zhuravskaya, 2002; Bennett et al., 2020). A stock is added to the S&P500 index only when another stock is excluded, mainly due to major corporate actions such as mergers, bankruptcy, and spin-offs. The sudden surge in institutional investors' attention for the newly added stock creates an exogenous shock in viewership, disproportionately from human analysts and fund managers due to their limited capacity to follow stocks outside of their mandate (Farboodi et al., 2020).³² Human attention is a scarce cognitive resource (Barber and Odean, 2008; Kahneman, 1973) and influences the market's reaction to corporate announcements (e.g., Hirshleifer

³²Begenau et al. (2018) examine whether big data disproportionately benefits large firms by splitting firms by S&P500 index membership. Similarly, Farboodi et al. (2020) examine whether firms in the S&P500 index have different stock price informativeness.

et al. (2009); Kempf et al. (2017)). The existing literature also shows that a stock's inclusion in the S&P500 index attracts more sophisticated and skilled investors (Chen et al., 2004). This increase in sophistication and skill is another reason why inclusion in the S&P500 index is expected to improve the impact of humans on informational efficiency.

In contrast, machines are less likely to be constrained in their capacity to process information and therefore, as we demonstrated earlier, show less of a tendency to favor particular types of stocks in their information acquisition activities. We therefore expect inclusion in the S&P500 index to have little or no effect on how machines impact informational efficiency.

We use difference-in-differences regressions to examine the joint effects of 8-K viewership by humans or machines and the firm's inclusion in the S&P500 index on the price drift following 8-K filings. The dependent variable is the $DRIFT(2, T)_{i,t}$ measure. As the key independent variables, we have interactions of machine or human viewership of firm i 's 8-K with a dummy variable for S&P500 index inclusion ($InIndex_t$). This allows us to also include the $InIndex_t$ variable by itself to capture any index effects that are not associated specifically with human or machine viewership of 8-Ks. We also include the human and machine viewership variables by themselves to capture their effects that are not related to index inclusion, firm and year fixed effects, and a range of control variables. We also limit the sample in these tests to stocks that are included in the S&P500 index at some stage during our sample period to reduce the effect of other firm characteristics associated with the price drift.

Table (7) reports the results of the difference-in-difference tests. The primary coefficients of interest are for the interaction term between human viewership and index inclusion, $Human \times InIndex$. These coefficients are negative in all specifications (columns 3, 6 and 9) and statistically significant at the 5% level. Consistent with our hypothesis, the results suggest that index inclusion improves the impact of humans viewership on informational efficiency. That is, all else equal, human viewership in S&P500 stocks is more likely to improve efficiency and reduce drift than human viewership in non-S&P500 stocks.

In contrast, the coefficients of $Machine \times InIndex$ (columns 2, 5 and 8), are insignificant, sug-

gesting that S&P500 inclusion does not significantly affect the impact of machines on price drift. This result is also consistent with our hypothesis that, unlike human viewership, automated models are less resource-limited and are not concentrated in index stocks.

6.3. Instrumental variables approach

We also address potential endogeneity using a two-stage least-squares (2SLS) framework in which human viewership is instrumented by various market sentiment proxies and macroeconomic news. Our tests are based on the notion that investor sentiment (measured using the index of Baker and Wurgler (2006)), macroeconomic news announcements (Peng and Xiong, 2006; Kacperczyk et al., 2016), changes in the *VIX* during the window $(-20, -1)$ relative to 8-K publication date, and the CRSP value-weighted market return over the same window $(-20, -1)$ (*CRET* $(-20, -1)$) are likely to disproportionately affect human’s decision. For example, humans are more likely to be affected by sentiment and more likely to be constrained in their ability to process relevant information during periods of intense information arrival. While these instrumental variables are likely to affect market efficiency, their effects are through the human participants to access and assimilate relevant company-specific information. Therefore, our first-stage regression is:

$$V_{i,t}^H = \alpha_0 + \beta_1 InvSent + \beta_2 MacroNews + \beta_3 VIX(-20, -1) + \beta_4 CRET(-20, -1) + \sum_{j=1}^k \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \quad (5)$$

where $V_{i,t}^H$ is human viewership of firm i ’s 8-K on days $t \in \{0, 1\}$, measured by the Dechow et al. (2015) method. We estimate the first-stage equation for machines or humans separately. *InvSent*, *MacroNews*, *VIX* $(-20, -1)$, and *CRET* $(-20, -1)$ are our instruments. $\Gamma_{i,j,t}$ is a vector of controls j including *BM*, *SIZE*, *ROA*, *LEV*, *STD_{RET}*, *InstOwn* and *Analysts*. \tilde{f} and $\tilde{\tau}$ are firm and year fixed effects.

Appendix Table D reports the results. Columns (3) shows that our instruments significantly affect human viewership of 8-K filings as hypothesized. Human viewership is negatively related to investor sentiment, suggesting that humans are less likely to access information related to unscheduled events when investor sentiment is high. Human viewership is also negatively related to the release of

macroeconomic news at the time of the 8-K publication date. This result is consistent with humans having attention and cognitive constraints that make them less attentive to firm-specific information when there is substantial market-wide information, e.g., (Peng and Xiong, 2006; Kacperczyk et al., 2016). Lastly, human viewership is also negatively impacted by the $VIX(-20, -1)$ and the market return $CRET(-20, -1)$, consistent with humans having a tendency to read more 8-Ks in rising markets and during lower volatility. While all sentiment metrics suggested by previous research are significant, we have an overidentification problem for the endogenous variable, i.e., human viewership. We then test the null hypothesis of whether any of those instrumental variables is statistically weaker than the others. The Hensen J -statistics (reported in column 4-6 of Appendix Table D) indicate all sentiment instrumental variables have comparable impact on human viewership.

To formally test the relevant condition of instrumental variable, we report Kleibergen-Paap F -statistics (for non-i.i.d errors) in Appendix Table D. The F -statistics in Column (3) for the first stage is about 42.3 for human viewership, which is much greater than Stock and Yogo (2002)'s crucial values of 5% maximal IV relative bias and 10% maximal IV size tests.³³ Those results imply that human viewership identification does not suffer from an insufficient instrumental variable selection. To mitigate the concerns associated with alternative hypotheses in which the sentiment variables can affect information efficiency through the channels related to general market conditions, we conduct two falsification tests in which we regress the same set of instrumental variables on viewership on *Machine* or *CloudMachine*. The F -statistics in Columns (1) and (2) suggest our instrument variables do not affect information efficiency through non-human viewership.³⁴

The second-stage regressions are:

$$DRIFT(2, T)_{i,t} = \alpha_0 + \beta \widehat{V}_{i,t}^H + \sum_{j=1}^k \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \quad (6)$$

where $DRIFT(2, T)_{i,t}$ is our baseline measure of *inefficiency*, defined earlier in Equation (2), $\widehat{V}_{i,t}^H$ are the fitted values of human viewership, $\Gamma_{i,j,t}$ is the same vector of control variables, and \tilde{f} and $\tilde{\tau}$ are firm and year fixed effects. The second-stage results are also in Appendix Table D, columns

³³The critical values for the 5% maximal relative bias and the 10% maximal relative size are respectively 16.9 and 24.6 among our first stage sample.

³⁴We did not report the second stage results for those falsification tests, as the coefficients estimated of weak instruments in second stage would be inefficient, biased and misleading (Roberts and Whited, 2013).

(4-6). They are consistent with our baseline results discussed earlier. Namely, human viewership is significantly positively related to price drift. Hence the instrumental variables models reinforce the estimates reported in Table (4) and discussed in Section (5.1).

Overall, the results from the exogenous outages and the quasi-experimental tests of index inclusion events (in addition to the instrumental variables tests) support a causal interpretation of our baseline results about how humans and machines impact price drift following 8-K filings. In an unreported table, we repeat our tests with the machine and human viewership measured by DRT (Drake et al., 2015), and find qualitatively similar results.

7. Mechanisms

What is the mechanism by which viewership of information by cloud computing machine results in more efficient reflection of that information in prices? We conjecture that the mechanism is that machines access the information, they then process it and use it in their investment/trading decision models, resulting in trading by the machines, and that trading leads to the information being impounded into prices in a more timely and accurate manner. If this conjectured mechanism is correct, we should see (i) an increase in informed trading when cloud-based machines access information, reflecting more accurate interpretation of the information, and (ii) an increase in algorithmic trading when cloud-based machines access information, reflecting faster trading decisions through automation. In this section we test both of these links in the conjectured mechanism of how machines impact informational efficiency.

7.1. Does machine access to information increase informed trading?

Is the positive effect of cloud computing machines on informational efficiency the result of machines being better at interpreting the information (more informed), faster at interpreting the information, or both? To shed some light on this issue, we first test the relation between human and machine viewership and measures of informed trading (this subsection) and then the relation between viewership type and algorithm trading (next subsection).

The microstructure-based Probability of Informed Trading (PIN) is widely used as a measure of the time-varying information asymmetry in financial markets. It captures the information advantage held by better informed investors (Bharath et al., 2009) based on the properties of the order flow. We use the PIN to test how the level of informed trading varies depending on the levels of human and machine viewership of 8-Ks. We regress $PIN(0, T)$, which is the average daily Probability of Informed Trading of firm i over the window $(0, T)$ on machine and human viewership, controlling for BM , $SIZE$, ROA , LEV , STD_{RET} , $InstOwn$, and $Analysts$.

Table (8) presents our findings from this analysis. Machine viewership on days $t \in \{0, 1\}$ is positively related to $PIN(0, 1)$ and $PIN(0, 5)$ (columns 1 and 5). In contrast, we find a significantly negative relation between human viewership and $PIN(0, 1)$ and $PIN(0, 5)$ (columns 2 and 6). We further explore the results by machine category. We find that machine viewership from cloud computing services is predominant positive driver of the position relation with $PIN(0, 1)$ and $PIN(0, 5)$ (columns 3 and 7).

In summary, the results indicate that the machine (human) viewership is significantly positively (negatively) related to the probability of informed trading. These results support the earlier evidence that cloud computing machines in particular contribute to price discovery and improve informational efficiency, while humans tend to be less well informed traders.

7.2. Viewership and algorithmic trading

Automated investment decisions implemented through algorithmic trading are likely to be faster in responding to new information than manual trading decisions and in particular manual handling of order execution. This speed advantage of machines relative to humans is possibly one of the contributors to the positive effect of cloud computing machines on informational efficiency. To examine this notion, we test the relation between algorithmic trading, proxied by (a) the Odd Lot Ratio, and (b) the Trade Size, and contemporary machine and human 8-K viewership.³⁵ Algorithmic

³⁵As algorithmic trading based on newly arrived 8-K filings is likely to involve fast trading as investors race to trade on the basis of the information, we expect to see predominantly informed market orders. These should be reflected in the Odd Lot Ratio and the Trade Size suggested by Weller (2017), but not necessarily in other proxies of algorithmic trading that are more targeted at measuring algorithmic market making, such as the Order-to-Trade Ratio or the Cancel-to-Trade Ratio.

trading usually splits large trades into smaller order to spread out trade over time, blend in with other order flow, and avoid a large price change. For this reason, a smaller trade size and larger proportion of odd lots is often considered to reflect algorithmic trading. An automated model that improves market efficiency would involve both textual data analysis and algorithmic trading.

We regress either the Odd Lot Ratio or the Trade Size (as dependent variables in separate estimations) on machine and human viewership. We control for *BM*, *SIZE*, *ROA*, *LEV*, *STD_{RET}*, *InstOwn*, and *Analysts*. Table (9) presents our findings from this analysis. There is interesting heterogeneity in the impact of human and machine viewership of 8-Ks on days $t \in \{0, 1\}$ on the distribution of *OddLotRatio*(0,1). As expected we find that human 8-K viewership is negatively related to the amount of algorithmic trading proxied by the *OddLotRatio*(0,1) (column 2). In contrast, we find that the machine 8-K viewership from cloud computing services is positively related to the *OddLotRatio*(0,1) as indicated in column (3). Hence machine viewership from cloud computing services leads to more algorithmic trading.

We also examine how human and machine viewership of 8-Ks affects *TradeSize*(0,1) in columns (5-8). We find that human 8-K viewers are positively related to the trade size (column 6). In contrast, machine 8-K viewership from cloud computing services is negatively related to the trade size (column 7). Therefore, cloud machine (human) viewership is significantly and positively (negatively) related to algorithmic trading around the publication date of the 8-K.

8. When do machines have an edge over humans?

While machines have advantages such as being able to process more information and faster than humans, their advantage is unlikely to be universal, and in some setting humans may still have the upper hand. The key question in this section is therefore in what settings do machines have an edge, and conversely, when do humans have an advantage over machines?

First, we consider information that is difficult for humans to read such as linguistically complex information, which may increase the advantage of machines. Second, we consider the role of negativity content in 8-K filings and how they impact decision making – machines should be less susceptible

to emotions and therefore better able to handle information in the presence of negative sentiment content that is known to elicit bias in human decisions. Finally, we explore how well machines and humans process repeated information considering possible benefits for attention-constrained humans and potential errors from machines misinterpreting repeat information as new signals.

8.1. Information that is difficult for humans to read

We first test whether machines have an advantage over humans when information is difficult for humans to read. To do this we examine linguistically complex 8-Ks. We use two proxies of *readability* to test the effectiveness in processing and assimilating such information: the *Gunning FOG*, and the *Flesch-Kincaid*. We label them as *DifficultToRead*. We regress $DRIFT(2, T)_{i,t}$ on machine and human viewership activities of 8-K filings, the two measures of 8-K *DifficultToRead*, and importantly interactions between machine and human viewership and *DifficultToRead*. We control for *BM*, *SIZE*, *ROA*, *LEV*, STD_{RET} , *InstOwn*, and *Analysts*.

The results are reported in Table (10) using *Gunning FOG* in columns (1-4) and *Flesch-Kincaid* in columns (5-8). Consistent with our conjecture, we find that machine 8-K viewership (in columns 1 and 5) lessens the price drift following 8-K filings by a larger amount for harder-to-read 8-Ks, which is information that will be more challenging for human readers to interpret. Another possible interpretation of this finding is consistent with recent evidence suggesting that the increasing machine downloading activity motivates firms to prepare filings according to machines' readership, processing capacity and capability (Cao et al., 2020). In contrast, the interactions between human viewership of 8-Ks and *DifficultToRead* measures are not statistically different from zero (columns (3-4) and (7-8)).

8.2. Negative sentiment and the role of emotions

We first examine whether machine viewers, in particular those associated with automated investment decisions such as cloud machines, are better able to handle negative sentiment content in 8-Ks. We regress $DRIFT(2, T)_{i,t}$ on (a) machine viewership from cloud computing services, (b) machine viewership from financial institutions, and (c) negative sentiment in a 8-K defined by Loughran and

Mcdonald (2011), as well as interactions between cloud machine and institutional investor machine with negative sentiment. We include a wide range of controls such as *BM*, *SIZE*, *ROA*, *LEV*, *STD_{RET}*, *InstOwn*, and *Analysts*.

Table (11) reports the results. The results support the notion that machines are better at handling negative sentiment without bias and therefore have a larger positive effect on informational efficiency (larger negative effect on drift) when information is more likely to induce bias in humans due to negative content. Specifically, columns (1), (3), and (5) show that machine viewership from cloud computing services reduces the price drift more among the 8-K with considerable negative sentiment content. The interaction term coefficients are negative and significant up to two months following the 8-K publication date. Along similar lines, results in columns (2), (4), and (6) show that machine viewership from financial institutions also reduces the price drift following 8-Ks with higher negative content, even though the machine viewership from institutional investors does not affect price drift overall. When we replace machine viewership from cloud computing services with aggregated machine viewership, the interaction variables are no longer significant, suggesting machine viewership from cloud computers and financial institutions play an important role in correcting the emotional biases of humans.

8.3. Identifying incremental information

Company disclosures, including 8-K filings, often contain some overlap with existing information and some incremental information. Discerning what is incremental and what is repeated, and thus already reflected in the price, is an important step in using information in trading decisions. For example, Peress and Schmidt (2021) highlight how distinguishing between novel and stale information is a pervasive challenge for investors, even the most sophisticated informed traders.

For machines, this challenge is likely to be particularly acute. Some machines are likely to misinterpret repeat information as a new signal, particularly when machines are trained to respond to individual 8-K filings. For example, if a second 8-K filing arrives with a degree of overlap in the underlying information, machines trained to react to individual filings are likely to overestimate the information content of the second filing by failing to separate the repeated vs incremental parts of

the information.

In contrast, humans may benefit from repeated information given their attention and capacity constraints mean that some humans will miss information in the first instance, leading to a general tendency for humans to under-react to information. Repeating some or all of the content of a previous information disclosure may help the inattentive humans that missed the information initially, identify and react to it, helping impound the information in prices.

To examine these conjectures, we use item 2.02 of the 8-Ks, which largely reveals the same earnings information that has already been disclosed in the press release in an abbreviated format. We regress $DRIFT(2, T)_{i,t}$ on machine and human viewership activities, a dummy variable that is one for 8-K filings that include *item2.02*, and zero otherwise, as well as interactions between machine and human viewership and the *item2.02* dummy. We include standard controls: *BM*, *SIZE*, *ROA*, *LEV*, *STDRET*, *InstOwn*, and *Analysts*. All variables are defined in the Appendix Table A.

The results reported in Table (12) indicate that for incremental information (*item2.02*) neither machines in aggregate nor cloud-based machines make any significant contribution to reducing price drift, in contrast to their beneficial effects for 8-K filings overall. In contrast, the interaction term between human viewership and *item2.02* has a negative and significant effect on the price drift up to 10 or 20 days following the 8-K publication date. These results are consistent with the notion that machines do not suffer from attention constraints and hence they fully respond to information in the first instance of the information. Subsequent repeats of the information do not lead to a more efficient response from machines. Repeat information also does not appear to cause significant errors among machines, which face a risk of interpreting repeat information as a new signal. In contrast, the results suggest that the efficiency of the human reaction to information is improved when information is repeated, consistent with the presence of attention and capacity constraints among humans.

9. Conclusion

Advances in computing and machine learning are rapidly changing how investors acquire and use firm-specific information. We find that machine viewership of company 8-K filings has increased

exponentially during recent years and now represents a significant fraction of total information consumption via the SEC's EDGAR server.

Using novel data, we show that these changes in how information is accessed and used affect how the information is impounded into stock prices. Information acquired by machines from cloud computing servers significantly improves informational efficiency, decreasing price drifts following information releases, and decreasing noise in returns. In contrast, humans accessing the same information does not benefit informational efficiency as much and can even harm efficiency. Similarly, other types of machines are generally not as beneficial to market efficiency as cloud computing users. We address identification and find evidence of causality in these effects by examining exogenous outages of cloud computing services and major power disruptions, a quasi-natural experiment, and instrumental variables. Interestingly, while humans show variation in what information they access, consistent with attention and cognitive capacity constraints, machines are more uniform in accessing different types of information across different segments of the market, consistent with a higher capacity to process information.

The evidence points to the underlying mechanism being one where machines access information, use the information in decision models, leading to trading by the machines, and that trading drives the information to become impounded in prices. For example, we find that market microstructure measures of information trading increase immediately following cloud-based machine access to information as do measures of the amount of algorithmic trading.

Finally, on the issue of when do machines vs humans have an advantage, we find that the advantage of machines increases for information that is difficult for humans to read (linguistically complex). Also, we find that cloud-based machines are more effective in processing information that has negative sentiment content or tone compared to humans, consistent with the notion that machines are less susceptible to bias from emotions. Humans also react more efficiently when information is repeated, consistent with attention and capacity constraints among humans that result in some humans missing information in the first instance. In contrast, machines show no such tendencies and repeated information neither benefits nor harms the efficiency of machines' reactions as they are more likely to process information when it first arrives. Our findings are also consistent with theory that humans

specialise in slower decision making that combines sequential information whereas machines have an edge in making quick decisions on isolated information.

Overall, our findings uncover the important role of machine-based quantitative investment or trading strategies in assimilating and incorporating information in the underlying firm's stock prices. Our findings contribute to the ongoing debate on the role of machines in financial markets and their impact on informational efficiency.

References

- Abis, S., 2020. Man vs . machine : Quantitative and discretionary equity management. Columbia University Working Paper pp. 1–79.
- Abis, S., Veldkamp, L., 2021. The changing economics of knowledge production. Available at SSRN 3570130 .
- Admati, A. R., Pfleiderer, P., 1988. A Theory of Intraday Patterns: Volume and Price Variability. *The Review of Financial Studies* 1, 3–40.
- Akbas, F., Armstrong, W. J., Sorescu, S., Subrahmanyam, A., 2015. Smart money, dumb money, and capital market anomalies. *Journal of Financial Economics* 118, 355 – 382.
- Avramov, D., Cheng, S., Metzker, L., 2020. Machine learning versus economic restrictions: Evidence from stock return predictability. Working Paper pp. 1–72.
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance* 61, 1645–1680.
- Baldauf, M., Mollner, J., 2020. High-frequency trading and market performance. *The Journal of Finance* 75, 1495–1526.
- Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6, 159–178.
- Barber, B., Odean, T., 2002. Online investors: Do the slow die first? *Review of Financial Studies* 15, 455–488.
- Barber, B. M., Odean, T., 2008. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *The Review of Financial Studies* 21, 785–818.
- Begenau, J., Farboodi, M., Veldkamp, L., 2018. Big data in finance and the growth of large firms. *Journal of Monetary Economics* 97, 71–87.

- Ben-Rephael, A., Da, Z., Easton, P., Israelsen, R. D., 2020. Who pays attention to sec form 8-k? Working Paper pp. 1–51.
- Bennett, B., Stulz, R. M., Wang, Z., 2020. Does joining the s&p 500 index hurt firms? Working paper series, Ohio State University, Charles A. Dice Center for Research in Financial Economics.
- Bharath, S. T., Pasquariello, P., Wu, G., 2009. Does Asymmetric Information Drive Capital Structure Decisions? *The Review of Financial Studies* 22, 3211–3243.
- Bradshaw, M. T., Lock, B., Wang, X., Zhou, D., 2020. Soft information in the financial press and analyst revisionsanalyst revisions. *The Accounting Review* .
- Brogaard, J., Nguyen, T. H., Putniņš, T. J., Wu, E., 2020. What moves stock prices? the role of news, noise, and information. In: AFA Annual Meeting.
- Bybee, L., Kelly, B. T., Manela, A., Xiu, D., 2020. The structure of economic news. Tech. rep., National Bureau of Economic Research.
- Cao, C., Liang, B., Lo, A. W., Petrasek, L., 2018. Hedge fund holdings and stock market efficiency. *The Review of Asset Pricing Studies* 8, 77–116.
- Cao, S., Jiang, W., Yang, B., Zhang, A. L., 2020. How to talk when a machine is listening: Corporate disclosure in the age of ai. Working Paper 27950, National Bureau of Economic Research.
- Chen, H., Cohen, L., Gurun, U., Lou, D., Malloy, C., 2020a. Iq from ip: Simplifying search in portfolio choice. *Journal of Financial Economics* 138, 118 – 137.
- Chen, H., Noronha, G., Singal, V., 2004. The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation. *The Journal of Finance* 59, 1901–1930.
- Chen, Y., Kelly, B., Wu, W., 2020b. Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics* 138, 316 – 341.
- Chordia, T., Miao, B., 2020. Market efficiency in real time: Evidence from low latency activity around earnings announcements. *Journal of Accounting and Economics* 70, 101335.

- Cookson, J. A., 2018. When saving is gambling. *Journal of Financial Economics* 129, 24–45.
- Dechow, P. M., Lawrence, A., Ryans, J. P., 2015. SEC Comment Letters and Insider Sales. *The Accounting Review* 91, 401–439.
- Dorn, D., Sengmueller, P., 2009. Trading as entertainment? *Management Science* 55, 591–603.
- Drake, M. S., Roulstone, D. T., Thornock, J. R., 2015. The determinants and consequences of information acquisition via edgar. *Contemporary Accounting Research* 32, 1128–1161.
- Dugast, J., Foucault, T., 2018. Data abundance and asset price informativeness. *Journal of Financial Economics* 130, 367 – 391.
- Dyer, T., 2019. Does public information acquisition level the playing field or widen the gap? an analysis of local and non-local investors. Working Paper pp. 1–63.
- Fama, E. F., Fisher, L., Jensen, M. C., Roll, R., 1969. The adjustment of stock prices to new information. *International Economic Review* 10, 1–21.
- Farboodi, M., Matray, A., Veldkamp, L., Venkateswaran, V., 2020. Where Has All the Data Gone? NBER Working Papers 26927, National Bureau of Economic Research, Inc.
- García, D., 2013. Sentiment during recessions. *The Journal of Finance* 68, 1267–1300.
- Grinblatt, M., Keloharju, M., 2009. Sensation seeking, overconfidence, and trading activity. *Journal of Finance* 64, 549–578.
- Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70, 393–408.
- Gu, S., Kelly, B., Xiu, D., 2020. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies* 33, 2223–2273.
- Gunawi, H. S., Hao, M., Suminto, R. O., Laksono, A., Satria, A. D., Adityatama, J., Eliazar, K. J., 2016. Why does the cloud stop computing? lessons from hundreds of service outages. In: *Proceedings of the Seventh ACM Symposium on Cloud Computing*, pp. 1–16.

- Hendershott, T., Jones, C. M., Menkveld, A. J., 2011. Does algorithmic trading improve liquidity? *The Journal of Finance* 66, 1–33.
- Hirshleifer, D., Lim, S. S., Teoh, S. H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance* 64, 2289–2325.
- Hwang, B.-H., Kim, H. H., 2017. It pays to write well. *Journal of Financial Economics* 124, 373–394.
- Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84, 571–626.
- Kahneman, D., 1973. *Attention and effort*. Prentice-Hall.
- Kempf, E., Manconi, A., Spalt, O., 2017. Distracted Shareholders and Corporate Actions. *The Review of Financial Studies* 30, 1660–1695.
- Kokkonen, J., Suominen, M., 2015. Hedge funds and stock market efficiency. *Management Science* 61, 2890–2904.
- Kumar, A., 2009. Who gambles in the stock market? *The Journal of Finance* 64, 1889–1933.
- Kyle, A. S., 1985. Continuous auctions and insider trading. *Econometrica* 53, 1315–1335.
- Kyle, A. S., 1989. Informed speculation with imperfect competition. *The Review of Economic Studies* 56, 317–355.
- Lee, C., Watts, E. M., 2020. Tick size tolls: Can a trading slowdown improve price discovery? *The Accounting Review* Forthcoming.
- Linnainmaa, J. T., 2011. Why do (some) households trade so much? *Review of Financial Studies* 24, 1630–1666.
- Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance* 66, 35–65.
- Meulbroek, L. K., 1992. An empirical analysis of illegal insider trading. *The Journal of Finance* 47, 1661–1699.

- Mukherjee, S., Nateghi, R., Hastak, M., 2018. Data on major power outage events in the continental u.s. Data in Brief 19, 2079 – 2083.
- Peng, L., 2005. Learning with information capacity constraints. *The Journal of Financial and Quantitative Analysis* 40, 307–329.
- Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80, 563–602.
- Peress, J., Schmidt, D., 2021. Noise traders incarnate: Describing a realistic noise trading process. *Journal of Financial Markets* 54, 100618.
- Roberts, M. R., Whited, T. M., 2013. Endogeneity in empirical corporate finance¹. In: Handbook of the Economics of Finance, Elsevier, vol. 2, pp. 493–572.
- Shleifer, A., 1986. Do demand curves for stocks slope down? *The Journal of Finance* 41, 579–590.
- Sims, C. A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50, 665 – 690, swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.
- Sims, C. A., 2006. Rational inattention: Beyond the linear-quadratic case. *The American Economic Review* 96, 158–163.
- Stock, J. H., Yogo, M., 2002. Testing for weak instruments in linear iv regression.
- Tetlock, P. C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance* 62, 1139–1168.
- Van Kervel, V., Menkveld, A. J., 2019. High-frequency trading around large institutional orders. *The Journal of Finance* 74, 1091–1137.
- Veldkamp, L. L., 2006. Information Markets and the Comovement of Asset Prices. *The Review of Economic Studies* 73, 823–845.

Weller, B. M., 2017. Does Algorithmic Trading Reduce Information Acquisition? *The Review of Financial Studies* 31, 2184–2226.

Wurgler, J., Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks? *The Journal of Business* 75, 583–608.

Xiong, W., 2013. *Bubbles, Crises, and Heterogeneous Beliefs*. Cambridge University Press.

Figure 1: **8-K viewership activity by machines and humans**

This figure presents total, machine, and human viewership activities of firm i 's 8-K files on days $t \in \{0, 1\}$ (relative to the 8-K publication date). Total visits is the natural logarithm of one plus the number of views by both machines and humans. Machine (Human) visits is the natural logarithm of one plus the number of views by machines (humans) of 8-K filings of firm i on days $t \in \{0, 1\}$, measured by methods proposed by Dechow et al. (2015).

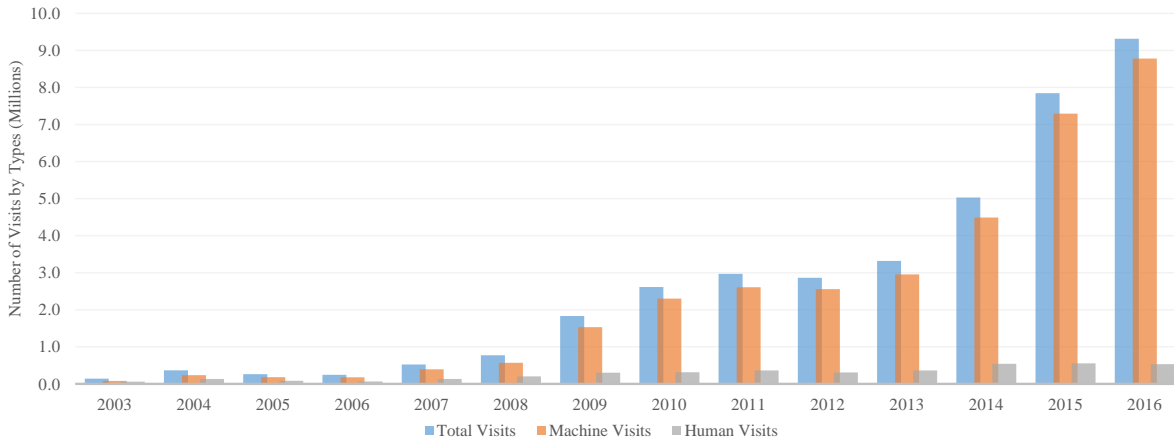


Figure 2: **8-K viewership activity by organization.**

This figure presents machine viewership activity of firm i 's 8-K filings by Cloud Computing, Institutional Investor, and Internet Provider on days $t \in \{0, 1\}$ (relative to the 8-K publication date), measured by methods proposed by Dechow et al. (2015).

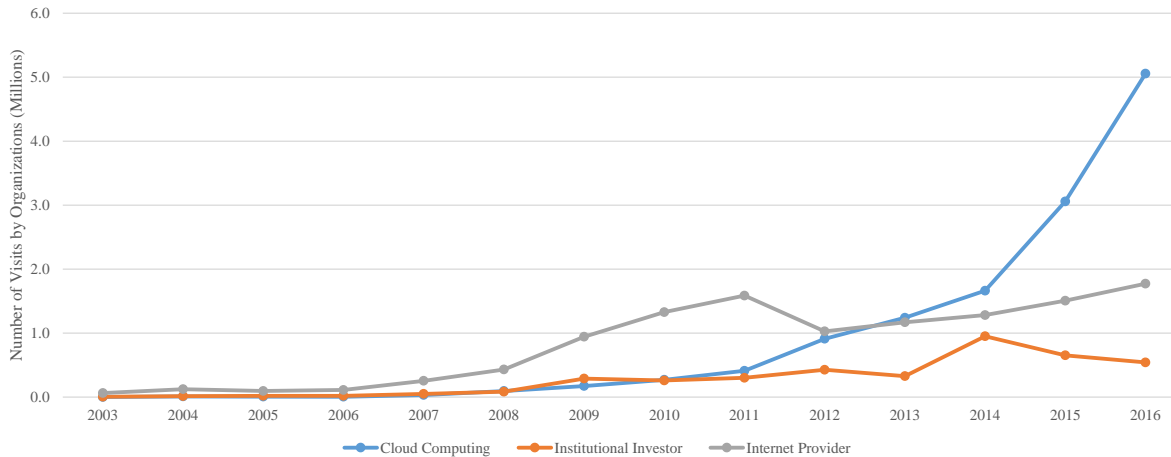


Table 1: Summary Statistics

This table reports summary statistics of our main variables. *Machine*, *CloudMachine*, *InstInvMachine*, *IntProvMachine*, *DataVendMachine* and *Human* (Panel A) denote viewership activities of firm *i*'s 8-K filing(s) on days $t \in \{0, 1\}$, relative to 8-K publication date, by machine (and sub-categories of machine) and human viewers, measured as in Dechow et al. (2015). Alternatively, viewership is measured as in Drake et al. (2015), and denoted by $ViewerType_{DRT}$. Our main (control) variables, which are winsorized at the 1st and 99th percentiles, include the firm *i*'s *BM*, *SIZE*, *ROA*, *LEV*, *STDRET*, *InstOwn* and *Analysts*. $DRIFT(2, 10)$, $DRIFT(2, 20)$, and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a (short) window right after the 8-K announcement, as described in Equation (2). *Item#.0#* (Panel B) refers to a dummy variable equal to one if a firm *i*'s 8-K filing containing an *Item#.0#*, and zero otherwise. Accordingly, the *Item1.01* refers to *Entry into a Material Definitive Agreement*, the *Item2.02* refers to *Results of Operations and Financial Condition*, the *Item5.02* refers to *Departure of Directors or Certain Officers Election of Directors Appointment*, the *Item7.01* refers to *Regulation FD Disclosure*, and the *Item8.01* refers to *Other Events*. Lastly, visiting organizations (Panel C) refer to the number of 8-K visits by organizational types on days $t \in \{0, 1\}$ relative to 8-K publication date. All variables are defined in the Appendix Table A.

	N	Mean	Std Dev	25th Pctl	Median	75th Pctl
<i>Panel A: Summary Statistics</i>						
<i>Machine</i>	547,883	3.30	1.42	2.08	3.61	4.41
<i>Machine_{DRT}</i>	547,883	3.35	1.41	2.08	3.66	4.45
<i>CloudMachine</i>	517,745	1.98	1.63	0.00	1.79	3.30
<i>CloudMachine_{DRT}</i>	517,745	1.99	1.62	0.00	1.79	3.30
<i>InstInvMachine</i>	517,745	1.43	1.12	0.69	1.61	2.30
<i>InstInvMachine_{DRT}</i>	517,745	1.47	1.13	0.69	1.61	2.40
<i>IntProvMachine</i>	517,745	2.41	1.19	1.39	2.83	3.30
<i>IntProvMachine_{DRT}</i>	517,745	2.48	1.19	1.61	2.89	3.37
<i>DataVendMachine</i>	517,745	1.10	0.89	0.00	1.10	1.79
<i>DataVendMachine_{DRT}</i>	517,745	1.11	0.90	0.00	1.39	1.79
<i>Human</i>	547,883	1.35	1.07	0.69	1.39	2.08
<i>Human_{DRT}</i>	547,883	1.05	0.96	0.00	1.10	1.61
<i>BM</i>	547,883	0.75	0.99	0.30	0.55	0.89
<i>SIZE</i>	547,883	6.44	2.05	4.99	6.40	7.80
<i>ROA</i>	547,883	-0.03	0.20	-0.02	0.02	0.06
<i>LEV</i>	547,883	0.23	0.22	0.04	0.19	0.37
<i>STDRET</i>	547,883	0.12	0.10	0.07	0.10	0.15
<i>InstOwn</i>	547,883	0.61	0.32	0.34	0.66	0.87
<i>Analysts</i>	547,883	1.63	0.99	0.69	1.79	2.40
$DRIFT(2, 10)$	547,883	0.06	0.09	0.02	0.04	0.07
$DRIFT(2, 20)$	547,883	0.09	0.10	0.02	0.05	0.11
$DRIFT(2, 40)$	547,883	0.13	0.15	0.04	0.08	0.17
<i>Panel B: 8-K Items</i>						
	<i>#Filings</i>			<i>#Visits per Filing</i>		
Item 1.01	62,396	77.58	143.79	9	40	101
Item 2.02	193,330	69.91	138.98	9	41	90
Item 5.02	70,920	74.17	121.78	12	47	97
Item 7.01	117,747	78.84	184.95	9	45	102
Item 8.01	144,925	71.28	159.56	8	39	91
Other Items	89,218	82.72	147.72	15	52	108
<i>Panel C: Visiting Organizations(1000)</i>						
	<i>#Org.</i>			<i>#Visits per Organization (Org.)</i>		
Auditing & Law Firms	130	3.49	19.49	0.03	0.45	2.29
Cloud Computing Services	938	13.78	204.00	0.00	0.01	0.07
Data Vendor & Media	381	4.92	39.22	0.00	0.01	0.08
Education and Regulator	3,172	0.10	1.50	0.00	0.00	0.01
Institutional Investor	992	3.98	34.14	0.01	0.04	0.32
Internet Service Provider	1,153	10.14	77.68	0.01	0.03	0.48
Others	18,930	0.07	0.79	0.00	0.00	0.02

Table 2: Determinants of machine and human viewership of 8-Ks

This table reports OLS regression estimates from the analysis of determinants of firm i 's 8-K machine or human viewership activities on days $t \in \{0, 1\}$, relative to 8-K publication date. *TotalVisits* is the natural logarithm of one plus the number of firm i 's visits of 8-K(s) by both machines and humans. *Machine* and *Human* are the corresponding machine and human viewership activities of firm i 's 8-K(s) filings on days $t \in \{0, 1\}$, relative to 8-K publication date, which are measured as in Dechow et al. (2015). Alternatively, our machine and human viewership activities are measured as in Drake et al. (2015) and denoted by *Machine_{DRT}* and *Human_{DRT}*, respectively. *FinNeg* is the proportion of negative words in a firm i 's 8-K filing that is defined by Loughran and McDonald (2011), *FOG* is the Gunning fog index, a readability measure, that is computed based on the words used in a firm i 's 8-K filing, *WordCount* is the number of words in a firm i 's 8-K filing, *DayRelease* is the number of days between the event date to the firm i 's 8-K publication date, *#Item* is the number of topics included in a firm i 's 8-K filing, *BM* is the ratio of book value over the market value of common equity in the year end prior to a firm i 's 8-K filing date, *SIZE* is the natural logarithm of market value of common equity in the year end prior to a firm i 's 8-K filing date, *InstOwn* is the institutional ownership percentage calculated in the quarter prior to a firm i 's 8-K filing date, and *Analysts* is the natural logarithm of one plus the number of analyst covering a firm in the quarter prior to a firm i 's 8-K filing date, are control variables across all columns, which are winsorized at the 1st and 99th percentiles. All variables are defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj. R²* is the adjusted R^2 value. All t -statistics reported in parentheses are computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>TotalVisits</i>	<i>Machine</i>	<i>Human</i>	<i>Machine_{DRT}</i>	<i>Human_{DRT}</i>
	(1)	(2)	(3)	(4)	(5)
<i>FinNeg</i>	1.026*	-0.009	4.883***	0.195	5.608***
	(1.84)	(-0.02)	(9.08)	(0.35)	(10.05)
<i>FOG</i>	0.000	0.000	-0.002	0.000	-0.001
	(0.29)	(0.46)	(-1.21)	(0.39)	(-1.08)
<i>WordCount</i>	0.049***	0.050**	0.059***	0.051***	0.046***
	(3.13)	(2.94)	(5.84)	(3.11)	(5.00)
<i>DayRelease</i>	-0.020***	-0.013***	-0.056***	-0.016***	-0.042***
	(-6.81)	(-4.69)	(-8.12)	(-5.42)	(-8.39)
<i>#Item</i>	0.101***	0.083***	0.156***	0.095***	0.105***
	(11.58)	(8.34)	(15.90)	(9.66)	(12.62)
<i>BM</i>	0.009**	0.001	0.037***	0.003	0.037***
	(2.32)	(0.25)	(5.97)	(0.90)	(6.49)
<i>SIZE</i>	0.012	0.007	0.034***	0.008	0.032***
	(1.42)	(0.94)	(3.10)	(0.98)	(3.22)
<i>InstOwn</i>	0.004	-0.007	-0.148***	-0.001	-0.167***
	(0.08)	(-0.19)	(-3.15)	(-0.03)	(-3.65)
<i>Analysts</i>	-0.001	-0.003	0.007	-0.002	0.008
	(-0.05)	(-0.26)	(0.62)	(-0.24)	(0.73)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Nobs	553,339	553,339	553,339	553,339	553,339
Adj. R^2	0.845	0.872	0.405	0.862	0.349

Table 3: Machine and human viewership of different 8-K items

This table reports OLS regression estimates from our analysis on the determinants of viewership activities of firm i 's 8-K(s) on days $t \in \{0, 1\}$ (relative to the 8-K publication date) by *Machine* and *Human* viewers. *Machine* and *Human* denote their corresponding viewership activities of firm i 's 8-K(s) on days $t \in \{0, 1\}$ (relative to the 8-K publication date), measured as in [Dechow et al. \(2015\)](#). *Item#.0#* refers to a dummy variable equal to one if an 8-K filing containing an 8-K item#.0#, zero otherwise. Accordingly, the *Item1.01* refers to *Entry into a Material Definitive Agreement*, the *Item2.02* refers to *Results of Operations and Financial Condition*, the *Item5.02* refers to *Departure of Directors or Certain Officers Election of Directors Appointment*, the *Item7.01* refers to *Regulation FD Disclosure*, and the *Item8.01* refers to *Other Events*. Our control variables include the *FinNeg* that refers to the proportion of negative words defined by [Loughran and Mcdonald \(2011\)](#) in a firm i 's 8-K filing, the *FOG* that refers to the Gunning fog index, a readability measure, is computed based on the words used in a firm i 's 8-K filing, the *WordCount* that refers to the number of words used in a firm i 's 8-K filing, the *DayRelease* that refers to the number of days between the firm i 's 8-K related event date to the same firm i 's 8-K publication date, the *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to a firm i 's 8-K filing date, the *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to a firm i 's 8-K filing date, the *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to a firm i 's 8-K filing date, and the *Analysts* that refers to the natural logarithm of one plus the number of analyst covering a firm in the quarter prior to the firm i 's 8-K filing date, which are winsorized at the 1st and 99th percentiles. All variables are defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj.R²* is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Machine					Human				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
item2.02	-0.005 (-0.28)					0.227*** (6.77)				
item8.01		0.075** (2.85)					0.015 (1.59)			
item7.01			0.102*** (8.65)					0.137*** (8.91)		
item5.02				-0.032 (-1.43)					-0.205*** (-7.74)	
item1.01					0.114*** (5.64)					0.176*** (5.37)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	553,397	553,397	553,397	553,397	553,397	553,397	553,397	553,397	553,397	553,397
Adj. R^2	0.869	0.869	0.869	0.869	0.869	0.393	0.384	0.386	0.388	0.386

Table 4: Impact of machine and human viewership on price drift

This table reports OLS regression estimates from our analysis of the impact of *Machine* and *Human* viewership activities of firm i 's 8-K filings on days $t \in \{0, 1\}$ (relative to the 8-K publication date) on $DRIFT(2, 10)$, $DRIFT(2, 20)$, and $DRIFT(2, 40)$. $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *Machine* and *Human* denote machine and human viewership activities, measured as in Dechow et al. (2015). *BM* is the ratio of book value over the market value of common equity in the year end prior to a firm i 's 8-K filing date, *SIZE* is the natural logarithm of one plus the market value of common equity in the year end prior to a firm i 's 8-K filing date, *ROA* is the return on assets, *LEV* is the total debt over total assets, *STDRET* is the standard deviation of monthly return over the year prior to a firm i 's 8-K filing date, *InstOwn* is the institutional ownership percentage calculated in the quarter prior to a firm i 's 8-K filing date, and *Analysts* is the natural logarithm of one plus the number of analyst covering a firm i in the quarter prior to the firm i 's 8-K filing, are control variables across all columns, which are winsorized at the 1st and 99th percentiles. All variables are defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. $Adj.R^2$ is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>DRIFT(2, 10)</i>			<i>DRIFT(2, 20)</i>			<i>DRIFT(2, 40)</i>		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Machine</i>	0.003 (1.00)		0.001 (0.32)	0.004 (1.03)		0.002 (0.49)	0.006 (1.27)		0.003 (0.66)
<i>Human</i>		0.003*** (4.11)	0.003*** (6.75)		0.004*** (3.76)	0.003*** (6.85)		0.005*** (4.44)	0.004*** (6.47)
<i>BM</i>	0.004*** (6.16)	0.004*** (6.14)	0.004*** (6.12)	0.005*** (5.84)	0.005*** (5.75)	0.005*** (5.80)	0.007*** (5.70)	0.007*** (5.62)	0.007*** (5.63)
<i>SIZE</i>	-0.009*** (-10.88)	-0.009*** (-11.18)	-0.009*** (-11.06)	-0.012*** (-13.14)	-0.013*** (-13.55)	-0.013*** (-13.41)	-0.019*** (-13.23)	-0.019*** (-13.60)	-0.019*** (-13.50)
<i>ROA</i>	-0.030*** (-7.12)	-0.030*** (-7.19)	-0.030*** (-7.15)	-0.040*** (-8.51)	-0.040*** (-8.63)	-0.040*** (-8.56)	-0.061*** (-9.53)	-0.061*** (-9.60)	-0.061*** (-9.56)
<i>LEV</i>	0.020*** (4.72)	0.019*** (4.59)	0.019*** (4.66)	0.029*** (5.82)	0.028*** (5.72)	0.028*** (5.78)	0.046*** (5.63)	0.044*** (5.53)	0.045*** (5.59)
<i>STDRET</i>	0.028*** (3.39)	0.028*** (3.29)	0.028*** (3.32)	0.046*** (3.85)	0.045*** (3.76)	0.045*** (3.79)	0.078*** (4.97)	0.076*** (4.88)	0.076*** (4.91)
<i>InstOwn</i>	-0.012*** (-4.12)	-0.011*** (-4.02)	-0.012*** (-4.06)	-0.017*** (-4.83)	-0.017*** (-4.70)	-0.017*** (-4.78)	-0.021*** (-3.26)	-0.020*** (-3.15)	-0.021*** (-3.21)
<i>Analysts</i>	0.000 (0.08)	0.000 (0.06)	0.000 (0.07)	-0.000 (-0.50)	-0.001 (-0.53)	-0.001 (-0.53)	-0.002 (-1.28)	-0.002 (-1.31)	-0.002 (-1.30)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	547,883	547,883	547,883	547,883	547,883	547,883	547,883	547,883	547,883
Adj. R^2	0.203	0.204	0.204	0.235	0.236	0.236	0.257	0.258	0.258

Table 5: Machine types and price drift

This table reports OLS regression estimates from our analysis of the impact of *CloudMachine*, *InstMachine*, *DataMachine*, *OtherMachine* viewership activities of firm i 's 8-K filings on days $t \in \{0, 1\}$ (relative to the 8-K publication date) on $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$. $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *CloudMachine* is the machine viewership by cloud computing services; *InstMachine* is the machine viewership by financial institutions; *DataMachine* is the machine viewership by data and media publishers; finally *OtherMachine* is the machine viewership by the remaining categories. Control variables across all models include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; *STD_{RET}* that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. Adj. R^2 is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>DRIFT(2, 10)</i>			<i>DRIFT(2, 20)</i>			<i>DRIFT(2, 40)</i>		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>CloudMachine</i>	-0.002*** (-3.77)			-0.002** (-2.27)			-0.002 (-1.42)		
<i>InstInvMachine</i>		0.002 (1.23)			0.002 (1.23)			0.003 (1.26)	
<i>DataVendMachine</i>			-0.001 (-0.41)			-0.001 (-0.46)			-0.002 (-0.50)
<i>BM</i>	0.004*** (5.86)	0.004*** (5.88)	0.004*** (5.81)	0.005*** (5.49)	0.005*** (5.51)	0.005*** (5.43)	0.007*** (5.55)	0.007*** (5.56)	0.007*** (5.50)
<i>SIZE</i>	-0.008*** (-10.81)	-0.008*** (-10.91)	-0.008*** (-10.93)	-0.012*** (-12.95)	-0.012*** (-13.07)	-0.012*** (-13.11)	-0.018*** (-12.71)	-0.018*** (-12.82)	-0.018*** (-12.86)
<i>ROA</i>	-0.029*** (-7.21)	-0.029*** (-7.24)	-0.029*** (-7.20)	-0.041*** (-9.10)	-0.041*** (-9.17)	-0.041*** (-9.12)	-0.063*** (-9.61)	-0.063*** (-9.66)	-0.063*** (-9.62)
<i>LEV</i>	0.018*** (3.90)	0.018*** (3.87)	0.018*** (3.82)	0.028*** (5.46)	0.028*** (5.45)	0.028*** (5.35)	0.046*** (5.46)	0.046*** (5.45)	0.046*** (5.33)
<i>STD_{RET}</i>	0.028*** (3.53)	0.028*** (3.50)	0.028*** (3.52)	0.045*** (3.75)	0.045*** (3.73)	0.045*** (3.74)	0.079*** (4.59)	0.079*** (4.57)	0.079*** (4.59)
<i>InstOwn</i>	-0.011*** (-3.27)	-0.011*** (-3.28)	-0.011*** (-3.18)	-0.016*** (-4.08)	-0.016*** (-4.09)	-0.016*** (-3.93)	-0.020** (-2.69)	-0.020** (-2.69)	-0.020** (-2.64)
<i>Analysts</i>	-0.000 (-0.24)	-0.000 (-0.30)	-0.000 (-0.26)	-0.001 (-1.12)	-0.001 (-1.16)	-0.001 (-1.14)	-0.003* (-2.16)	-0.003** (-2.18)	-0.003** (-2.21)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	517,745	517,745	517,745	517,745	517,745	517,745	517,745	517,745	517,745
Adj. R^2	0.204	0.204	0.204	0.235	0.235	0.235	0.258	0.258	0.258

Table 6: Evidence from cloud and power outages

The table presents estimates from our analysis of the impact of disruption proxies on the price drift post to 8-K publication date as a test to our information-assimilation channel. *Cloud Outage Day* refers to the day characterised by a reported service disruption to one or more major cloud service providers; *Power Outage Day* refers to the day characterised by the loss of the electrical power network supply to an end user. We label them as *Outage day*. *DRIFT*(2, 10) and *DRIFT*(2, 20) are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *CloudMachine*, *Machine* and *Human* denote cloud machine, machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. Control variables include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; *STDEBT* that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj. R²* is the adjusted *R²* value. *t*-statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Cloud Outage Day					Power Outage Day						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>CloudMachine</i>	-0.002*** (-4.11)			-0.002*** (-2.40)			-0.002*** (-3.70)			-0.002*** (-2.26)		
<i>CloudMachine</i> × <i>OutageDay</i>	0.004*** (3.63)			0.004*** (2.50)			0.000 (0.32)			-0.001 (-0.40)		
<i>Machine</i>		0.003 (0.95)			0.005 (1.06)			0.003 (0.97)			0.005 (1.09)	
<i>Machine</i> × <i>OutageDay</i>		0.002*** (3.20)			0.002*** (2.22)			0.002*** (2.35)			0.002 (1.12)	
<i>Human</i>			0.003*** (4.35)			0.004*** (4.12)			0.003*** (4.84)			0.004*** (4.42)
<i>Human</i> × <i>OutageDay</i>			0.001 (1.21)			0.001 (0.70)			0.004 (1.48)			0.003 (1.06)
<i>OutageDay</i>	-0.008** (-2.83)	-0.006* (-1.86)	0.000 (0.19)	-0.010* (-2.02)	-0.009 (-1.70)	-0.001 (-0.20)	0.001 (0.34)	-0.006* (-2.01)	-0.003* (-1.93)	0.003 (0.41)	-0.005 (-1.04)	-0.002 (-0.88)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	517,739	517,739	517,739	517,739	517,739	517,739	517,739	517,739	517,739	517,739	517,739	517,739
Adj. <i>R²</i>	0.206	0.206	0.207	0.235	0.235	0.236	0.205	0.206	0.207	0.235	0.235	0.236

Table 7: The effects of S&P500 inclusion on machines and humans

This table reports regression estimates from difference-in-difference models that examine the joint effects of 8-K viewership and S&P500 inclusion on post-event price drift. $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *Machine* and *Human* denote machine and human viewership activities of firm i 's 8-K filings on days $t \in \{0, 1\}$. *InIndex* is a binary variable that is assigned the value of one if a stock is added into the S&P500 index, and zero otherwise. *BM* is the ratio of book value over the market value of common equity in the year end prior to an 8-K filing, *SIZE* is the natural logarithm of market value of common equity in the year end prior to an 8-K filing, *ROA* is the return on assets, *LEV* is the total debt over total assets, *STD_{RET}* is the standard deviation of monthly return over the year prior to an 8-K filing, *InstOwn* is the institutional ownership percentage calculated in the quarter prior to an 8-K filing, and *Analysts* is the number of analyst covering a firm in the quarter prior to an 8-K filing, are control variables across all columns. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. $Adj.R^2$ is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>DRIFT</i> (2, 10)			<i>DRIFT</i> (2, 20)			<i>DRIFT</i> (2, 40)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>CloudMachine</i>	-0.002** (-2.19)			-0.001 (-0.98)			-0.000 (-0.16)		
<i>CloudMachine</i> × <i>InIndex</i>	0.000 (0.70)			-0.001 (-0.64)			-0.001 (-0.41)		
<i>Machine</i>		0.003 (1.32)			0.005 (1.68)			0.007 (1.45)	
<i>Machine</i> × <i>InIndex</i>		-0.001 (-1.63)			-0.002 (-1.57)			-0.002 (-0.83)	
<i>Human</i>			0.003** (2.81)			0.004** (2.29)			0.006** (2.58)
<i>Human</i> × <i>InIndex</i>			-0.002** (-2.26)			-0.002** (-2.21)			-0.004** (-2.59)
<i>InIndex</i>	0.001 (0.47)	0.003* (1.77)	0.004** (2.17)	0.005* (1.98)	0.009** (2.49)	0.008** (2.59)	0.010* (1.97)	0.014* (1.81)	0.015** (2.44)
<i>BM</i>	0.003* (2.10)	0.003* (2.10)	0.002* (2.03)	0.004 (1.48)	0.004 (1.48)	0.003 (1.43)	0.011*** (3.16)	0.011*** (3.15)	0.010*** (3.20)
<i>SIZE</i>	-0.004** (-2.52)	-0.004** (-2.47)	-0.004** (-2.71)	-0.006*** (-3.14)	-0.006*** (-3.11)	-0.006*** (-3.18)	-0.005 (-1.16)	-0.005 (-1.12)	-0.005 (-1.25)
<i>ROA</i>	-0.033*** (-3.87)	-0.033*** (-3.97)	-0.033*** (-4.04)	-0.039*** (-4.37)	-0.039*** (-4.50)	-0.039*** (-4.48)	-0.053*** (-5.16)	-0.054*** (-5.21)	-0.053*** (-5.09)
<i>LEV</i>	0.004 (0.73)	0.004 (0.74)	0.003 (0.69)	0.010 (1.47)	0.010 (1.48)	0.010 (1.43)	0.012 (1.52)	0.013 (1.54)	0.012 (1.51)
<i>STD_{RET}</i>	0.046** (2.22)	0.046** (2.22)	0.044* (2.14)	0.061** (2.18)	0.061** (2.19)	0.059* (2.08)	0.105 (1.67)	0.105 (1.66)	0.101 (1.60)
<i>InstOwn</i>	-0.013** (-2.67)	-0.013** (-2.73)	-0.013** (-2.65)	-0.019*** (-3.67)	-0.019*** (-3.87)	-0.018*** (-3.78)	-0.019** (-2.21)	-0.019** (-2.31)	-0.018** (-2.17)
<i>Analysts</i>	0.001 (0.39)	0.000 (0.25)	0.000 (0.25)	0.000 (0.01)	-0.000 (-0.07)	-0.000 (-0.05)	-0.005 (-1.52)	-0.006 (-1.60)	-0.006 (-1.60)
Constant	0.080*** (7.30)	0.068*** (5.39)	0.074*** (7.45)	0.121*** (7.57)	0.102*** (6.17)	0.114*** (8.13)	0.144*** (3.95)	0.120*** (3.68)	0.138*** (3.95)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	33,653	33,653	33,653	33,653	33,653	33,653	33,653	33,653	33,653
Adj. R^2	0.158	0.159	0.160	0.169	0.170	0.170	0.194	0.195	0.196

Table 8: Viewership and informed trading

The table presents estimates from our analysis of the impact of machine (and sub-categories including *cloud machine* and *financial institution machine*) and human viewership activities on the average daily Probability of Informed Trading on days (0, 1) and (0, 5), relative to the 8-K publication date, as a validation test. Daily Average $PIN(0, 1)$ and Daily Average $PIN(0, 5)$ is the average of daily probability of information-based trading (PIN) over windows (0, 1) and (0, 5), relative to the 8-K publication date. *Machine* and *Human* denotes machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. *CloudMachine* is the machine viewership by cloud computing services, and *InstMachine* is the machine viewership by financial institutions. Control variables across all columns include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; *STD_{RET}* that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. Adj. R^2 is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Daily Average $PIN(0, 1)$				Daily Average $PIN(0, 5)$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Machine</i>	0.002** (2.54)				0.002** (2.95)			
<i>Human</i>		-0.002*** (-5.96)				-0.002*** (-4.68)		
<i>CloudMachine</i>			0.001* (1.88)				0.002** (2.39)	
<i>InstInvMachine</i>				-0.000 (-0.05)				0.000 (0.08)
<i>BM</i>	-0.002** (-2.55)	-0.002** (-2.39)	-0.002** (-2.29)	-0.002** (-2.29)	-0.002** (-2.72)	-0.002** (-2.60)	-0.002** (-2.52)	-0.002** (-2.52)
<i>SIZE</i>	-0.028*** (-23.11)	-0.028*** (-23.07)	-0.029*** (-23.60)	-0.029*** (-23.53)	-0.028*** (-24.39)	-0.028*** (-24.42)	-0.029*** (-25.18)	-0.029*** (-25.11)
<i>ROA</i>	-0.011** (-2.96)	-0.011*** (-3.08)	-0.011** (-2.92)	-0.011** (-2.91)	-0.010** (-2.87)	-0.010** (-2.98)	-0.010** (-2.99)	-0.010** (-2.98)
<i>LEV</i>	-0.012*** (-4.86)	-0.011*** (-4.48)	-0.011*** (-4.45)	-0.011*** (-4.45)	-0.011*** (-4.08)	-0.010*** (-3.77)	-0.010*** (-4.10)	-0.010*** (-4.10)
<i>STD_{RET}</i>	-0.042*** (-6.08)	-0.041*** (-6.00)	-0.044*** (-6.16)	-0.044*** (-6.17)	-0.038*** (-6.14)	-0.037*** (-6.05)	-0.040*** (-6.19)	-0.040*** (-6.19)
<i>InstOwn</i>	-0.098*** (-29.72)	-0.098*** (-29.29)	-0.099*** (-29.59)	-0.099*** (-29.54)	-0.097*** (-28.80)	-0.098*** (-28.37)	-0.098*** (-28.99)	-0.098*** (-28.96)
<i>Analysts</i>	-0.018*** (-14.19)	-0.018*** (-14.09)	-0.017*** (-12.05)	-0.016*** (-12.04)	-0.017*** (-13.61)	-0.017*** (-13.53)	-0.016*** (-11.33)	-0.016*** (-11.33)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	498,289	498,289	469,735	469,735	498,289	498,289	469,735	469,735
Adj. R^2	0.599	0.599	0.597	0.597	0.682	0.682	0.679	0.679

Table 9: Viewership and algorithmic trading

The table presents estimates from our analysis of the impact of machine (and sub-categories including *cloud machine* and *financial institution machine*) and human viewership activities on algorithmic trading (including the *OddLotRatio(0, 1)* and the *TradeSize(0, 1)*) as a validation test. *OddLotRatio(0, 1)* is the fraction of volume associated with abnormally small trades (less than 100 shares) over the day 0 and 1 relative to 8-K publication date. *TradeSize(0, 1)* is the number of shares traded divided by the number of trades over the day 0 and 1 relative to 8-K publication date. *Machine* and *Human* denote machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. *CloudMachine* is the machine viewership by cloud computing services, and *InstMachine* is the machine viewership by financial institutions. Control variables across all columns include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; *STD_{RET}* that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj. R²* is the adjusted *R²* value. *t*-statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>OddLotRatio(0, 1)</i>				<i>TradeSize(0, 1)</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Machine</i>	0.000 (0.19)				0.000 (0.45)			
<i>Human</i>		-0.009** (-2.93)				0.005*** (7.47)		
<i>CloudMachine</i>			0.005** (3.92)				-0.002** (-2.78)	
<i>InstMachine</i>				-0.002 (-0.57)				0.002 (1.67)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	165,821	165,821	165,821	165,821	165,821	165,821	165,821	165,821
Adj. <i>R²</i>	0.101	0.101	0.101	0.101	0.534	0.534	0.534	0.534

Table 10: Readability and price drift

The table presents estimates from our analysis of the impact of readability (*DifficultToRead*) proxies on the price drift post to 8-K filings date as a test to our information-assimilation channel. *Gunning FOG* and *Flesch-Kincaid* are readability measures computed based on the textual contents of a 8-K to proximate how hard the filing is to read. *DRIFT(2, 10)* and *DRIFT(2, 20)* are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *Machine* and *Human* denote machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. Control variables across all columns include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; *STD_{RET}* that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj. R²* is the adjusted *R²* value. *t*-statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>Gunning FOG</i>				<i>Flesch-Kincaid</i>			
	<i>DRIFT(2, 10)</i>		<i>DRIFT(2, 20)</i>		<i>DRIFT(2, 10)</i>		<i>DRIFT(2, 20)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Machine</i> × <i>DifficultToRead</i>	-0.006** (-2.48)		-0.003 (-0.75)		-0.006** (-2.52)		-0.004 (-0.96)	
<i>Machine</i>	0.004 (1.24)		0.005 (1.10)		0.004 (1.20)		0.005 (1.12)	
<i>Human</i> × <i>DifficultToRead</i>		-0.003 (-1.37)		-0.003 (-1.29)		-0.004 (-1.47)		-0.004 (-1.51)
<i>Human</i>		0.004*** (3.95)		0.005*** (3.79)		0.004*** (4.05)		0.005*** (3.90)
<i>DifficultToRead</i>	0.009 (1.76)	-0.001 (-0.19)	0.006 (0.69)	0.004 (0.84)	0.010* (1.91)	-0.001 (-0.32)	0.007 (0.85)	0.004 (0.89)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	517,848	517,848	517,848	517,848	517,848	517,848	517,848	517,848
Adj. <i>R²</i>	0.206	0.207	0.235	0.236	0.205	0.206	0.237	0.237

Table 11: Negative sentiment and price drift

The table presents estimates from our analysis of the impact of negative sentiment on the price drift post to 8-K filing date as a test to our information-assimilation channel. $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). *CloudMachine* is the machine viewship by cloud computing services. *InstMachine* is the machine viewship by financial institutions. *FinNeg* is the proportion of negative words defined by Loughran and Mcdonald (2011) in an 8-K filing. Control variables across all columns include: *BM* that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; *SIZE* that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; *ROA* that refers to the return on assets; *LEV* that refers to the total debt over total assets; STD_{RET} that refers to the standard deviation of monthly return over the year prior to an 8-K filing; *InstOwn* that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and *Analysts* that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. $Adj. R^2$ is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$DRIFT(2, 10)$		$DRIFT(2, 20)$		$DRIFT(2, 40)$	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>CloudMachine</i> × <i>FinNeg</i>	-0.033** (-2.83)		-0.058** (-2.77)		-0.071* (-1.87)	
<i>CloudMachine</i>	-0.002*** (-3.04)		-0.002* (-1.78)		-0.002 (-1.15)	
<i>InstInvMachine</i> × <i>FinNeg</i>		-0.056* (-2.06)		-0.095* (-2.13)		-0.131* (-2.00)
<i>InstInvMachine</i>		0.002 (1.59)		0.003 (1.38)		0.003 (1.47)
<i>FinNeg</i>	0.146*** (3.12)	0.155** (2.69)	0.288*** (3.21)	0.303** (2.77)	0.429*** (3.10)	0.470*** (3.08)
<i>BM</i>	0.003*** (7.96)	0.003*** (7.96)	0.005*** (6.16)	0.005*** (6.17)	0.007*** (6.47)	0.007*** (6.47)
<i>SIZE</i>	-0.007*** (-11.12)	-0.007*** (-11.20)	-0.010*** (-13.70)	-0.010*** (-13.95)	-0.016*** (-12.97)	-0.016*** (-13.10)
<i>ROA</i>	-0.024*** (-7.90)	-0.024*** (-7.94)	-0.037*** (-8.40)	-0.037*** (-8.47)	-0.058*** (-9.03)	-0.058*** (-9.08)
<i>LEV</i>	0.015*** (4.56)	0.015*** (4.55)	0.025*** (5.58)	0.025*** (5.58)	0.042*** (5.52)	0.042*** (5.51)
STD_{RET}	0.032*** (4.57)	0.032*** (4.54)	0.046*** (3.76)	0.045*** (3.75)	0.081*** (4.70)	0.080*** (4.68)
<i>InstOwn</i>	-0.009*** (-3.32)	-0.009*** (-3.33)	-0.014*** (-3.88)	-0.014*** (-3.90)	-0.017** (-2.41)	-0.017** (-2.41)
<i>Analysts</i>	-0.000 (-0.28)	-0.000 (-0.34)	-0.001 (-1.05)	-0.001 (-1.09)	-0.003* (-2.02)	-0.003* (-2.04)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	515,547	515,547	515,547	515,547	515,547	515,547
Adj. R^2	0.222	0.222	0.239	0.239	0.262	0.262

Table 12: Incremental earnings information and price drift

The table presents estimates from our analysis of the impact of Incremental Earnings Information on the price drift post to 8-K publication date as a test to our information-assimilation channel. $DRIFT(2, 10)$ and $DRIFT(2, 20)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). $Machine$, $Human$, and $CloudMachine$ denote machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. $Item2.02$ is a dummy variable assigned the value of one if a 8-K contain $Item2.02$, zero otherwise. Control variables across all columns include: BM that refers to the ratio of book value over the market value of common equity in the year end prior to an 8-K filing; $SIZE$ that refers to the natural logarithm of market value of common equity in the year end prior to an 8-K filing; ROA that refers to the return on assets; LEV that refers to the total debt over total assets; STD_{RET} that refers to the standard deviation of monthly return over the year prior to an 8-K filing; $InstOwn$ that refers to the institutional ownership percentage calculated in the quarter prior to an 8-K filing; and $Analysts$ that refers to the number of analyst covering a firm in the quarter prior to an 8-K filing. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. $Firm FE$ and $Year FE$ denote firm and year fixed effects. $Nobs$ refers to the number of observations. $Adj.R^2$ is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$DRIFT(2, 10)$			$DRIFT(2, 20)$			$DRIFT(2, 40)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Machine$	0.003 (1.02)			0.004 (1.04)			0.006 (1.26)		
$Machine \times Item2.02$	-0.000 (-0.10)			0.000 (0.20)			0.000 (0.52)		
$Human$		0.003*** (4.50)			0.004*** (4.17)			0.006*** (4.94)	
$Human \times Item2.02$		-0.001** (-2.76)			-0.001** (-2.17)			-0.001*** (-3.24)	
$CloudMachine$			-0.002*** (-3.36)			-0.002** (-2.19)			-0.003 (-1.49)
$CloudMachine \times Item2.02$			-0.000 (-0.70)			-0.000 (-0.10)			0.000 (0.53)
$Item2.02$	-0.001 (-0.76)	-0.001 (-1.16)	-0.001 (-0.85)	-0.004* (-1.79)	-0.004** (-2.90)	-0.005** (-2.54)	-0.009** (-2.56)	-0.007*** (-3.73)	-0.009*** (-3.18)
BM	0.004*** (6.38)	0.004*** (6.37)	0.004*** (6.21)	0.005*** (5.89)	0.005*** (5.80)	0.005*** (5.73)	0.007*** (5.68)	0.007*** (5.59)	0.007*** (5.85)
$SIZE$	-0.009*** (-11.68)	-0.009*** (-12.07)	-0.008*** (-10.91)	-0.012*** (-13.12)	-0.013*** (-13.58)	-0.012*** (-12.72)	-0.019*** (-13.20)	-0.019*** (-13.63)	-0.018*** (-12.58)
ROA	-0.030*** (-7.22)	-0.029*** (-7.32)	-0.029*** (-7.30)	-0.040*** (-8.53)	-0.040*** (-8.68)	-0.040*** (-8.50)	-0.061*** (-9.51)	-0.060*** (-9.62)	-0.063*** (-9.53)
LEV	0.019*** (4.80)	0.018*** (4.65)	0.017*** (4.43)	0.029*** (5.81)	0.028*** (5.70)	0.028*** (5.55)	0.046*** (5.59)	0.044*** (5.48)	0.045*** (5.40)
STD_{RET}	0.029*** (3.66)	0.028*** (3.54)	0.028*** (3.43)	0.046*** (3.84)	0.045*** (3.75)	0.044*** (3.44)	0.077*** (4.98)	0.076*** (4.87)	0.078*** (4.52)
$InstOwn$	-0.012*** (-4.17)	-0.011*** (-4.06)	-0.010*** (-3.43)	-0.017*** (-4.83)	-0.017*** (-4.69)	-0.016*** (-3.99)	-0.021*** (-3.25)	-0.020*** (-3.13)	-0.019*** (-2.56)
$Analysts$	0.000 (0.09)	0.000 (0.10)	-0.000 (-0.16)	-0.000 (-0.39)	-0.000 (-0.40)	-0.001 (-0.84)	-0.002 (-1.16)	-0.002 (-1.18)	-0.003* (-2.05)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	547,879	547,879	515,446	547,879	547,879	515,446	547,879	547,879	515,446
Adj. R^2	0.204	0.205	0.204	0.236	0.236	0.233	0.258	0.258	0.258

Appendix Table A: Variables, definitions, and sources

Variable	Definition	Source
<i>Panel A: Viewership Variables</i>		
<i>Machine</i>	The natural logarithm of non-human visits to an 8-K filing in day 0 and 1, defined by Dechow et al. (2015) .	SEC Log File
<i>Human</i>	The natural logarithm of human visits to an 8-K filing in day 0 and 1, defined by Dechow et al. (2015) .	SEC Log File
<i>TotalVisit</i>	The natural logarithm of both machine and human visits to an 8-K filing in day 0 and 1.	SEC Log File
<i>MachineDRT</i>	The natural logarithm of non-human visits to an 8-K filing in day 0 and 1, defined by Drake et al. (2015) .	SEC Log File
<i>HumanDRT</i>	The natural logarithm of human visits to an 8-K filing in day 0 and 1, defined by Drake et al. (2015) .	SEC Log File
<i>CloudMachine</i>	The natural logarithm of non-human visits (defined by Dechow et al. (2015)) to an 8-K filing in day 0 and 1 from cloud computing services.	SEC Log File, MaxMind, Thomson Reuters Ownership, & CIQ
<i>Machine/Total</i>	The ratio of machine viewership to total visits of 8K filings on days $t \in \{0, 1\}$ relative to 8K publication date.	Our estimates
<i>CloudMachine/Total</i>	The ratio of cloud machine viewership to total visits of 8K filings on days $t \in \{0, 1\}$ relative to 8K publication date.	Our estimates
<i>InstMachine</i>	The natural logarithm of non-human visits (defined by Dechow et al. (2015)) to an 8-K filing in day 0 and 1 from institutional investors.	SEC Log File, MaxMind, Thomson Reuters Ownership, & CIQ
<i>DataMachine</i>	The natural logarithm of non-human visits (defined by Dechow et al. (2015)) to an 8-K filing in day 0 and 1 from data and media vendor.	SEC Log File, MaxMind, Thomson Reuters Ownership, & CIQ
<i>OtherMachine</i>	The natural logarithm of non-human visits (defined by Dechow et al. (2015)) to an 8-K filing in day 0 and 1 from the IP we cannot identify entity type by Capital IQ, Thomson data, and Google.	SEC Log File, MaxMind, Thomson Reuters Ownership, & CIQ
<i>Panel B: Dependent Variables / Research outcome variables</i>		
<i>DRIFT(2, T)</i>	The absolute difference between post-announcement price variation $CAR(0, T)$ and short variation right after the announcement $CAR(0, 2)$, where cumulative abnormal returns (CAR) are from Fama French 3 Factor + Momentum model.	CRSP & Compustat
<i>PIN(0, T)</i>	Average of daily probability of information-based trading (PIN) over window (0, T). The PIN is calculated daily over 5-minute interval	TAQ
<i>Noise</i>	The share of stock return variance that is attributable to noise.	CRSP & Compustat
<i>OddLotRatio(0, 1)</i>	The fraction of volume associated with abnormally small trades (less than 100 shares) over the day 0 and 1 relative to 8-K publication date.	SEC MIDAS
<i>TradeSize(0, 1)</i>	The number of shares traded divided by the number of trades over the day 0 and 1 relative to 8-K publication date.	SEC MIDAS

Appendix Table A: Variables, definitions, and source (Continued)

Variable	Definition	Source
Panel C: 8-K variables		
<i>#Item</i>	Number of topics included in an 8-K filing	WRDS SEC Suite
<i>WordCount</i>	Number of words used in an 8-K filing	WRDS SEC Suite
<i>FinNeg</i>	The proportion of negative words defined by Loughran and McDonald (2011) in an 8-K filing.	WRDS SEC Suite
<i>DayRelease</i>	The number of days between 8-K event date to publication date	WRDS SEC Suite
<i>Item#-#</i>	A dummy variable equal to one if an 8-K filing containing a 8-K item#-#, zero otherwise.	WRDS SEC Suite
Panel D: Control Variables		
<i>LEV</i>	Total debt over total assets	Compustat
<i>BM</i>	The ratio of book value over the market value of common equity in the year end prior to an 8-K filing	CRSP & Compustat
<i>SIZE</i>	The natural logarithm of market value of common equity in the year end prior to an 8-K filing	CRSP
<i>InstOwn</i>	The institutional ownership percentage calculated in the quarter prior to an 8-K filing	Factset 13F
<i>Analysts</i>	The natural logarithm of one plus the number of analyst covering a firm i in the quarter prior to the firm i 's 8-K filing date	I/B/E/S
<i>STDRET</i>	Standard Deviation of Monthly Return over the year prior to an 8-K filing	CRSP
<i>InIndex</i>	A dummy variable equal to one if a stock has been added into S&P500 index, zero otherwise.	CRSP
Panel E: Instrumental Variables		
<i>InvSent</i>	Investor sentiment index based on first principal component of FIVE (standardized) sentiment proxies where each of the proxies has first been orthogonalized with respect to a set of six macroeconomic indicators	Jeffrey Wurgler's website
<i>MacroNews</i>	A dummy variable equal to one if a major macroeconomic indicator update is announced the 8-K publication data, zero otherwise.	Bloomberg
<i>VIX(-20, -1)</i>	Return based on VIX level over window (-20, -1) relative to 8-K publication day	OptionMetrics
<i>CRET(-20, -1)</i>	CRSP value weighted return over window (-20, -1) relative to 8-K publication day	CRSP
Panel F: Disruption and Readability Measures		
<i>Cloud Outage Day</i>	A day characterised by a reported service disruption to one or more major cloud service providers.	Gunawi et al. (2016)
<i>Power Outage Day</i>	A day characterised by loss of the electrical power network supply to an end user.	Mukherjee et al. (2018)
<i>Gunning FOG</i>	The Gunning fog index, a readability measure, is computed based on the words used in an 8-K filing.	WRDS SEC Suite
<i>Flesch-Kincaid</i>	Flesch-Kincaid grade level designed to indicate how difficult a passage in English is to understand.	WRDS SEC Suite

Appendix Table B: Top 8-K Viewers by Organisation Types

Organization Type	Organization Name	# Visits (Millions)
Cloud Computing	AMAZON.COM	5.635
Cloud Computing	1&1 INTERNET AG	1.836
Cloud Computing	VICTORY NETWORKS	1.476
Cloud Computing	SOFTLAYER TECHNOLOGIES	1.098
Cloud Computing	FUSIONSTORM	0.482
Cloud Computing	CORESITE	0.320
Cloud Computing	RACKSPACE LTD.	0.300
Cloud Computing	DIGITAL OCEAN	0.274
Cloud Computing	NET ACCESS CORPORATION	0.203
Cloud Computing	SUNGARD AVAILABILITY NETWORK SOLUTIONS	0.181
Institutional Investor	BARCLAYS CAPITAL	0.656
Institutional Investor	TWO SIGMA INVESTMENTS, LLC	0.496
Institutional Investor	THE GOLDMAN SACHS GROUP	0.403
Institutional Investor	HUTCHIN HILL CAPITAL	0.285
Institutional Investor	SCHONFELD TOOLS, LLC.	0.236
Institutional Investor	HARTFORD LIFE INSURANCE COMPANY	0.230
Institutional Investor	HUTCHIN HILL CAPITAL, LP	0.222
Institutional Investor	D. E. SHAW & CO.	0.180
Institutional Investor	WILLIAM O'NEIL & COMPANY	0.164
Institutional Investor	KNIGHT CAPITAL GROUP	0.098
Internet Provider	CENTURYLINK	1.765
Internet Provider	COMCAST	1.042
Internet Provider	SPECTRUM	0.922
Internet Provider	VERIZON	0.584
Internet Provider	OPTIMUM ONLINE	0.523
Internet Provider	AT&T SERVICES	0.518
Internet Provider	OXFORD NETWORKS	0.411
Internet Provider	ROGERS CABLE	0.357
Internet Provider	ILIAD-ENTREPRISES	0.331
Internet Provider	RESILANS AB	0.330
Rest/Other	DOW JONES & COMPANY	0.588
Rest/Other	MARKIT ON DEMAND	0.366
Rest/Other	BLOOMBERG, LP	0.258
Rest/Other	MORRIS, NICHOLS, ARSHT AND TUNNEL	0.221
Rest/Other	THOMSON REUTERS U.S. LLC	0.147
Rest/Other	MCGRAW-HILL	0.127
Rest/Other	CHENGXI MIDDLE SCHOOL	0.061
Rest/Other	GOOGLEBOT*	0.060
Rest/Other	REGUS GROUP SERVICES LTD	0.053
Rest/Other	GODADDY.COM, LLC	0.049

* Googlebot is the generic name for Google's web crawler.

Appendix Table C: Alternative Measure of Machine Viewership

This table reports OLS regression estimates from our analysis of the alternative measures of machine and cloud machine viewership activities of firm i 's 8-K filings on days $t \in \{0, 1\}$ (relative to the 8-K publication date) on $DRIFT(2, 10)$, $DRIFT(2, 20)$, and $DRIFT(2, 40)$. $DRIFT(2, 10)$, $DRIFT(2, 20)$ and $DRIFT(2, 40)$ are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). $Machine/Total$ and $CloudMachine/Total$ denote viewership fraction from machine and cloud machine, and $TotalVisit$ is the logarithm of total viewership. BM is the ratio of book value over the market value of common equity in the year end prior to an 8-K filing, $SIZE$ is the natural logarithm of market value of common equity in the year end prior to an 8-K filing, ROA is the return on assets, LEV is the total debt over total assets, STD_{RET} is the standard deviation of monthly return over the year prior to an 8-K filing, $InstOwn$ is the institutional ownership percentage calculated in the quarter prior to an 8-K filing, and $Analysts$ is the number of analyst covering a firm in the quarter prior to an 8-K filing, are control variables across all columns. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. $Firm FE$ and $Year FE$ denote firm and year fixed effects. $Nobs$ refers to the number of observations. $Adj.R^2$ is the adjusted R^2 value. t -statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$DRIFT(2, 10)$		$DRIFT(2, 20)$		$DRIFT(2, 40)$	
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine/Total$	-0.010** (-2.18)		-0.011 (-1.61)		-0.015 (-1.63)	
$CloudMachine/Total$		-0.017** (-2.36)		-0.020* (-1.98)		-0.023 (-1.50)
$TotalVisit$	0.003 (1.14)	0.003 (1.45)	0.005 (1.23)	0.006 (1.51)	0.008 (1.52)	0.008* (1.91)
BM	0.003*** (6.64)	0.003*** (6.68)	0.005*** (5.58)	0.005*** (5.54)	0.007*** (5.63)	0.007*** (5.67)
$SIZE$	-0.008*** (-12.93)	-0.008*** (-13.32)	-0.012*** (-13.58)	-0.012*** (-13.82)	-0.018*** (-13.41)	-0.018*** (-13.57)
ROA	-0.026*** (-7.91)	-0.026*** (-8.01)	-0.039*** (-9.24)	-0.040*** (-9.36)	-0.061*** (-9.68)	-0.062*** (-9.74)
LEV	0.015*** (4.19)	0.015*** (4.15)	0.027*** (5.53)	0.027*** (5.51)	0.045*** (5.51)	0.045*** (5.49)
STD_{RET}	0.031*** (4.66)	0.031*** (4.70)	0.045*** (3.83)	0.046*** (3.85)	0.079*** (4.62)	0.079*** (4.66)
$InstOwn$	-0.010*** (-3.45)	-0.009*** (-3.44)	-0.016*** (-4.12)	-0.015*** (-4.12)	-0.019** (-2.66)	-0.019** (-2.65)
$Analysts$	-0.000 (-0.35)	-0.000 (-0.33)	-0.001 (-1.19)	-0.001 (-1.17)	-0.003** (-2.21)	-0.003** (-2.19)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	517,848	517,848	517,848	517,848	517,848	517,848
Adj. R^2	0.218	0.219	0.237	0.238	0.260	0.260

Appendix Table D: Instrumental variables approach

The table presents estimates from our 2-Stage Least Squares (2SLS) analysis using four market sentiment metrics as instrumental variables. *Machine* and *Human* denote machine and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. *DRIFT*(2, 10), *DRIFT*(2, 20) and *DRIFT*(2, 40) are measured by the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as described in Equation (2). Our instruments are: *InvSent*, *MacroNews*, *VIX*(-20, -1), and *CRET*(-20, -1). *BM* is the ratio of book value over the market value of common equity in the year end prior to an 8-K filing, *SIZE* is the natural logarithm of market value of common equity in the year end prior to an 8-K filing, *ROA* is the return on assets, *LEV* is the total debt over total assets, *STD_RET* is the standard deviation of monthly return over the year prior to an 8-K filing, *InstOwn* is the institutional ownership percentage calculated in the quarter prior to an 8-K filing, and *Analysts* is the number of analyst covering a firm in the quarter prior to an 8-K filing, are control variables across all columns. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. Weak ID *F*-stat is for the Stock-Yogo test and *t*-statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	First Stage			Second Stage		
	<i>Machine</i>	<i>CloudMachine</i>	<i>Human</i>	<i>DRIFT</i> (2, 10)	<i>DRIFT</i> (2, 20)	<i>DRIFT</i> (2, 40)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Human</i>				0.091** (2.42)	0.125** (2.34)	0.164** (2.80)
<i>InvSent</i>	-0.256 (-1.08)	0.381*** (2.84)	-0.223* (-1.89)			
<i>MacroNews</i>	-0.021 (-1.42)	-0.004 (-0.46)	-0.057*** (-4.41)			
<i>VIX</i> (-20, -1)	0.003 (0.06)	0.120 (1.38)	-0.102* (-1.72)			
<i>CRET</i> (-20, -1)	-0.618 (-1.37)	0.885* (1.72)	-0.742*** (-2.71)			
<i>BM</i>	-0.000 (-0.11)	-0.006 (-1.24)	0.045*** (6.84)	-0.001 (-0.49)	-0.001 (-0.32)	-0.000 (-0.18)
<i>SIZE</i>	0.004 (0.51)	-0.004 (-0.55)	0.055*** (4.59)	-0.013*** (-6.12)	-0.019*** (-6.00)	-0.027*** (-7.71)
<i>ROA</i>	0.009 (0.50)	0.028 (1.58)	-0.139*** (-3.30)	-0.014** (-2.60)	-0.022** (-2.94)	-0.037*** (-3.90)
<i>LEV</i>	0.009 (0.33)	0.004 (0.19)	0.314*** (6.24)	-0.012 (-1.09)	-0.010 (-0.70)	-0.005 (-0.35)
<i>STD_RET</i>	0.007 (0.23)	0.050* (1.85)	0.286*** (5.38)	0.008 (0.63)	0.013 (0.66)	0.034 (1.38)
<i>InstOwn</i>	0.005 (0.14)	-0.037 (-1.12)	-0.121*** (-2.73)	0.000 (0.06)	-0.002 (-0.21)	-0.001 (-0.12)
<i>Analysts</i>	0.001 (0.16)	0.009 (1.25)	0.005 (0.50)	-0.000 (-0.14)	-0.001 (-0.85)	-0.002 (-1.34)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	546,949	516,909	546,949	546,949	546,949	546,949
Weak ID F-Value	1.322	3.976	42.329	-	-	-
Hansen J-Stats	-	-	-	0.128	0.118	0.102

Appendix Table E: Impact of viewership on price variance due to noise trading

This table reports OLS regression estimates from our analysis of the impact of *Machine*, *Human* and *CloudMachine* viewership activities on *Noise*. *Noise* (the dependent variable) is the share of stock return variance that is attributable to noise, which is developed by Brogaard et al. (2020). *Machine*, *Human* and *CloudMachine* denote machine, human, and cloud machine viewership activities, respectively, of firm *i*'s 8-K filings on days $t \in \{0, 1\}$. *BM* is the ratio of book value over the market value of common equity in the year end prior to an 8-K filing, *SIZE* is the natural logarithm of market value of common equity in the year end prior to an 8-K filing, *ROA* is the return on assets, *LEV* is the total debt over total assets, *STD_{RET}* is the standard deviation of monthly return over the year prior to an 8-K filing, *InstOwn* is the institutional ownership percentage calculated in the quarter prior to an 8-K filing, and *Analysts* is the number of analyst covering a firm in the quarter prior to an 8-K filing, are control variables across all columns. All variables are winsorized at the 1st and 99th percentiles. All variables are also defined in the Appendix Table A. *Firm FE* and *Year FE* denote firm and year fixed effects. *Nobs* refers to the number of observations. *Adj. R²* is the adjusted R-squared value. *t*-statistics are reported in parentheses and computed based on adjusted standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	<i>Noise</i>		
	(1)	(2)	(3)
<i>Machine</i>	0.116 (0.94)		
<i>Human</i>		0.101*** (4.38)	
<i>CloudMachine</i>			-0.081** (-2.38)
<i>BM</i>	0.127*** (8.08)	0.123*** (8.05)	0.127*** (8.12)
<i>SIZE</i>	-0.243*** (-8.62)	-0.247*** (-8.80)	-0.242*** (-8.39)
<i>ROA</i>	-0.642*** (-8.53)	-0.628*** (-8.64)	-0.640*** (-8.44)
<i>LEV</i>	0.576*** (4.90)	0.547*** (4.76)	0.575*** (4.85)
<i>STD_{RET}</i>	0.771*** (3.05)	0.746** (2.96)	0.776*** (3.05)
<i>InstOwn</i>	-0.196** (-2.52)	-0.186** (-2.39)	-0.197** (-2.50)
<i>Analysts</i>	-0.000 (-0.01)	-0.002 (-0.07)	0.001 (0.04)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Nobs	455,607	455,607	455,607
Adj. <i>R</i> ²	0.486	0.487	0.485