Market Efficiency When Machines Access Information*

Leonidas G. Barbopoulos† Rui Dai‡ Tālis J. Putniņš§ Anthony Saunders¶

This version
August 22, 2023

*We are grateful for comments and suggestions from Simona Abis (EFA discussant), Yakov Amihud, Seth Armitage, Sandra Betton (FMA discussant), Bruce I. Carlin, Ranjan Chakravarty, Gjergji Cici, Jo Danbolt, Rayan L. Devis, Xavier Gabaix, Angelica Gonzalez, Alistair Haig, Pia Helbing, Ryan D. Israelsen (AFA discussant), Pantelis Kazakis, Paul-Olivier Klein, Woon Sau Leung, Richard M. Levich, Elisa Luciano, Andrew Marshall, Georgios Panos, Matthew Richardson, Khaladdin Rzayev, Alexi Savov, Georgios Sermpinis, Charalampos Stasinakis, Julian Thimme, Jeffrey Wurgler, David Yermack, and participants in the seminar series of Adam Smith Business School (University of Glasgow), University of Bergamo (Bergamo, Italy), Cyprus University of Technology (Limassol, Cyprus), Franklin Templeton Investments (virtual), iaelyon School of Management (Lyon, France), NYU Stern School of Business (NY, USA), Surrey Business School (Surrey, UK), UBS Quant Finance Insights (virtual), and University of Western Australia (Perth, Australia), as well as the American Finance Association (AFA, 2022), the European Finance Association (EFA, 2021), the AI & Machine Learning in Finance conference, Swedish House of Finance, SSE (2022), the Financial Management Association Annual Meeting (2021), the Financial Management Association European Meeting (2022), the Financial Markets and Corporate Governance Conference (Monash University, 2022), the 7th International Conference on Accounting and Finance (keynote, 2022), the 15th International Risk Management Conference (2022), the third One-Day International Workshop on Machine Learning for Finance (Ca' Foscari University of Venice, 2021), the 8th SAFE Asset Pricing Workshop (2021), and the World Finance Conference (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 When Machines Access Information

Abstract

As machines replace humans in financial markets, how is informational efficiency impacted? We shed light on this issue using a unique dataset that allows us to separately identify when machines and humans access company information (8-K filings). We find that increased information access by cloud computing services significantly improves informational efficiency, reducing price drift (or reversal) and noise return variance following information events. We address identification through exogenous cloud outages, a quasi-natural experiment, and instrumental variables. We show that machines are better at handling numerical information, are less biased by negative sentiment, and are less capacity constrained. Conversely, humans can better handle sequential and soft information.

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, and 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) can interpret is lost, and consequently, prices become less informative. While humans make errors in interpreting information, machines are also not infallible and 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 (e.g., robots, algorithms)[2] on informational efficiency, with a unique dataset in which we can distinguish between when a machine reads information released by a company versus when a human accesses the same information. Specifically, we focus on machine viewership from cloud computing services, as anecdotal evidence indicates that sophisticated investors tend to deploy advanced algorithms from the cloud, which has been shown to boost processing efficiency and artificial intelligence capabilities.[3] We use server log files via the SEC system, which records viewership of corporate events disclosed in Form 8-K filings, the means through which firms notify the market of important unscheduled corporate events, including changes to material agreements, financial information, M&A, and any other events deemed material to shareholders.[4]

---

[1]This question has also attracted significant attention from the business press. For example, an article in *The Economist* (October 5, 2019) highlights concerns arising from increased robotisation of financial markets: "The rise of financial robotisation is not only changing the speed and makeup of the stock market. It also raises questions about the function of markets, the impact of markets on the wider economy, how companies are governed and financial stability."

[2]According to PwC, in 2020 cloud computing services were among the biggest FinTech trends, with public cloud (a virtual space on common hardware) poised to become the dominant infrastructure model.

[3]For example, at the Amazon cloud computing event (re:Invent) in 2019, Vanguard's CTO, Jeff Dowds, stated that system infrastructure, machine learning facilities, etc. at Amazon cloud services allow Vanguard to increase the speed of building algorithms by 30%, and deploy those algorithms 30 times faster, resulting in a better ability to innovate along the way to improve resiliency.

[4]We use the terms 'viewership', 'readership,' and 'downloading' of 8-K filings interchangeably throughout.

What distinguishes our paper from other studies of these company filings is that we separate each viewing of an 8-K filing into machine viewers and human viewers based on the downloading behavior captured by the server logs.[5] We further partition viewership at the level of the viewer's computing facility, using IP addresses. Unlike previous research that focused on the role of machine automation in accelerating price discovery,[6] this study aims to directly compare the pricing efficiency associated with information acquisition venues, namely algorithms via the cloud and human viewership. No prior study, to the best of our knowledge, has examined material information acquisition at this granular level. We then examine how information viewership, and subsequent trading, impacts the efficiency of the price reaction to the information release.

First, we show that cloud computing viewership of 8-K filings has experienced significant growth in recent years.[7] Specifically, the viewership of 8-K filings from cloud computing services, such as Amazon Web Services (AWS), was only 1% of the total machine viewership before 2008, whereas this figure increased to 14% in 2008, to 35% in 2012, and to 62% in 2016. Similarly, the total machine viewership of 8-Ks increased from 0.57 million in 2008 to 8.78 million in 2016, while the total human viewership of 8-K remained stable throughout the sample period 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.[8]

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

[5]Following 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.

[6]For example, Brogaard et al. (2014) show that high-frequency trading improves market efficiency, yet at intraday horizons. More recently, Chakrabarty et al. (2022) find that high-frequency trading improves price efficiency following earnings announcements. Baldauf and Mollner (2021) find that more fast traders are associated with faster price adjustment following unanticipated news. von Beschwitz et al. (2020) also show that computer-readable news (news analytics) speeds up stock price adjustment following a news release.

[7]An 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."

[8]Due 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) report that 75% of trading volume in the US in 2009 was conducted through computer-based trading.

information. Where humans and machines differ is that humans show a tendency to disproportionately allocate their attention by accessing the 8-K of large firms and value firms, consistent with a need to ration attention due to cognitive and capacity constraints. By 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 with a greater information processing capacity. We also find that on Fridays, human viewership of 8-K is significantly lower, whereas viewership from cloud machines remains unaffected, consistent with existing studies pointing to limited human attention on Fridays (Dellavigna and Pollet, 2009).

A noticeable difference between machines and humans that helps explain our findings is the type of information that they choose to focus on. A useful feature of 8-K filings is that they arrange events into topical categories. We find that while human viewers pay significantly more attention to earnings information ($Item2.02$), machine viewers pay more attention to corporate events not listed by the SEC but deemed to be material non-public information by the firm ($Item8.01$). 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 (and humans) on informational efficiency, we measure the magnitude of absolute price drift (or reversal) following the 8-K publication date over horizons of 10 and 20 days, hereinafter referred as PI or *price inefficiency*, and relate its extent to human and machine viewership.[9] Our key finding is that more viewership by cloud-based machines tends to decrease PI, consistent with more efficient price discovery around the information release events. By contrast, relatively more human viewership of 8-K is significantly associated with increased PI following the 8-K publication date, consistent with an inefficient assimilation of the information. This evidence is robust to the inclusion of various controls and fixed effects, and remains qualitatively similar to alternative ways of defining *price inefficiency*. The results are also robust to an alternative measure of informational efficiency based on a variance decomposition that separates noise

[9] While companies can publish the 8-K Form to the SEC at any time within four days following the actual event, we keep in our analysis only 8-K that have been published and downloaded on the same day of the event in order to avoid conflating effects from information that may be released between the actual event date and the 8-K publication date.

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 PI over the month following the 8-K filing date, whereas a one standard deviation increase in human viewership (by holding fixed the level of machine viewership) yields a 13.23% aggregate increase in PI.

We find that not all machines are beneficial to informational efficiency. We partition machine viewership (using IP addresses) into six categories, namely: 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 to machine-based investment strategies. The beneficial effects of machines on price discovery stem almost entirely from 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 PI, 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 outages within an instrumental variable framework, and a quasi-natural experiment. As a further strategy, we also consider evidence from instrumental variables that impact human viewership. In our first approach, we estimate two-stage least-squares (2SLS) models in which cloud machine viewership is instrumented with exogenous cloud outages, which can directly affect the cloud machine ability to access and assimilate information in prices. 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 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 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 improve informational efficiency.

What is the channel through which machine access to information leads to more efficient price reactions? We conjecture it is through trading – machines use the information in quantitative investment strategies and decision models, leading to trading by the machines, which drives the information to become impounded in prices. We find support for this channel. Machine viewership is significantly positively related to measures of informed trading following the 8-K publication date, whereas for human viewership we get the opposite result.

We also examine whether the effects of machines accessing information are distinct from the impacts of high-frequency traders (HFT), which also rely on algorithms to automate trading decisions. We augment our tests of how machine access to information impacts efficiency by placing two measures of HFT activity on the right-hand side. Consistent with existing literature on HFTs (e.g., Carrion (2013); Brogaard et al. (2014)), we find that one of the proxies of HFT activity is associated with more efficient price reactions. Importantly, however, the positive impacts of cloud machine access to 8-K filings on informational efficiency is not diminished by controlling for HFT activity. These results suggest that the effects in this paper are different to what has been previously found for HFT. Taken together, the literature on HFT and this paper suggest machines can benefit informational efficiency in multiple ways and at different horizons. While HFT studies suggest that HFT improve intraday price discovery, largely by inferring the information in order flow (e.g., large orders of institutional investors sliced into a series of trades), i.e., the information gathering activities of other investors, this paper suggests machines also play a role in accessing new corporate information and then more efficiently incorporating it into prices at daily horizons.

Finally, we turn to the issue of determining the settings in which humans (or 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 settings humans may still have the upper hand. We find that the advantage of machines increases for highly numerical "hard" information. Machines are less constrained in terms of attention and processing capacity, compared to humans. We also find that cloud-based machines are less susceptible to bias when processing information that has negative sentiment content compared to humans, consistent with the notion that emotions can impede human decision-making. 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. By contrast, repeated information has no significant effect on machines, as they are more likely to process this information when it first arrives. Our findings are also consistent with theory that humans specialize in slower decision-making that combines sequential information, whereas machines have an edge in making quick decisions based on isolated information.

Our paper makes three 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 the 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 to improve market efficiency.

Our results provide insight on the regulators' efforts to mandate machine-readable disclosures, such as XML and XBRL, by providing unique evidence on the effect of various venues of data access on market efficiency. Prior studies (e.g, Blankespoor et al., 2014; Gao and Huang, 2020) suggest greater machine-friendly information may benefit more sophisticated investors. Our findings echo this argument in that financial institutions behind the cloud can leverage advanced technology to improve market efficiency. Nonetheless, the market-wide gain may come at the price of a relative increase in informational barriers and, as a result, information asymmetry among less sophisticated investors who rely on conventional information acquisition means.

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). So too do HFT, although their effects are largely attributable to speeding up price discovery with respect to information brought to the market by other investors (Carrion, 2013; Brogaard et al., 2014). 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, mitigate the impairment of market efficiency caused by analyst coverage reductions. Begenau et al. (2018) argue that certain investors can process larger amounts of data more effectively and therefore help to price assets of firms with big data more accurately. Our paper expands this array of studies by showing that information acquisition by machine-based investment strategies impacts post-announcement drifts or reversals differently from information acquisition by humans.

2. Theoretical Framework

Several branches of the literature help us to 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, as well as studies that examine who access the information content of 8-K filings and how that impacts price formation.[10]

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; hence, investor efforts in becoming informed are compensated (Grossman and Stiglitz, 1980; Kyle, 1985, 1989; Admati and Pfleiderer,

[10] A growing literature utilizes Form 8-K filings to analyze when and who receives information. Early studies employing large samples of 8-K reveal a more substantial market reaction on event day than on filing day (Carter and Soo, 1999; Lerman and Livnat, 2010). Recently, Ben-Rephael et al. (2021) found that price discovery generally occurs on the event day and the days between the event date and the filing date, albeit 8-K filings may have a limited informational benefit, particularly for retail investors. Watkins (2022) reports a decreased in length and quantitative information after the SEC expanded and accelerated 8-K filings.

1988).[11] A direct consequence is that a reduction in the cost of information acquisition and processing will tend to increase the amount of information acquired 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 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 quantitative investment 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. Machines are also likely to be faster than humans, especially for demanding and complex tasks (Sims, 2003, 2006; Bradshaw et al., 2020).[12] 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 to stock prices, lowering market efficiency. Humans also trade for non-investment reasons.[13] 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

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

[12] Bradshaw 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 "hard" news as the latter is more likely to have already been factored into prices.

[13] For 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).

equity ownership in 2019.[14]

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 increases informational efficiency (relative to human viewership) and decreases the price inefficiency* (PI) *following information releases.*

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 biases, under-performing traditional methods in the long run.[15] 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. 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 from slower and more deliberated decision-making. 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,

[14] Anonymous. "March of the machines. The stock market is now run by computers" *The Economist*, Oct 5th 2019.
[15] For example, see Mark Hulbert (2020, Jan 5). Use AI for Picking Stocks? Not So Fast. *Wall Street Journal.* Retrieved from http://online.wsj.com.

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 HFT 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 were better able to predict the news of the upcoming earnings releases, consistent with improved price discovery.[16] 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.

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 news of firms with big data. They find that large firms benefit disproportionately from supplying big data that allows 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 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 "hard" (quantitative, numerical) information versus "soft" (qualita-

[16] On 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.

tive, textual) information. We conjecture that machines have a comparative advantage in processing hard information, whereas humans may have a comparative advantage in interpreting soft information.

Another strand of the literature suggests that investor sentiment affects stock prices. Tetlock (2007) uses 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 unresolved. We anticipate that the positive impact of machines on stock price efficiency is stronger when the information disclosure contains negative sentiment, which is when human decisions are most likely to be biased by sentiment.

Other recent studies have theoretically investigated whether lower information processing cost and increased information transparency leads to more precise signals. Dugast and Foucault (2018) demonstrate 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 deeper analyses of sequential information pieces. Hence, we expect that machines will not improve stock price efficiency and could even harm it when it comes to the sequential arrival of information that lends itself to deeper analysis by human experts.

3. Data and Summary Statistics

3.1. Data sources and identifying machines vs humans

Our sample includes all firms in the CRSP and Compustat universe during a 14-year period from January 2003 to December 2016.[17] We collect data on each viewing (referred to as a visit) of each

[17] At the time of writing, the 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.

8-K filing (a notification to investors of an unscheduled material or extraordinary event that is important to shareholders) from the SEC's EDGAR server, giving us nearly 4 billion visits.[18] Each observation contains a partially anonymized IP address, time stamp, HTTP status codes (e.g., 200 for successful delivery), and crawler flag.[19] We obtain 8-K filings from WRDS SEC Analytics Suite (approximately 1.2 million 8-K filings). We restrict our sample to 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%).[20] 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 dataset 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", "Cloud Computing Services", "Media & Data Vendors", "Educations & Regulators", "Internet Services Providers," and "Others."[21]

[18]To 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 the EDGAR server (HTTP code≠200), only visit the index page (index dummy=1), or that come from search engines and other similar web crawlers (crawler=1).

[19]The 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).

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

[21]To 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 entitles), 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' websites obtained from the first ten Google

We use two approaches to separate human viewership from machine viewership. We first follow Drake et al. (2015) to define visits from IP addresses with more than 5 downloads per minute, 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 more than 25 downloads per minute, 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 event date, 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 (e.g., $Item9.01$). 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., 2021). To focus on the 8-K filings that carry new information, we limit the sample to filings that contain events that are released within the same business day of the event date.[22]

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 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, and cloud service disruption event data from Gunawi et al. (2016).[23] To calculate informed trading and HFT 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 Appendix Table A.

search hits. Our manual search allows us to reduce 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 for less than 3.86% of total 8-K visits in our final sample. Appendix Table B reports the top ten organizations of "Cloud Computing", "Institutional Investors", "Internet Provider Services", and "Others."

[22]On 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 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.

[23]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 2016 through Ravenpack databases. We drop outage events related to gaming (e.g., Playstation), business storage (e.g., Salesforce), and social media (e.g., Facebook), resulting in a collection of Infrastructure as a Service (IaaS) outage events. We ended up with 60 cloud outage incidents, most of which, based on our manual search, lasted less than 24 hours.

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 the post-event *price inefficiency* (PI), which is net of predicted returns from a factor model. We proceed by first computing the cumulative abnormal return, as follows,

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

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 event information that enters prices through the announcement day to the day T. To measure the PI 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 following the 8-K publication date. Specifically, we use the absolute difference between post-announcement price variation and variation over a short window right after the announcement, as follows,

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

with $T > 1$ to capture post-announcement *price inefficiency*.[24] We also examine a scaled version of the PI measure, normalizing it by the total information content of the information release, as follows,

$$ScaledPI(2, T) = \left| CAR_{i,t}^{0,T} - CAR_{i,t}^{0,1} \right| \Big/ CAR_{i,t}^{0,T} \tag{3}$$

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. (2022). This measure is based on a variance decomposition model that partitions stock return variance into information of various

[24] Our metric, like Weller (2017), is more closely connected to "informativeness" than the "informational efficiency" proposed by Brunnermeier (2005). In other words, this metric reveals how informative the pricing process is in absolute terms rather than the extent to which information dispersed across multiple traders may be inferred from the price (process) in conjunction with other publicly available information.

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, and the 25th, median, and 75th percentiles of our key variables. To focus on the immediate viewership right after the 8-K publication date, we restrict our sample to visits on the SEC's 8-K filing day and the day following ($t \in \{0, 1\}$). Because companies can file the 8-K Form to the SEC within four days following the date of the actual event, we also restrict our sample to 8-K published on the event day to avoid potential conflating effects from information that could be released between the event date and the 8-K publication date.[25] This leaves us with 295,540 8-K filings with all main control variables available. Panel A presents the main dependent variables used in our analyses, including the measures of PI, $ScaledPI$, $Noise$, and PIN. Among other notable statistics reported in Panel B, machines view more 8-K filings (log viewership of 3.47 and 3.52, based on Dechow et al. (2015) and Drake et al. (2015) methods, respectively), on average, compared to humans (log viewership of 1.46 and 1.14, 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-K than humans. Among other statistics reported in Panel C, the average firm size is about \$6.63m, while the average book-to-market ratio is 0.75. The average institutional holding is 63%, with 1.72 analysts covering a given firm, on average. Panel D reports viewership divided into six item types that appear in 8-K filings. The most frequently type of item viewed is $Item2.02$, which is largely related to quarterly and annual earnings announcements. The mean (median) 2-day viewership ($t \in \{0, 1\}$) of a filing with $Item2.02$ is 69.91 (41) views.

Panel E reports viewership categorised by viewer type, using seven categories defined in Section (*3.1*). Viewership from "Cloud Computing" and "Internet Provider" are considerably greater than those from the remaining types. The proportion of machine viewership from Cloud Computing is

[25] When we study the determinants of the 8-K viewership we drop this restriction and we control for the number of days between the actual event date and the 8-K publication date.

98.33%, the highest among all seven categories.[26] 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 – a limited number of facilities serve a big number of active viewers.

4. What Information is Accessed by Machines vs Humans?

First, we examine how machine and human access to information has changed over 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. It is important to note that we chose a two-day event window, defined shortly, to capture not only the visits from machines or analysts constantly monitoring the EDGAR system but also the later traffic initiated by social/press media, market activity, and various less-frequent information retrieval programs, triggered by proprietary conditions and algorithms. Indeed, if the bulk of viewers in the first two days are constant watchers who arrived within the first few seconds after the 8-K publication, there should be no correlation between the filing characteristics and the number of viewers.

4.1. Human and machine information access over time

Figure (1) plots the 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 filings since 2007. By contrast, human viewership has remained at relatively low levels, compared to machines, over time. Therefore, while machines now account for the vast majority of information access, they have increased their dominance compared to humans.

Cloud computing services, such as Amazon Web Services (AWS), 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, help avoid internet service disruption, and access servers from desired geographic

[26] The "Others" category has the lowest proportion of robot viewers, at roughly 46.02%.

locations. Therefore, cloud computing services are appealing for various financial institutions that would 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.[27]

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. Moreover, Abis and Veldkamp (2022) 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 increased about 13 times.

Figure (3) plots the geographical location of 8-K readers based on their IP address. More than 76% of 8-K visits are from the United States, 9% from the United Kingdom, 3% from India, 3% from China, and another 3% from Canada, with the remaining 6% from the rest of the world.

Finally, Figure (4) plots the percentage of the total and human visits to newly published 8-K filings per weekday, and the percentage of 8-K filings published daily in our sample, indicating that human readers, in contrast to machines, pay far less attention to 8-K filings on Fridays.

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:

$$V_{i,j,t}^{M,H} = \alpha_0 + \sum_{j=1}^{k} \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,j,t} \tag{4}$$

[27]The 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 measure of the number of 8-K views by machines, 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 word count in the 8-K ($\#Words$), the Gap that captures the number of days between the event date and the 8-K filing date (Gap), the number of items in the 8-K ($\#Items$), the number of news mentions within two weeks prior to the 8-K publication ($\#News$), the average percent effective spread over the week before the 8-K publication ($Spread$), a dummy variable equal to one for 8-K released on Fridays ($Friday$), the firm's book-to-market ratio (BM), the firm's market capitalization ($SIZE$), the firm's institutional ownership ($InstOwn$), the number of analysts following the firm ($Analysts$), and a dummy variable for days with one or more outages for major cloud service providers ($Cld.Outage$). All variables are defined in 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 (1.072). This tendency is strongest for humans (Columns 2 and 5) and is insignificant for machines (Columns 3 and 6). These results suggest that negative sentiment content tends to attract more attention from human viewers. Column (1) also shows that larger 8-K are more likely to have a significantly larger total number of visits, as proxied by the word count in the 8-K ($\#Words$) and the number of items included in the 8-K ($\#Items$). High-dimensional and voluminous 8-K are likely to reveal more material information, prompting both machine and human readers to examine them more closely (Columns 2-7). Column (1) also shows that timely 8-K are likely to receive more attention, as the number of days between the event date and the 8-K publication date (Gap) is negatively related to total 8-K viewership.

Table (2) also shows that larger firms ($SIZE$) have greater human viewership (Columns 2 and 5) but not significantly greater machine or cloud machine viewership (Columns 3-4 and 6-7). 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 (2 and 5) also reveal that value firms (BM) have greater human viewership but not significantly greater machine viewership (Columns 3-4 and 6-7). Similarly, humans pay more attention to the 8-K of firms with low institutional ownership ($InstOwn$), which are likely to lack information transparency, yet machines show no preference for 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, 2022). Lastly, we find that both machines and cloud machines are negatively impacted by cloud outages, as per the negative and significant coefficients for ($Cld.Outage$).

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.[28] 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).[29]

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, Columns (1 and 6) show that machines and cloud machines do not pay significant attention to $Item2.02$, which mostly refers to non-GAAP earning disclosures, as reflected in the negative and significant $Item2.02$ coefficients. By contrast, humans pay significant attention to $Item2.02$ (Column 11). 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 (12) show that machines, but not humans, pay significantly more attention to $Item8.01$ (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, or anticipated, 8-K. Human viewers, in contrast, are more likely to target specific information disclosure events rather than unspecified ones.

Humans tend to pay less attention to $Item5.02$ relative to other items in 8-K filings (Column 14). Therefore, human readers appear to be less attentive to unanticipated events related to the departure

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

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

or election of officers or directors. By contrast, machines do not significantly under-allocate attention to $Item5.02$. Lastly, 8-K viewership by both machines and humans is higher for voluntary disclosure $Item7.01$ (i.e., Regulation FD Disclosure), while 8-K viewership by humans is significantly higher for mandatory disclosure $Item1.01$ (i.e., Entry into a Material Definitive Agreement) as reported in Column (15).

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 information disclosed in non-standardized 8-K items.

5. Impact of Machines on Market Efficiency

In this section we investigate 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 alternative efficiency measures. We examine endogeneity issues in the subsequent section.

We hypothesize that machines contribute to the faster and more efficient reaction of stock prices to new information. Therefore, we expect that the information inefficiency in price 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 investors using cloud computing services are likely to be among the most sophisticated and therefore will have the most significant impact towards improving informational efficiency.

5.1. Machine and Human Viewership

We estimate the following regression:

$$PI(2,T)_{i,t} = \alpha_0 + \sum_{j=1}^{2} \beta_j V_{i,t}^{M,H} + \sum_{j=1}^{k} \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \tag{5}$$

where $PI(2,T)_{i,t}$ 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 measure of 8-K viewership by humans, all the machines, or machines from various facilities, such as *Cloud Computing Services*, *Financial Institutions*, and *Data Vendors*. 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 publication date, institutional ownership, analyst coverage, the average percent effective spread, a dummy variable equal to one for 8-Ks released on Friday, the number of news mentions, the $|CAR(0,1)|$, the negative sentiment content in the 8-K, and the number of items in the 8-K. All variables are defined in Appendix Table A. $\tilde{f}$ and $\tilde{\tau}$ are firm and year fixed effects, respectively.

Table (4) reports the results. Columns (1-2) provide the first evidence that the $PI(2,10)$ following the 8-K publication date is mainly associated with human viewership of the 8-K, rather than aggregated machine viewership. In particular, Column (1) shows that 8-K viewership by humans has a positive and highly significant effect on the $PI(2,10)$. Column (2) simultaneously accounts for viewership by aggregated machines and humans, indicating that the *aggregate* machine viewership (including all categories of machines) of an 8-K has no significant effect on the $PI(2,10)$ following the 8-K publication date.

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 (such as viewership from universities, regulators, and auditing firms). While the same can be said of humans, the focus on this paper is on better understanding the role machines play in markets; therefore, this section examines whether different types of machines have different effects.[30]

We expect that 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 machine-based 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 on oper-

[30]In Appendix Table E, we present by-facility analysis results for human viewership. There is no notable heterogeneity in the correlation between price inefficiency and human viewership associated with IP addresses from various institutions. To some extent, Columns (1) and (5) in Appendix Table E can also be viewed as a placebo test for machine viewership from cloud computing, as it investigates viewership from cloud machines services by humans rather than a machine algorithm.

ating expenses and minimize capital investment on hardware.[31] 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.[32] Additionally, stock exchanges such as Nasdaq, Euronext, and Deutsche Börse, as well as brokerage firms such as Robinhood and iSTOX, rely extensively on cloud storage. Investment banks such as Goldman Sachs and data vendors have also begun to offer services to institutional investors through cloud computing.[33] However, how institutional investors implement their investment strategies through cloud computing facilities remains largely 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. Column (3) of Table (4) reports the results from this analysis. The results show that machine viewership from cloud computing services does indeed have a significant positive association with informational efficiency (negative association with $PI(2,10)$). In terms of economic significance, a one standard deviation increase in the machine-based viewership of 8-Ks from cloud computing services is associated with a 15.73% aggregate decrease in the $PI(2,10)$ over the post-filing month.[34]

In Columns (4-5) of Table (4) we extend our analysis by increasing the $PI(2,T)$ window from $PI(2,10)$ to $PI(2,20)$, with the same control variables. We find consistent and even stronger evidence regarding the impact of viewership by machines and humans on the price inefficiency.[35] We also

[31] "Mirae Asset Global Investments Case Study," https://aws.amazon.com/solutions/case-studies/mirae-asset/, accessed August 23, 2021.

[32] For more details, please see AWS re:Invent Talks of Jeff Dowds at 2019 and Coline Mazzola at 2018.

[33] More 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.

[34] The unreported standard deviation of raw machine visits from cloud machine services in the first two days is 110.98, and its log is $ln(1+110.98) = 4.72$. Given the average number of post-event price inefficiency, $PI(2,20)$, is 0.09, a one standard deviation increase in machine visits from cloud services leads to a 15.73% ($= 0.003 \times 4.72/0.09$) decrease in post-event price inefficiency.

[35] We 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 Appendix Table C, are qualitatively similar. We also

examine an alternative measure of inefficiency in Section (*5.2*) and find similar results. The effects of human viewership on the $PI(2, 20)$ are also economically significant. For example, a one standard deviation increase in human viewership is associated with a 13.23% aggregate increase in the $PI(2, 20)$ over the post-publication month.[36] Therefore, the well-established inefficiencies from overreaction and underreaction to information in financial markets are driven by humans and not machines. It is worth noting that a sizable amount of human viewers may not only slow down information integration but also contribute to more heterogeneity in belief, resulting in more noise components in stock price. According to the models of the heterogeneous belief (e.g., Kandel and Pearson, 1995, among others), a larger body of human viewers would extend the variation in interpretation of the same information, and sequentially reduce the proportion of effective information in price.

Column (6) illustrates the statistical significance of machine viewership from cloud computing services coefficient, indicating that investors behind clouds likely to improve market efficiency in the short-term. The estimates in Columns 3 and 6 of Table (**4**) also show that machine viewership from financial institutions other than those using cloud computing facilities ($InstInvMachine$) and machine viewership from data & media publishers ($DataVendMachine$), respectively, do not have a significant effect on the price inefficiency.[37]

The regressions control for several firm characteristics. Across Columns (1-6) we find that firm size and ROA are negatively related to the price inefficiency following the 8-K filing date, suggesting that larger and more profitable firms have more efficient prices. Consistent with Ben-Rephael et al. (2021), Columns (1-6) show a negative and highly significant relation between institutional ownership and price inefficiency following the 8-K filing date, with the coefficient of $InstOwn$ ranging between -0.007 and -0.014.[38] Other noticeable results suggest that, across all columns, the coefficients of $Spread$ and $Friday$ are positive and significant at the 1% confidence level, while the coefficient of

repeat all tests with the machine and human viewership measured by the DRT (Drake et al., 2015) method, and the results are qualitatively similar.

[36] The 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 the average number of post-event price inefficiency, $PI(2, 20)$, is 0.09, a one standard deviation increase in human visits leads to a 13.23% ($0.003 \times 3.97/0.09$) increase in post-event price inefficiency.

[37] Once 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.

[38] In their analysis of abnormal institutional attention and price discovery during the filing period, Ben-Rephael et al. (2021) 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.

$\#News$ variable is negative and significant at the 1% confidence level.

5.2. Evidence from alternative measures of informational efficiency

To test the robustness of our main results, we re-estimate the regressions from the previous subsections using a different measure of information (in)efficiency – the *Noise* measure developed by Brogaard et al. (2022). This measure is the standard deviation of estimated pricing errors obtained from a variance decomposition model that separates information and noise. It captures both underreaction and overreaction to information as two forms of inefficiency that contribute to "noise" in prices. In a series of validation tests, Brogaard et al. (2022) show that the *Noise* measure captures informational inefficiency in prices.

Additionally, we also test a $ScaledPI$ measure, which is $|CAR(2,T)|$ normalized by $|CAR(0,T)|$ (as described in Equation (3)). The normalization removes the total magnitude of the information, making the PI a relative or proportional measure. However, a downside of the normalization, and the reason why we do not use it in the baseline tests, is that its denominator may be near zero, leading to extreme values and some undefined values.[39]

Similar to the previous subsections, we regress *Noise* and $ScaledPI$ on measures of viewership by each of the machine types and include the same control variables as before. The results, reported in Table (**5**), support the conclusions made based on the PI measures. Specifically, human viewership has a significant positive association with noise in prices (coefficient of 0.062, significant at the 5% level) (Column 9), while cloud computing machine viewership has a significant negative association with noise (coefficient of -0.089, significant at the 5% level) (Column 8). The results for $ScaledPI$ are also consistent with our baseline tests (coefficient of $CloudMachine$ is negative and significant in both Columns (2 and 4)), except that the human coefficient is only statistically significant in the analysis of $ScaledPI(2,20)$ in Column (6).

The results therefore support the notion that access to information by machines operating from cloud computing services tends to improve the informational efficiency of prices.

[39] To mitigate this issue, we exclude observations with small denominator values $|CAR(0,T)| < 0.1$ percent).

5.3. Directional events and market reactions

Our PI measure also allows us to separately investigate the impacts of human or cloud viewership on price inefficiency in positive and negative events and when the market overacts or underacts. We first define an event as negative (positive) if the market's initial reaction to its announcement is less than or equal to zero (greater than zero) in the first two days. We then consider a stock experiencing a price drift when its price develops in the same direction, or price reversal otherwise. This treatment partitions our sample into four groups, tilting towards the price drift and negative quadrants.

Panel A (B) of Table (6) presents the outcomes of negative (positive) events, while Columns (1-4) and Columns (5-8) of both panels provide price drift and reversal quadrant, respectively. The findings reveal that the association between human viewing and PI metrics is statistically significant at the 1% level and is more prominent in adverse events, especially in the price drift quadrant. Those results present consistent evidence that human viewership deteriorates market efficiency more in negative news through under-reaction.

On the other hand, the viewership from cloud computing is significantly and negatively correlated with $PI(0,10)$ across four quadrants, suggesting that investors behind clouds improve market efficiency over a short period. Interestingly, this effect is more pronounced in negative events, presumably through interactions with other investors related to human viewership. The results also exhibit that viewership from cloud computing helps the most in reducing price inefficiency due to the market's overreaction to negative news. In contrast, the informational role of cloud viewership disappears in two weeks, $PI(0,20)$, among the good news, suggesting less profitable quantitative strategies for cloud investors in positive events over this window.

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 that directly impact machines. Cloud computing outages unambiguously reduce the capacity for cloud-based machines to access information. To focus on the events that are likely to affect many investors using automated models, we consider days on which one or several of the major cloud service providers (e.g., Amazon Web Services) have a significant outage, *Cloud Outage*. Such outages, by interrupting the connectivity of computing facilities used to run automated investment models, are expected to reduce machines' ability to improve price discovery.

We estimate two-stage least-squares (2SLS) models in which machine, cloud machine, and human viewerships are instrumented with cloud outages. Machine and cloud machine viewership of 8-K are likely to be affected by cloud outages. While cloud outages are likely to affect market efficiency, their effects are through the machine and cloud machine participants' ability to access and assimilate relevant company-specific information. We estimate the first-stage equation for machines, cloud machines, and humans separately. We control for BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$, and we include firm and year fixed effects.

Table (7) reports the results. Columns (1-2) show that our instrument significantly affects machine and cloud machine viewership of 8-K filings, as expected. Both machines and cloud machines are less likely to access information related to unscheduled events during cloud outage days. Moreover, we find that human viewership is not related to cloud outage (Column 3). To formally test the relevance condition of an instrumental variable, we report Kleibergen-Paap F-statistics (for non-i.i.d. errors). The F-statistic in Column (2) for the first stage is about 13.16 for cloud machine viewership, which is much greater than Stock and Yogo (2002) critical values of 5% maximal IV relative bias and 10% maximal IV size tests. Those results imply that cloud machine viewership identification does not suffer from an insufficient instrumental variable. The F-statistics in Columns (1) and (3) suggest our instrumental variable does not affect informational efficiency through general machine and human viewership, respectively.

In the second stage we regress $PI(2,T)_{i,t}$, which is our baseline measure of *inefficiency*, defined earlier in Equation (2), on the fitted values of cloud machine viewership (Column 5) and a set of control variables and firm and year fixed effects. The evidence is consistent with our baseline results discussed earlier. Namely, cloud machine viewership is significantly negatively related to PI. Hence, the instrumental variable model supports a causal interpretation of the baseline result that cloud computing machines that access company information improve the efficiency of price reactions.

In Appendix Table D we present results from an alternative approach to using the exogenous cloud outages. In this framework, we regress $PI(2,T)_{i,t}$ on machine and human viewership and a dummy variable that is one on *Cloud Outage* days. We include interactions between cloud computing, machine viewership, human viewership, and the *Cloud Outage* dummy variable, and control for our standard list of stock characteristics. As expected, we find that the coefficients of the interaction between $Machine$ or $CloudMachine$ viewership and *Cloud Outage*, $\hat{\beta}_2$ and $\hat{\beta}_4$, are positive and significant in Columns (1) and (2) (and also in Columns 4 and 5). These estimates suggest that during unexpected cloud outage days the viewership through some cloud machines is less effective in reducing the PI following the 8-K publication, and there is a corresponding deterioration in market efficiency. By contrast, we find no significant effects for the interaction between human viewership and major cloud outages (insignificant $\hat{\beta}_6$ coefficients in Columns 3 and 6).

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, acquisitions, bankruptcy, and spin-offs. The sudden surge in institutional investors' attention to 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).[40] Human attention is a scarce cognitive resource (Barber and Odean, 2008;

[40] Begenau 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.

Kahneman, 1973) and influences the market's reaction to corporate announcements (e.g., Hirshleifer 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 increase the impact of human viewership on informational efficiency.

By 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 PI following 8-K filings. The dependent variable is the $PI(2,T)_{i,t}$ 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 inefficiency.

Table (8) reports the results of the difference-in-difference tests. The primary coefficients of interest are the interaction term between human viewership and index inclusion, $Human \times InIndex$. These coefficients are negative in all specifications (Columns 3 and 6) and statistically significant at the 5% and 10% levels. Consistent with our hypothesis, the results suggest that index inclusion improves the impact of human viewership on informational efficiency. That is, all else equal, human viewership of S&P500 stocks is more likely to improve efficiency than human viewership of non-S&P500 stocks.

By contrast, the coefficients of $Machine \times InIndex$ (Columns 1 and 4), as well as the coefficients of $CloundMachine \times InIndex$ (Columns 2 and 5), are insignificant, suggesting that S&P500 inclusion does not significantly affect the impact of machines or cloud machines on PI. 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. Additional instrumental variables approach

Finally, we consider one further set of instrumental variables tests. This time, instead of drawing on exogenous variation in cloud computing, we consider instruments that affect human access to information. The details of these additional tests are in Appendix F, with the results reported in Appendix Table F; here we provide a summary.

We instrument human viewership with 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)) and macroeconomic news announcements (Peng and Xiong, 2006; Kacperczyk et al., 2016) are likely to disproportionately affect humans' decisions. 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 (Barbopoulos et al., 2020). 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.

We find that these instruments significantly affect human viewership of 8-K filings, as hypothesized. Human viewership is negatively related to investor sentiment and macroeconomic news released at the time of the 8-K publication date. 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 results suggest our instrument variables do not affect information efficiency through non-human viewership.

The second-stage results show that human viewership is significantly positively related to PI. Hence, the instrumental variables models support the baseline regression 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 PI following 8-K filings. In unreported tables, 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 machines results in more efficient reflection of that information in prices? We conjecture that the mechanism is that machines access the information, and 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 an increase in informed trading when cloud-based machines access information, reflecting more accurate interpretation of the information. Conversely, given the results about efficiency, when humans access information we do not expect to see a relative increase in measures of informed trading. In this section we test both of these aspects of the conjectured mechanism of how machines impact informational efficiency. We then examine whether the effects of machines accessing company information are different from the effects of HFTs in markets. Lastly, we test whether machines and humans impact market efficiency differently when they process good vs bad news.

7.1. Does machine access to information increase informed trading?

First, we test the relation between human and machine viewership and measures of informed trading. 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 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}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$.

Table (9) 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 4). The results also show that machine viewership specifically from cloud computing services is the predominant driver of the positive relation between machine viewership and PIN (Columns 2 and 5). By contrast, we find a significantly negative relation between human viewership and $PIN(0,1)$ and $PIN(0,5)$ (Columns 3 and 6).

In summary, the results indicate that 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. They support the conjectured channel that machine access to information improves the efficiency of prices as a result of machines using the information to make trading decisions.

7.2. The role of high-frequency trading

Earlier studies show that HFTs, which are also automated machine-based traders, improve price discovery at intraday horizons (Carrion, 2013; Brogaard et al., 2014). We therefore examine whether the effects of machine access to information are driven by, or are distinct from, the effects of HFTs. To do this, we augment our tests of how machine viewership impacts efficiency by including two common measures of HFT activity on the right-hand side of the regressions: the *Cancel-to-Trade* ratio and the *Trade-to-Order* ratio (an inverse measure of HFT activity). Both measures are based on the fact that HFTs tend to submit a much larger number of orders and order cancellations per trade because they have a speed advantage in dynamically managing orders in the limit order book. Using these two measures obtained from the SEC MIDAS dataset, we rerun our baseline regressions for the period of available data (2012 to 2016).

The results in Table (10) show that, consistent with existing literature on HFTs, one of the proxies of HFT activity is associated with more efficient price reactions. The *Trade-to-Order* ratio is positively related to the price inefficiency, suggesting that when HFTs are more active (lower values of *Trade-to-Order*), PI is smaller, consistent with a more efficient price reaction.

Importantly, however, the positive impact of cloud machine access of 8-K filings on informational efficiency is not diminished by controlling for HFT activity. In Columns (3-4) and (9-10) we find that irrespective of which HFT proxy we control for, the effects of cloud machines accessing information on informational efficiency remains statistically significant and of a similar magnitude to our baseline tests. These results suggest that the effects in this paper are different to what has been previously found for HFTs. Taken together, the literature on HFTs and this paper suggest machines can benefit informational efficiency in multiple ways and at different horizons. HFT studies suggest HFTs improve intraday price discovery, largely by inferring the information in order flow (e.g., large orders of institutional investors sliced into a series of trades), i.e., the information gathering activities of other investors. By contrast, our results suggest machines also play a role in accessing new corporate information and more efficiently incorporating it into prices at daily horizons.

Although the effects of machines accessing company information are distinct from the role of HFTs, when machines access information and trade on it, there is a subsequent increase in algorithmic trading. Table G in the Appendix shows that when machines access company information, there is an increase in the ratio of odd lots and a decrease in the average trade size, consistent with machines engaging in algorithmic trading based on the (newly accessed) information. Algorithmic trading is a broader category of automated trading (HFT is a subset) that, in contrast to HFTs, includes order execution algorithms that optimally slice a larger order into smaller trades and automatically executes then on the market.

8. Comparative advantage analysis

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

First, we consider hard, numerical information, for which we conjecture machines will have a comparative advantage as it feeds naturally into quantitative models. Conversely, soft, qualitative information is where we expect humans may still have an advantage. Second, we consider the role of negative sentiment in 8-K filings and how that can impact decision-making – machines should be less susceptible to emotions and therefore better able to handle such information, which is known to elicit bias in human decisions (Chakrabarty et al., 2022). Finally, we explore how well machines and humans process repeated information, considering possible benefits for attention-constrained humans and potential errors from machines misinterpreting repeated information as new signals.

8.1. Hard, numeric information

To test whether machines have a comparative advantage in handling hard, numeric or quantitative information, we construct a measure of the proportion of numerical content in each 8-K filing. We do this by using the Named Entity Recognition (NER) method, which identifies the proportion of numeric values, including dollar amounts, percentages, and cardinal numbers (mostly tabulated items) in both Arabic and English numerals, while discarding irrelevant numeric values like phone numbers, zip codes, dates, and page numbers.[41] We label this machine-readable numerical proportion as *NER*. We regress $PI(2,T)_{i,t}$ on machine and human viewership of 8-K filings, our measure of *NER*, and interactions between machine and human viewership and *NER*. We include the standard control variables and firm and year fixed effects.

Table (**11**) reports the results. Consistent with our conjecture, Columns (2) and (5) show that cloud machine 8-K viewership reduces the PI following 8-K filings with significantly more numerical content. These results align with Hope et al. (2016), who document that the market reaction to regulatory disclosures significantly correlates with the number of specific terms that the machine can recognize in the risk disclosure section of 10-K filings.

By contrast, the interactions between human viewership of 8-Ks and their percentage of numerical content are not statistically different from zero (Columns 3 and 6).

[41] We use a publicly accessible and free Python module, spaCy, which has a 90% NER accuracy (F1-score), 5% more than the Stanford NER method. For more information on NER applications in textual analysis, see Dai et al. (2022).

8.2. Negative sentiment and the role of emotions

To examine whether machines are better able to handle negative sentiment content in 8-Ks, we regress $PI(2,T)_{i,t}$ on our machine and human viewership measures, a measure of negative sentiment in the 8-K based on Loughran and Mcdonald (2011), as well as interactions between the viewership measures and negative sentiment. We include the standard control variables and firm and year fixed effects.

Table (**12**) 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 PI) when information is more likely to induce bias in humans due to negative content. Specifically, Columns (2) and (5) show that machine viewership from cloud computing services reduces the PI more among 8-Ks with considerable negative sentiment content. The interaction term coefficients are negative and significant following the 8-K publication date. Along similar lines, results in Columns (1) and (4) show that machine viewership (overall machine) also reduces the PI following 8-Ks with higher negative content.

Lastly, when we use human viewership (in Columns 3 and 6) the interaction variables are no longer significant, suggesting human viewership of 8-Ks with significant amounts of negative sentiment does not improve price discovery.

8.3. 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 repeated 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.

By contrast, humans may benefit from repeated information, given that 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 initially missed the information, identify and react to it, helping impound the information in prices.

To examine these conjectures, we use $Item2.02$ of the 8-Ks, which often reveals the similar earnings information that has already been disclosed in conference calls and other press release venues. We regress $PI(2,T)_{i,t}$ on machine and human viewership measures, a dummy variable for 8-K filings that includes $Item2.02$, and interactions between machine and human viewership and the $Item2.02$ dummy. We include the standard controls and firm and year fixed effects.

The results reported in Table (**13**) indicate that for incremental information (released via $Item2.02$), neither machines in aggregate nor cloud-based machines make any significant contribution to reducing PI, showing that their impact on market inefficiency is indifferent to this type of 8-K. By contrast, the interaction term between human viewership and $Item2.02$ has a negative and significant effect on the PI 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. Subsequent repetition of the information does not lead to a more efficient response from machines. Repeated information also does not appear to cause significant errors among machines, which face a risk of interpreting repeated information as a new signal. By 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 or reversals following information releases, and decreasing noise variance in returns. By contrast, humans accessing the same information do not benefit informational efficiency as much and can even harm efficiency. Moreover, other types of machines are generally not as beneficial to market efficiency as cloud computing servers. We address identification and find evidence of causality in these effects by examining exogenous outages of cloud computing services, 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 machines vs humans have an advantage, we find that the advantage of machines increases for hard, numerical, and quantitative data. Conversely, humans have a comparative advantage in interpreting soft, qualitative information. We also find that cloud-based machines are more effective in processing information that has negative sentiment compared to humans, consistent with the notion that machines are less susceptible to bias from emotions. On the other hand, humans 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. By 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 predicting that humans specialize in slower, deeper learning and decision-making that combines sequential information, whereas machines have an edge in making quick decisions based on isolated information.

Overall, our findings illustrate 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., 2022. Man vs. machine: Quantitative and discretionary equity management. Columbia University Working Paper pp. 1–79.

Abis, S., Veldkamp, L., 2022. 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. 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. Journal of Finance 61, 1645–1680.

Baldauf, M., Mollner, J., 2020. High-frequency trading and market performance. The Journal of Finance 75, 1495–1526.

Baldauf, M., Mollner, J., 2021. Fast traders make a quick buck: The role of speed in liquidity provision. Journal of Financial Markets 58, 100621.

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. Review of Financial Studies 21, 785–818.

Barbopoulos, L. G., Adra, S., Saunders, A., 2020. Macroeconomic news and acquirer returns in M&As: The impact of investor alertness. Journal of Corporate Finance 64.

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. D., Israelsen, R. D., 2021. Who Pays Attention to SEC Form 8-K? The Accounting Review 97, 59–88.

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? Review of Financial Studies 22, 3211–3243.

Blankespoor, E., Miller, B. P., White, H. D., 2014. Initial evidence on the market impact of the xbrl mandate. Review of Accounting Studies 19, 1468–1503.

Bradshaw, M. T., Lock, B., Wang, X., Zhou, D., 2020. Soft information in the financial press and analyst revisions. Accounting Review .

Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. Review of Financial Studies 27, 2267–2306.

Brogaard, J., Nguyen, T. H., Putniņš, T. J., Wu, E., 2022. What moves stock prices? the role of news, noise, and information. Review of Financial Studies .

Brunnermeier, M. K., 2005. Information leakage and market efficiency. The Review of Financial Studies 18, 417–457.

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. 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.

Carrion, A., 2013. Very fast money: High-frequency trading on the nasdaq. Journal of Financial Markets 16, 680–711, high-Frequency Trading.

Carter, M. E., Soo, B. S., 1999. The relevance of form 8-k reports. Journal of Accounting Research 37, 119–132.

Chakrabarty, B., Moulton, P. C., (Frank), Wang, X., 2022. Attention: How high-frequency trading improves price efficiency following earnings announcements. Journal of Financial Markets 57, S138641812100063X.

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. 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.

Dai, R., Donohue, L., Drechsler, Q. F., Jiang, W., 2022. Dissemination, publication, and impact of finance research: When novelty meets conventionality. Review of Finance Forthcoming.

Dechow, P. M., Lawrence, A., Ryans, J. P., 2015. SEC comment letters and insider sales. Accounting Review 91, 401–439.

Dellavigna, S., Pollet, J. M., 2009. Investor inattention and friday earnings announcements. Journal of Finance 64, 709–749.

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.

Gao, M., Huang, J., 2020. Informing the market: The effect of modern information technologies on information production. The Review of Financial Studies 33, 1367–1411.

García, D., 2013. Sentiment during recessions. 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. 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? 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. Journal of Finance 64, 2289–2325.

Hope, O.-K., Hu, D., Lu, H., 2016. The benefits of specific risk-factor disclosures. Review of Accounting Studies 21, 1005–1045.

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.

Kandel, E., Pearson, N., 1995. Differential interpretation of public signals and trade in speculative markets. Journal of Political Economy 103, 831–72.

Kempf, E., Manconi, A., Spalt, O., 2017. Distracted shareholders and corporate actions. 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? 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. Review of Economic Studies 56, 317–355.

Lee, C., Watts, E. M., 2020. Tick size tolls: Can a trading slowdown improve price discovery? Accounting Review Forthcoming.

Lerman, A., Livnat, J., 2010. The new form 8-k disclosures. Review of Accounting Studies 15, 752–778.

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. Journal of Finance 66, 35–65.

Meulbroek, L. K., 1992. An empirical analysis of illegal insider trading. Journal of Finance 47, 1661–1699.

Peng, L., 2005. Learning with information capacity constraints. 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? 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. 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. Journal of Finance 62, 1139–1168.

Van Kervel, V., Menkveld, A. J., 2019. High-frequency trading around large institutional orders. Journal of Finance 74, 1091–1137.

Veldkamp, L. L., 2006. Information markets and the comovement of asset prices. Review of Economic Studies 73, 823–845.

von Beschwitz, B., Keim, D. B., Massa, M., 2020. First to "Read" the News: News Analytics and Algorithmic Trading. The Review of Asset Pricing Studies 10, 122–178.

Watkins, J., 2022. Consequences of Prescribed Disclosure Timeliness: Evidence from Acceleration of the Form 8-K Filing Deadline. The Accounting Review .

Weller, B. M., 2017. Does algorithmic trading reduce information acquisition? Review of Financial Studies 31, 2184–2226.

Wurgler, J., Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks? 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 $t \in \{0, 1\}$.

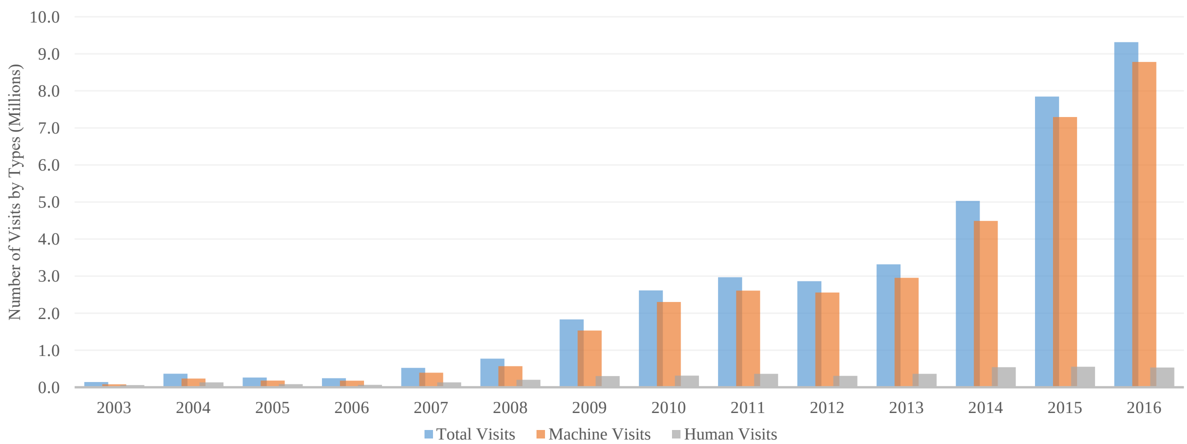

Figure 2: **8-K viewership activity by organization**

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).

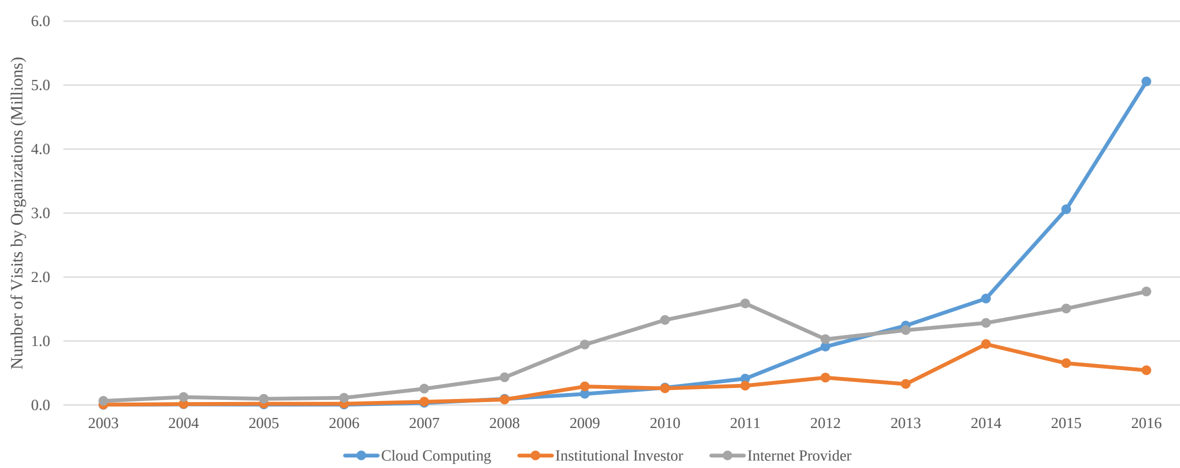

Figure 3: **Geographic locations of 8-K visits**

This figure presents the geographic locations of 8-K visits on days $t \in \{0, 1\}$ (relative to the 8-K publication date) based on the IP address of the 8-K visit.

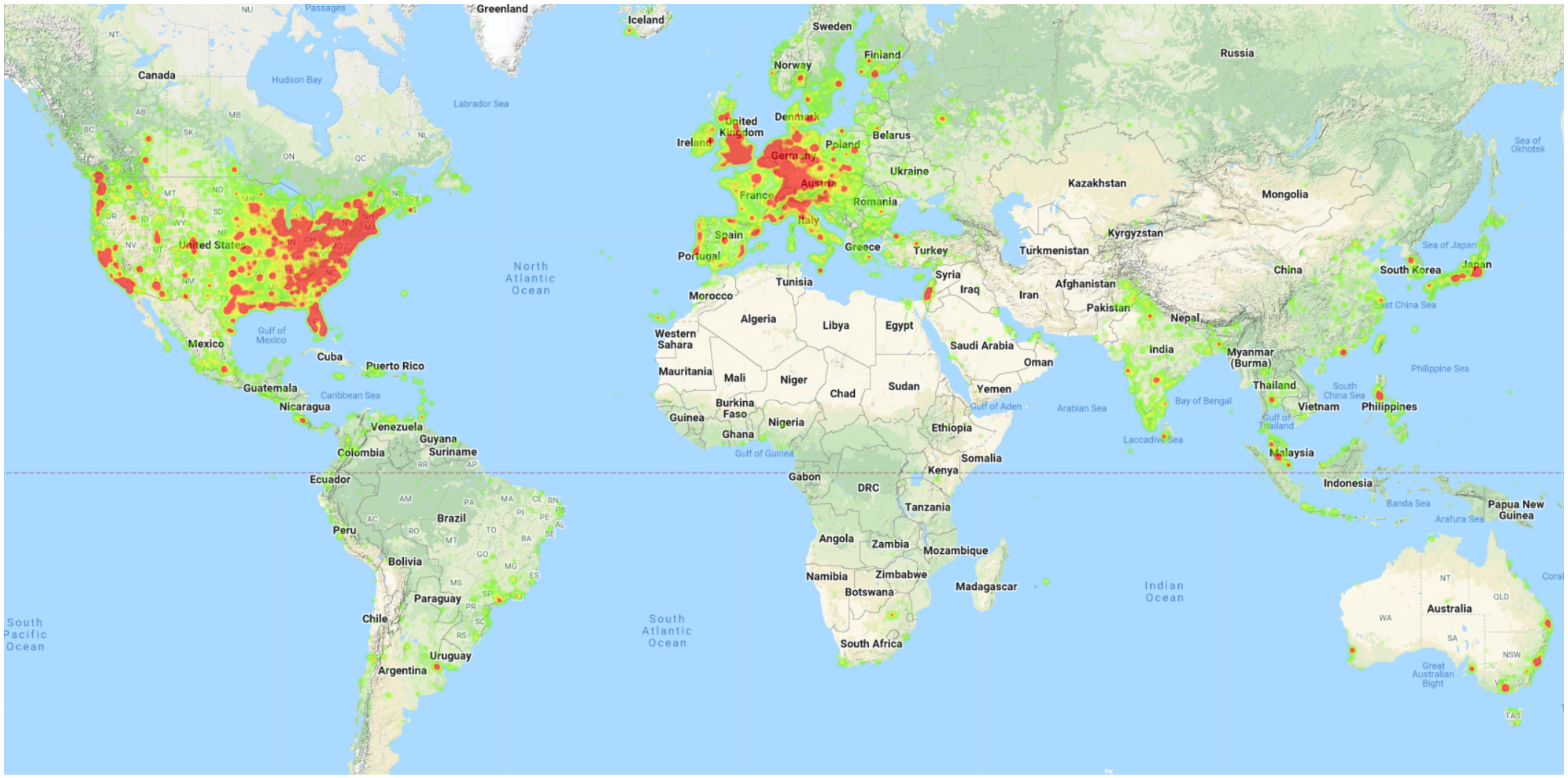

Figure 4: **Publications and visits of 8-Ks by weekday**

This figure presents the percentage of the total and human visits to the latest 8-K filings (published within one business day) by weekday, as well as the percentage of newly published 8-K filings each weekday in our sample.

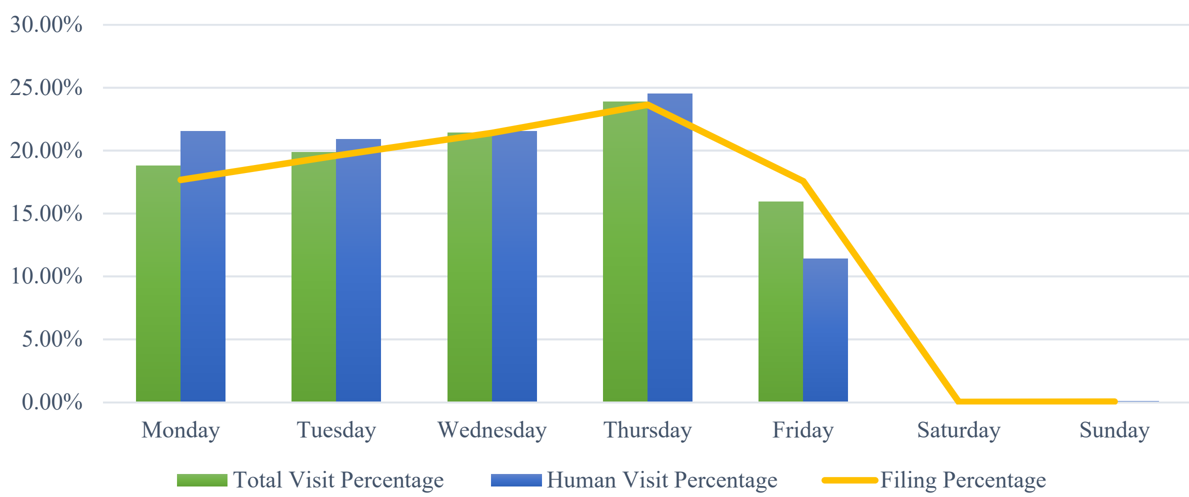

Table 1: **Summary statistics**

This table reports summary statistics (*N*, *Mean*, *Standard Deviation*, *25th Percentile*, *Median*, and *75th Percentile*) for the dependent variables (Panel A), independent (8-K Viewership) variables (Panel B), control variables (Panel C), 8-K items (Panel D), and visiting organizations (Panel E). All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles.

	N	Mean	Std Dev	25th Pctl	Median	75th Pctl
	Panel A: Dependent Variables					
$PI(2,10)$	287,249	1.28	1.93	0.42	0.79	1.24
$PI(2,20)$	287,249	1.29	1.81	0.55	0.88	1.23
Scaled $PI(2,10)$	286,852	1.42	3.27	0.55	0.88	1.24
Scaled $PI(2,20)$	286,852	1.40	3.36	0.68	0.94	1.20
$Noise$	259,846	2.42	1.71	1.24	1.90	3.02
$PIN(0,1)$	271,454	0.31	0.15	0.20	0.27	0.39
$PIN(0,5)$	277,889	0.32	0.14	0.21	0.28	0.39
	Panel B: Main Independent (8-K Viewership) Variables					
$TotalVisits$	295,540	3.63	1.29	2.49	3.89	4.62
$Machine$	295,540	3.47	1.36	2.20	3.81	4.51
$Machine_{DRT}$	295,540	3.52	1.35	2.30	3.85	4.55
$CloudMachine$	295,540	2.08	1.63	0.00	2.08	3.43
$InstInvMachine$	295,540	1.51	1.13	0.69	1.61	2.40
$DataVendMachine$	295,540	1.17	0.90	0.00	1.39	1.79
$Human$	295,540	1.46	1.08	0.69	1.39	2.20
$Human_{DRT}$	295,540	1.14	0.98	0.00	1.10	1.79
	Panel C: Control Variables					
BM	295,540	0.75	0.97	0.29	0.54	0.90
$SIZE$	295,540	6.63	2.05	5.18	6.59	7.99
ROA	295,540	-0.03	0.19	-0.01	0.02	0.06
LEV	295,540	0.23	0.22	0.04	0.19	0.37
STD_{RET}	295,540	0.12	0.09	0.06	0.10	0.15
$InstOwn$	295,540	0.63	0.31	0.38	0.69	0.88
$Analysts$	295,540	1.72	0.97	1.10	1.79	2.49
$Spread$	295,540	0.01	0.02	0.00	0.00	0.01
$Friday$	295,540	0.13	0.34	0.00	0.00	0.00
$\#News$	295,540	2.11	2.04	0.00	2.08	4.03
$\|CAR(0,1)\|$	295,540	0.05	0.07	0.01	0.02	0.06
$FinNeg$	295,540	0.01	0.01	0.00	0.00	0.01
$\#Items$	295,496	1.99	0.63	2.00	2.00	2.00
$\#Words$	295,540	6.22	0.99	5.61	5.87	6.52
NER	292,203	0.04	0.03	0.02	0.03	0.06
Cancel-to-Trade	105,499	32.13	42.81	13.47	19.64	31.39
Trade-to-Order	106,030	3.93	2.75	2.00	3.30	5.15

Table 1 (continued)

	N	Mean	Std Dev	25th Pctl	Median	75th Pctl
	Panel D: 8K Items					
	#Filings	#Visits Per Filing				
*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
	Panel E: Visiting Organizations					
	#Org.	#Visits Per Organization (Org.)				
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 of the determinants of machine or human viewership of firm i's 8-K filing. $TotalVisits$ is the natural log of one plus the number of visits of firm i's 8-K by both machines and humans. $Human$, $Machine$, and $CloudMachine$ are measures of human, machine, and cloud machine viewership of firm i's 8-K on days $t \in \{0, 1\}$, relative to 8-K publication date. In columns "DLR" these are measured using the approach of Dechow et al. (2015) and in columns "DRT" these are measured using the approach of Drake et al. (2015). Explanatory variables include $FinNeg$, $\#Words$, Gap, $\#Items$, $\#News$, $Spread$, $Friday$, BM, $SIZE$, $InstOwn$, $Analysts$, and $Cld.Outage$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	DLR				DRT		
	Total Visits	*Human*	*Machine*	*Cloud Machine*	*Human*	*Machine*	*Cloud Machine*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FinNeg	1.072*	5.404***	0.040	-1.445**	6.082***	0.229	-1.415**
	(1.99)	(12.75)	(0.08)	(-2.73)	(11.84)	(0.43)	(-2.65)
#Words	0.048***	0.048***	0.048**	0.054**	0.039***	0.051***	0.054***
	(3.06)	(6.29)	(2.77)	(2.94)	(5.31)	(3.08)	(3.05)
Gap	-0.014***	-0.042***	-0.009***	-0.011**	-0.034***	-0.011***	-0.011**
	(-5.35)	(-7.38)	(-3.49)	(-2.55)	(-7.34)	(-4.09)	(-2.54)
#Items	0.108***	0.160***	0.089***	0.049***	0.111***	0.101***	0.053***
	(11.89)	(19.04)	(8.48)	(4.11)	(13.10)	(9.82)	(4.15)
#News	0.006	0.053***	-0.000	-0.009	0.028***	0.004	-0.009
	(1.16)	(11.03)	(-0.05)	(-1.69)	(7.66)	(0.81)	(-1.58)
Spread	0.408	-0.851	0.764	-1.134**	-0.612	0.626	-1.175**
	(0.32)	(-1.12)	(0.56)	(-2.49)	(-1.15)	(0.46)	(-2.44)
Friday	-0.082***	-0.163***	-0.062***	-0.014	-0.120***	-0.072***	-0.017*
	(-7.35)	(-16.69)	(-5.39)	(-1.52)	(-15.00)	(-6.10)	(-1.78)
BM	0.016**	0.057***	0.002	-0.007	0.058***	0.006	-0.006
	(2.90)	(6.93)	(0.31)	(-1.06)	(7.55)	(1.18)	(-0.95)
SIZE	0.015	0.031**	0.009	-0.009	0.036***	0.010	-0.008
	(1.70)	(2.96)	(1.07)	(-1.51)	(3.76)	(1.14)	(-1.37)
InstOwn	-0.008	-0.149***	-0.015	-0.035	-0.178***	-0.012	-0.034
	(-0.26)	(-3.59)	(-0.51)	(-0.99)	(-4.01)	(-0.40)	(-0.93)
Analysts	0.004	0.007	0.002	0.012	0.008	0.002	0.011
	(0.42)	(0.62)	(0.23)	(1.65)	(0.73)	(0.21)	(1.61)
Cld.Outage	-0.045**	0.009	-0.048**	-0.070***	-0.007	-0.048**	-0.071***
	(-2.42)	(0.68)	(-2.23)	(-4.38)	(-0.48)	(-2.43)	(-4.71)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	514,731	514,731	514,731	514,731	514,731	514,731	514,731
Adj. R^2	0.854	0.400	0.879	0.912	0.345	0.867	0.907

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

This table reports OLS regression estimates of human and machine viewership of specific types of 8-K filings. *Machine*, *Cloud Machine*, and *Human* are measures of viewership of firm i's 8-K(s) on days $t \in \{0, 1\}$ (relative to the 8-K publication date), based on the Dechow et al. (2015) method. $Item\#.0\#$ are dummy variables equal to one if an 8-K filing contains an 8-K item#.0#, zero otherwise. $Item1.01$ refers to *Entry into a Material Definitive Agreement*, $Item2.02$ refers to *Results of Operations and Financial Condition*, $Item5.02$ refers to *Departure of Directors or Certain Officers Election of Directors Appointment*, $Item7.01$ refers to *Regulation FD Disclosure*, and $Item8.01$ refers to *Other Events*. Control variables in all regressions include $FinNeg$, $\#Words$, Gap, $\#Items$, $\#News$, $Spread$, $Friday$, BM, $SIZE$, $InstOwn$, $Analysts$, and $Cld.Outage$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Machine					Cloud Machine					Human				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
$Item2.02$	-0.043**					-0.036**					0.179***				
	(-2.29)					(-2.66)					(7.69)				
$Item8.01$		0.056*					0.003					-0.026***			
		(2.11)					(0.20)					(-3.87)			
$Item7.01$			0.067***					0.102***					0.060***		
			(4.91)					(4.66)					(5.60)		
$Item5.02$				-0.023					-0.003					-0.151***	
				(-1.09)					(-0.12)					(-8.15)	
$Item1.01$					0.023					0.001					0.048**
					(1.44)					(0.08)					(2.37)
Controls	Yes	Yes	Yes	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	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731	514,731
Adj. R^2	0.879	0.879	0.879	0.879	0.879	0.912	0.912	0.913	0.912	0.912	0.405	0.401	0.401	0.403	0.401

Table 4: **Impact of machine and human viewership on price inefficiency**

This table reports OLS regression estimates of the impact of all $Machine$, $Human$, $CloudMachine$, $InstInvMachine$, and $DataVendMachine$ viewership activities 8-K viewership on post-announcement PI. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement PI over different horizons, defined in Equation (2). $Machine$ and $Human$ denote machine and human viewership of firm i's 8-K filings on days $t \in \{0,1\}$ (relative to the 8-K publication date), measured as in Dechow et al. (2015). Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$		0.002			0.003	
		(0.68)			(0.79)	
$CloudMachine$			-0.003***			-0.003**
			(-3.14)			(-2.67)
$InstInvMachine$			0.001			0.001
			(1.12)			(1.02)
$DataVendMachine$			-0.002			-0.002
			(-0.72)			(-0.73)
$Human$	0.002***	0.002***	0.002***	0.003***	0.003***	0.004***
	(4.27)	(5.87)	(4.26)	(4.26)	(5.55)	(4.15)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.80)	(6.60)	(6.78)	(5.60)	(5.59)	(5.56)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.008***
	(-5.24)	(-5.25)	(-5.38)	(-8.68)	(-8.73)	(-8.87)
ROA	-0.027***	-0.027***	-0.026***	-0.037***	-0.037***	-0.037***
	(-6.42)	(-6.43)	(-6.46)	(-6.98)	(-7.02)	(-6.99)
LEV	0.013***	0.013***	0.013***	0.022***	0.023***	0.022***
	(3.21)	(3.31)	(3.18)	(5.02)	(5.14)	(4.95)
STD_{RET}	0.027***	0.027***	0.027***	0.045***	0.045***	0.045***
	(3.08)	(3.09)	(3.08)	(3.23)	(3.25)	(3.23)
$InstOwn$	-0.007*	-0.007*	-0.007**	-0.013***	-0.014***	-0.014***
	(-2.12)	(-2.13)	(-2.19)	(-3.60)	(-3.62)	(-3.62)
$Analysts$	0.001	0.001	0.001	-0.000	-0.000	-0.000
	(0.64)	(0.63)	(0.71)	(-0.09)	(-0.11)	(-0.04)
$Spread$	0.625***	0.623***	0.620***	0.682***	0.678***	0.675***
	(4.56)	(4.61)	(4.65)	(5.24)	(5.24)	(5.47)
$Friday$	0.002***	0.002***	0.002***	0.002***	0.003***	0.003***
	(4.44)	(4.30)	(4.34)	(5.46)	(5.24)	(5.59)
$\#News$	-0.001***	-0.001***	-0.001***	-0.003***	-0.002***	-0.003***
	(-3.82)	(-4.10)	(-3.75)	(-5.14)	(-5.55)	(-5.03)
$\|CAR(0,1)\|$	0.102***	0.102***	0.101***	0.104***	0.104***	0.102***
	(9.30)	(9.29)	(9.42)	(8.88)	(8.87)	(8.97)
$FinNeg$	0.027	0.025	0.024	0.071	0.068	0.066
	(0.55)	(0.53)	(0.49)	(0.90)	(0.89)	(0.84)
$\#Items$	-0.000***	-0.001**	-0.000*	-0.001***	-0.002***	-0.001***
	(-3.21)	(-2.66)	(-1.88)	(-3.85)	(-4.51)	(-4.35)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.222	0.222	0.222	0.243	0.243	0.244

Table 5: **Evidence from alternative measures of price inefficiency**

This table reports OLS regression estimates of the impact of *Machine*, *CloudMachine* and *Human* 8-K viewership on alternative measures of post-announcement *PI* and informational efficiency. *Machine*, *CloudMachine*, and *Human* denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0, 1\}$. It reports the results with Scaled $PI(2, 10)$ and Scaled $PI(2, 20)$, which are measures of the post-announcement *PI* over different horizons, and are scaled by the cumulative abnormal return during and after the 8-K filing. *Noise* is the share of stock return variance that is attributable to noise using the method in Brogaard et al. (2022). Control variables in all regressions include *BM*, *SIZE*, *ROA*, *LEV*, STD_{RET}, *InstOwn*, *Analysts*, *Spread*, *Friday*, *#News*, $|CAR(0, 1)|$, *FinNeg*, and *#Items*. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Scaled $PI(2, 10)$		Scaled $PI(2, 20)$		*Noise*	
	(1)	(2)	(3)	(4)	(5)	(6)
Machine	-0.028		-0.011		0.090	
	(-1.47)		(-0.80)		(0.69)	
CloudMachine		-0.038**		-0.036***		-0.111**
		(-2.58)		(-4.00)		(-2.88)
Human	0.006	0.005	0.030***	0.032***	0.045***	0.074**
	(0.78)	(0.76)	(3.22)	(3.73)	(3.55)	(2.82)
BM	-0.009	-0.009	0.005	0.004	0.170***	0.167***
	(-0.40)	(-0.41)	(0.25)	(0.23)	(6.02)	(5.90)
SIZE	-0.027	-0.028	-0.004	-0.004	-0.141***	-0.142***
	(-1.42)	(-1.48)	(-0.15)	(-0.18)	(-4.11)	(-4.12)
ROA	-0.091	-0.091	0.111	0.112	-0.618***	-0.614***
	(-1.10)	(-1.11)	(1.24)	(1.24)	(-7.36)	(-7.47)
LEV	-0.041	-0.040	-0.097	-0.097	0.503***	0.496***
	(-0.50)	(-0.49)	(-1.07)	(-1.07)	(4.36)	(4.18)
STD_{RET}	0.431***	0.433***	-0.056	-0.054	0.708**	0.710**
	(3.03)	(3.06)	(-0.52)	(-0.51)	(2.39)	(2.39)
InstOwn	0.014	0.013	-0.029	-0.030	-0.067	-0.067
	(0.17)	(0.16)	(-0.53)	(-0.55)	(-0.95)	(-0.97)
Analysts	0.002	0.002	0.054**	0.055**	0.018	0.020
	(0.06)	(0.08)	(2.87)	(2.88)	(0.97)	(1.10)
Spread	4.576***	4.500***	1.686*	1.631*	25.465***	25.432***
	(5.27)	(5.15)	(1.88)	(1.82)	(4.72)	(4.70)
Friday	-0.007	-0.006	-0.001	-0.000	0.042***	0.040***
	(-0.41)	(-0.36)	(-0.06)	(-0.02)	(4.99)	(5.39)
#News	0.006	0.006	0.029***	0.028***	-0.047***	-0.050***
	(0.75)	(0.72)	(5.29)	(5.08)	(-7.54)	(-6.72)
$\|CAR(0, 1)\|$	0.061	0.055	1.310***	1.303***	4.793***	4.763***
	(0.21)	(0.19)	(5.00)	(4.98)	(14.77)	(14.87)
FinNeg	0.989	0.986	-1.436**	-1.424**	1.637	1.827
	(0.74)	(0.74)	(-2.57)	(-2.43)	(1.10)	(1.20)
#Items	-0.007	-0.007	0.019	0.020	-0.024**	-0.014**
	(-0.49)	(-0.51)	(1.25)	(1.29)	(-2.26)	(-2.81)
Controls	Yes	Yes	Yes	Yes	Yes	Ye
Firm FE	Yes	Yes	Yes	Yes	Yes	Ye
Year FE	Yes	Yes	Yes	Yes	Yes	Ye
NObs	286,852	286,852	286,852	286,852	259,478	259,478
Adj. R^2	0.033	0.033	0.037	0.037	0.524	0.524

Table **6**: **Evidence from directional events and market reactions**

This table reports OLS regression estimates of the impact of *Human* and *CloudMachine* viewership activities of firm i's 8-K filings on days $t \in \{0, 1\}$ (relative to the 8-K publication date) on post-announcement price inefficiency. $PI(2, 10)$ and $PI(2, 20)$ are measures of price inefficiency over different horizons, as described in Equation (2). *CloudMachine* is the machine viewership by cloud computing services; and *Human* is the human viewership by financial institutions. We consider an event to be positive (negative) news if CAR(0, 1), the Cumulative Abnormal Return over the release date and the following day, is more than (less than) zero. We then categorize a negative event to have a price drift if $CAR(0, X) < CAR(0, 1)$ and price reversal if $CAR(0, X) > CAR(0, 1)$, whereas we classify a positive event to have a price drift if $CAR(0, X) > CAR(0, 1)$ and price reversal if $CAR(0, X) < CAR(0, 1)$. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, *InstOwn*, *Analysts*, *Spread*, *Friday*, *#News*, $|CAR(0, 1)|$, *FinNeg*, and *#Items*. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Drift				Reversal			
	$PI(2, 10)$		$PI(2, 20)$		$PI(2, 10)$		$PI(2, 20)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Negative News								
Human	0.002***		0.003***		0.002**		0.003***	
	(5.79)		(5.25)		(2.92)		(4.25)	
CloudMachine		-0.001**		-0.001*		-0.003***		-0.003***
		(-2.88)		(-1.85)		(-3.27)		(-3.38)
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	77,581	77,581	77,581	77,581	70,594	70,594	70,594	70,594
Adj. R^2	0.237	0.237	0.270	0.270	0.247	0.246	0.264	0.263
Panle B: Positive News								
Human	0.002***		0.004***		0.002***		0.003***	
	(3.18)		(3.16)		(3.18)		(3.33)	
CloudMachine		-0.002*		-0.002		-0.001*		-0.000
		(-2.04)		(-1.36)		(-2.13)		(-0.58)
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	72,791	72,791	72,791	72,791	70,997	70,997	70,997	70,997
Adj. R^2	0.249	0.249	0.269	0.269	0.241	0.241	0.255	0.254

Table 7: **Evidence from exogenous cloud outages**

The table presents estimates from 2-Stage Least Squares (2SLS) analysis using cloud outages as the instrumental variable. Columns (1-3) present estimates from the first stage and columns (4-9) present estimates from the second stage. *Machine*, *CloudMachine*, and *Human* denote machine, cloud machine, and human viewership activities of 8-K filings on days $t \in \{0, 1\}$. Cloud outage is a day with a service disruption to one or more major cloud service providers. $PI(2, 10)$ and $PI(2, 20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0, 1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	First Stage			Second Stage					
	Machine	*CloudMachine*	*Human*	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Machine				-0.050*			-0.002		
				(-2.03)			(-0.06)		
CloudMachine					-0.036**			-0.001	
					(-2.67)			(-0.06)	
Human						0.158			0.006
						(1.19)			(0.06)
CloudOutage	-0.050**	-0.068***	0.016						
	(-2.10)	(-3.63)	(1.57)						
BM	-0.000	-0.006	0.068***	0.004***	0.004***	-0.007	0.005***	0.005***	0.005
	(-0.06)	(-0.84)	(7.24)	(6.55)	(6.35)	(-0.68)	(5.61)	(5.41)	(0.69)
SIZE	0.013	-0.009	0.076***	-0.004***	-0.005***	-0.017	-0.008***	-0.008***	-0.009
	(1.45)	(-1.19)	(6.48)	(-3.44)	(-5.45)	(-1.57)	(-8.19)	(-7.58)	(-1.09)
ROA	0.012	0.010	-0.118***	-0.026***	-0.026***	-0.008	-0.037***	-0.037***	-0.036**
	(0.59)	(0.43)	(-3.10)	(-6.03)	(-6.86)	(-0.47)	(-6.87)	(-6.93)	(-2.66)
LEV	0.004	0.016	0.325***	0.014**	0.014***	-0.038	0.024***	0.024***	0.022
	(0.12)	(0.55)	(6.24)	(2.77)	(3.13)	(-0.81)	(5.15)	(5.24)	(0.63)
STD_{RET}	0.033	0.075**	0.303***	0.029***	0.030***	-0.020	0.046***	0.046***	0.044
	(1.16)	(2.24)	(5.59)	(3.19)	(3.25)	(-0.51)	(3.34)	(3.37)	(1.23)
InstOwn	-0.022	-0.039	-0.128**	-0.008*	-0.008**	0.013	-0.014***	-0.014***	-0.013
	(-0.73)	(-0.96)	(-2.57)	(-2.07)	(-2.46)	(0.62)	(-3.66)	(-3.64)	(-0.98)
Analysts	0.004	0.015*	-0.008	0.001	0.001	0.002	-0.000	-0.000	-0.000
	(0.44)	(1.94)	(-0.70)	(0.69)	(1.28)	(1.15)	(-0.11)	(-0.09)	(-0.06)
Spread	0.687	-1.448***	-1.984**	0.655***	0.568***	0.934**	0.676***	0.673***	0.687***
	(0.50)	(-2.79)	(-2.39)	(3.72)	(4.12)	(2.82)	(5.63)	(4.51)	(3.88)
Friday	-0.043***	-0.008	-0.121***	-0.001	0.001**	0.021	0.002	0.002***	0.003
	(-3.90)	(-0.69)	(-12.20)	(-0.43)	(2.32)	(1.27)	(1.33)	(4.26)	(0.23)
$\#News$	-0.001	-0.010*	0.057***	-0.001**	-0.001***	-0.010	-0.002***	-0.002***	-0.003
	(-0.22)	(-1.84)	(9.51)	(-2.41)	(-3.39)	(-1.30)	(-5.04)	(-4.72)	(-0.47)
$\|CAR(0,1)\|$	0.215**	-0.053	1.521***	0.116***	0.103***	-0.134	0.109***	0.109***	0.100
	(2.52)	(-1.61)	(10.95)	(8.86)	(9.05)	(-0.68)	(10.48)	(8.20)	(0.62)
FinNeg	2.069*	1.284	5.895***	0.142*	0.086	-0.890	0.094	0.092	0.056
	(1.73)	(0.93)	(5.32)	(1.84)	(1.60)	(-0.98)	(1.19)	(1.27)	(0.09)
$\#Items$	0.088***	0.062***	0.153***	0.004*	0.002**	-0.024	-0.001	-0.001	-0.002
	(10.49)	(5.96)	(18.74)	(1.98)	(2.57)	(-1.19)	(-0.23)	(-0.47)	(-0.11)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295142	295142	295142	295142	295142	295142	295142	295142	295142
Weak ID F-Value	4.361	13.157	2.661	-	-	-	-	-	-

Table 8: **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 *PI*. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement *PI* over different horizons, as described in Equation (2). *Machine*, *CloudMachine*, and *Human* denote machine, cloud machine and human viewership of 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. Control variables include *BM*, *SIZE*, *ROA*, *LEV*, STD_{RET}, *InstOwn*, *Analysts*, *Spread*, *Friday*, *#News*, $|CAR(0,1)|$, *FinNeg*, and *#Items*. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
Machine	0.003			0.005		
	(1.09)			(1.24)		
Machine × *InIndex*	-0.000			0.000		
	(-0.24)			(0.06)		
CloudMachine		-0.002***			-0.002**	
		(-3.35)			(-2.32)	
CloudMachine × *InIndex*		0.000			0.000	
		(1.39)			(0.95)	
Human			0.002***			0.004***
			(4.98)			(5.14)
Human × *InIndex*			-0.001**			-0.001*
			(-2.18)			(-1.82)
InIndex	0.003	0.002	0.004**	0.006**	0.005**	0.008***
	(1.58)	(0.99)	(2.46)	(2.56)	(2.51)	(3.39)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.54)	(6.61)	(6.60)	(5.53)	(5.53)	(5.44)
SIZE	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.009***
	(-5.14)	(-4.98)	(-5.20)	(-8.65)	(-8.20)	(-8.70)
ROA	-0.027***	-0.027***	-0.026***	-0.037***	-0.037***	-0.036***
	(-6.42)	(-6.45)	(-6.41)	(-7.02)	(-7.00)	(-6.99)
LEV	0.014***	0.014***	0.013***	0.023***	0.023***	0.022***
	(3.37)	(3.33)	(3.21)	(5.21)	(5.13)	(5.02)
STD_{RET}	0.028***	0.028***	0.027***	0.046***	0.046***	0.045***
	(3.14)	(3.16)	(3.08)	(3.29)	(3.30)	(3.24)
InstOwn	-0.007*	-0.007**	-0.006*	-0.013***	-0.013***	-0.013***
	(-2.16)	(-2.17)	(-2.09)	(-3.54)	(-3.59)	(-3.48)
Analysts	0.000	0.001	0.001	-0.000	-0.000	-0.000
	(0.60)	(0.63)	(0.64)	(-0.14)	(-0.11)	(-0.10)
Spread	0.618***	0.617***	0.624***	0.670***	0.670***	0.681***
	(4.59)	(4.54)	(4.56)	(5.20)	(5.22)	(5.23)
Friday	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(3.95)	(4.04)	(4.42)	(4.83)	(5.07)	(5.45)
#News	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***	-0.003***
	(-3.62)	(-3.62)	(-3.72)	(-5.13)	(-5.04)	(-5.01)
$\|CAR(0,1)\|$	0.105***	0.105***	0.102***	0.108***	0.109***	0.104***
	(9.53)	(9.25)	(9.27)	(9.18)	(8.71)	(8.82)
FinNeg	0.034	0.041	0.027	0.081	0.093	0.071
	(0.70)	(0.84)	(0.55)	(1.07)	(1.15)	(0.90)
#Items	-0.000	-0.000	-0.000***	-0.001***	-0.001	-0.001***
	(-1.74)	(-0.09)	(-3.20)	(-3.77)	(-1.76)	(-3.84)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.221	0.221	0.222	0.243	0.243	0.243

Table **9**: **Viewership of information and informed trading**

The table presents estimates of the impact of machine, cloud machine, and human viewership on the Probability of Informed Trading (PIN) on days $(0,1)$ (Columns 1-3) and $(0,5)$ (Columns 4-6), relative to the 8-K publication date. $PIN(0,1)$ and $PIN(0,5)$ is the average of daily probability of information-based trading (PIN) during windows $(0,1)$ and $(0,5)$, relative to the 8-K publication date. $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PIN(0,1)$			$PIN(0,5)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.002***			0.002***		
	(3.27)			(3.12)		
$CloudMachine$		0.002**			0.002**	
		(2.52)			(2.46)	
$Human$			-0.002***			-0.002***
			(-5.24)			(-3.67)
BM	-0.002*	-0.002*	-0.002*	-0.004**	-0.004**	-0.004**
	(-1.95)	(-1.94)	(-1.83)	(-2.74)	(-2.73)	(-2.63)
$SIZE$	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***
	(-22.46)	(-22.44)	(-22.22)	(-24.48)	(-24.42)	(-24.22)
ROA	-0.006	-0.006	-0.007	-0.007*	-0.007*	-0.008*
	(-1.65)	(-1.65)	(-1.70)	(-1.89)	(-1.88)	(-1.93)
LEV	-0.012***	-0.012***	-0.012***	-0.010***	-0.010***	-0.010***
	(-3.44)	(-3.45)	(-3.24)	(-3.29)	(-3.30)	(-3.10)
STD_{RET}	-0.044***	-0.045***	-0.044***	-0.042***	-0.042***	-0.041***
	(-6.25)	(-6.26)	(-6.16)	(-6.37)	(-6.38)	(-6.27)
$InstOwn$	-0.100***	-0.100***	-0.100***	-0.098***	-0.098***	-0.099***
	(-23.47)	(-23.44)	(-23.47)	(-24.19)	(-24.20)	(-24.24)
$Analysts$	-0.015***	-0.015***	-0.015***	-0.014***	-0.014***	-0.014***
	(-10.80)	(-10.80)	(-10.75)	(-9.47)	(-9.44)	(-9.44)
$Spread$	1.139***	1.145***	1.136***	0.872***	0.878***	0.870***
	(9.56)	(9.63)	(9.60)	(6.66)	(6.62)	(6.53)
$Friday$	-0.000	-0.000	-0.001	0.001	0.001	0.000
	(-0.56)	(-0.70)	(-1.14)	(1.58)	(1.39)	(0.92)
$\#News$	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(-6.41)	(-6.32)	(-6.18)	(-8.89)	(-8.72)	(-8.53)
$\|CAR(0,1)\|$	-0.103***	-0.102***	-0.100***	-0.099***	-0.098***	-0.096***
	(-6.66)	(-6.60)	(-6.49)	(-6.38)	(-6.28)	(-6.24)
$FinNeg$	0.093**	0.095**	0.108***	0.073*	0.075*	0.087**
	(2.82)	(2.90)	(3.10)	(2.05)	(2.13)	(2.36)
$\#Items$	-0.001**	-0.001*	-0.000	-0.001**	-0.001*	-0.000
	(-2.42)	(-2.14)	(-1.13)	(-2.28)	(-1.96)	(-0.96)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	271,079	271,079	271,079	277,526	277,526	277,526
Adj. R^2	0.633	0.633	0.633	0.686	0.686	0.686

Table **10**: **High-frequency traders (HFT) and price inefficiency**

The table presents estimates of the impact of High-Frequency Traders (HFT) on the *PI* following 8-K filings. *Cancel-to-Trade* is a proxy for HFT activity and *Trade-to-Order* is an inverse measure of HFT activity. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement *PI* over different horizons, as described in Equation (2). $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. Control variables in all regression include BM, $SIZE$, ROA, LEV, STD_{RET}, *InstOwn*, *Analysts*, *Spread*, *Friday*, *#News*, $|CAR(0,1)|$, $FinNeg$, and *#Items*. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$						$PI(2,20)$					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$Machine$	-0.001	-0.001					-0.001	-0.001				
	(-1.37)	(-1.38)					(-0.92)	(-1.02)				
$CloudMachine$			-0.002**	-0.002**					-0.002*	-0.002*		
			(-3.02)	(-2.80)					(-2.27)	(-2.15)		
$Human$					0.001**	0.001**					0.001**	0.001**
					(3.69)	(2.71)					(2.90)	(2.57)
Cancel-to-Trade	0.000		0.000		0.000		0.000		0.000		0.000	
	(0.18)		(0.16)		(0.15)		(0.69)		(0.67)		(0.66)	
Trade-to-Order		0.001**		0.001**		0.001**		0.001**		0.001**		0.001**
		(2.61)		(2.70)		(2.72)		(3.99)		(3.96)		(3.88)
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	101,368	101,368	101,368	101,368	101,368	101,368	101,368	101,368	101,368	101,368	101,368	101,368
Adj. R^2	0.250	0.250	0.249	0.250	0.250	0.250	0.279	0.279	0.279	0.279	0.279	0.279

Table **11**: **Does numerical information increase the impact of machines?**

This table reports OLS regression estimates of the impact of $Machine$, $CloudMachine$, and $Human$ viewership of 8-Ks, when interacted with the percentage of numerical information in the 8-K, NER. $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of firm i's 8-K filings on days $t \in \{0, 1\}$. $PI(2, 10)$ and $PI(2, 20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). NER is the percentage of numerical entities, including quantities, dollar amounts, and cardinal values contained in the sentence components of the 8-K filing. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0, 1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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. 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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2, 10)$			$PI(2, 20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.003			0.005		
	(1.24)			(1.45)		
$Machine \times NER$	-0.006			-0.012		
	(-1.27)			(-1.22)		
$CloudMachine$		-0.002**			-0.001	
		(-2.87)			(-1.66)	
$CloudMachine \times NER$		-0.011**			-0.015**	
		(-2.53)			(-2.21)	
$Human$			0.002***			0.004***
			(5.02)			(5.72)
$Human \times NER$			-0.008			-0.015
			(-1.19)			(-1.42)
NER	-0.022	-0.022	-0.039**	-0.053	-0.064*	-0.084**
	(-0.94)	(-1.48)	(-2.60)	(-1.04)	(-2.01)	(-2.54)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.87)	(6.95)	(6.96)	(5.62)	(5.59)	(5.56)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.008***
	(-5.20)	(-5.06)	(-5.28)	(-8.86)	(-8.38)	(-8.96)
ROA	-0.026***	-0.026***	-0.026***	-0.037***	-0.037***	-0.036***
	(-6.33)	(-6.35)	(-6.33)	(-7.08)	(-7.05)	(-7.05)
LEV	0.014***	0.014***	0.013***	0.023***	0.023***	0.022***
	(3.29)	(3.25)	(3.13)	(5.13)	(5.04)	(4.95)
STD_{RET}	0.027***	0.028***	0.027***	0.046***	0.046***	0.045***
	(3.12)	(3.15)	(3.06)	(3.32)	(3.33)	(3.25)
$InstOwn$	-0.006*	-0.007*	-0.006*	-0.013***	-0.013***	-0.013***
	(-2.08)	(-2.14)	(-2.01)	(-3.58)	(-3.66)	(-3.49)
$Analysts$	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.52)	(0.56)	(0.57)	(-0.22)	(-0.19)	(-0.17)
$Spread$	0.618***	0.617***	0.625***	0.669***	0.669***	0.680***
	(4.58)	(4.53)	(4.55)	(5.15)	(5.17)	(5.16)
$Friday$	0.002***	0.001***	0.002***	0.002***	0.002***	0.002***
	(3.99)	(3.96)	(4.46)	(4.35)	(3.99)	(4.74)
$\#News$	-0.001**	-0.001**	-0.001**	-0.002***	-0.002***	-0.002***
	(-2.64)	(-2.58)	(-3.00)	(-4.21)	(-4.12)	(-4.48)
$\|CAR(0, 1)\|$	0.106***	0.106***	0.103***	0.110***	0.111***	0.106***
	(9.52)	(9.24)	(9.26)	(9.11)	(8.63)	(8.77)
$FinNeg$	0.015	0.024	0.002	0.044	0.057	0.023
	(0.33)	(0.54)	(0.04)	(0.69)	(0.83)	(0.34)
$\#Items$	-0.000	0.000	-0.000**	-0.001***	-0.000	-0.001***
	(-1.17)	(0.62)	(-2.44)	(-3.34)	(-1.18)	(-3.19)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	291,848	291,848	291,848	291,848	291,848	291,848
Adj. R^2	0.222	0.221	0.222	0.243	0.243	0.243

Table **12**: **How does negative sentiment affect humans and machines?**

The table presents estimates of the effect of negative sentiment on how humans and machines impact PI following 8-K filings. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. $FinNeg$ is the proportion of negative words defined by Loughran and Mcdonald (2011) in an 8-K filing. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.003			0.005		
	(1.19)			(1.29)		
$Machine \times FinNeg$	-0.076**			-0.105*		
	(-2.24)			(-1.88)		
$CloudMachine$		-0.002**			-0.002*	
		(-2.93)			(-1.89)	
$CloudMachine \times FinNeg$		-0.045**			-0.079**	
		(-2.71)			(-2.37)	
$Human$			0.002***			0.004***
			(3.77)			(4.03)
$Human \times FinNeg$			-0.029			-0.055
			(-1.00)			(-1.48)
$FinNeg$	0.316**	0.148**	0.075	0.470*	0.279*	0.162
	(2.18)	(2.27)	(1.07)	(1.88)	(2.10)	(1.48)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.78)	(6.84)	(6.83)	(5.64)	(5.62)	(5.60)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.008***
	(-5.29)	(-5.12)	(-5.25)	(-8.83)	(-8.34)	(-8.70)
ROA	-0.027***	-0.027***	-0.027***	-0.037***	-0.037***	-0.037***
	(-6.42)	(-6.44)	(-6.42)	(-7.00)	(-6.96)	(-6.98)
LEV	0.014***	0.014***	0.013***	0.024***	0.024***	0.022***
	(3.37)	(3.34)	(3.23)	(5.19)	(5.11)	(5.03)
STD_{RET}	0.028***	0.028***	0.027***	0.046***	0.046***	0.045***
	(3.14)	(3.16)	(3.08)	(3.29)	(3.30)	(3.23)
$InstOwn$	-0.007**	-0.007**	-0.007*	-0.014***	-0.014***	-0.014***
	(-2.20)	(-2.26)	(-2.14)	(-3.71)	(-3.79)	(-3.62)
$Analysts$	0.000	0.001	0.001	-0.000	-0.000	-0.000
	(0.58)	(0.64)	(0.64)	(-0.15)	(-0.11)	(-0.10)
$Spread$	0.619***	0.618***	0.625***	0.671***	0.672***	0.682***
	(4.60)	(4.55)	(4.56)	(5.22)	(5.24)	(5.24)
$Friday$	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(3.94)	(4.02)	(4.48)	(4.84)	(5.00)	(5.44)
$\#News$	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***	-0.003***
	(-3.64)	(-3.57)	(-3.83)	(-5.28)	(-5.10)	(-5.16)
$\lvert CAR(0,1)\rvert$	0.105***	0.105***	0.102***	0.108***	0.109***	0.104***
	(9.54)	(9.27)	(9.30)	(9.20)	(8.72)	(8.88)
$\#Items$	-0.000	0.000	-0.000***	-0.001***	-0.001	-0.001***
	(-1.52)	(0.13)	(-3.09)	(-3.53)	(-1.49)	(-3.77)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.222	0.221	0.222	0.243	0.243	0.243

Table **13**: **Incremental earnings information and price inefficiency**

The table presents estimates of how incremental earnings information impacts PI post to 8-K publication and how well humans and machines assimilate such information. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. $Item2.02$ is a dummy variable that is one if an 8-K contains $Item2.02$, and zero otherwise. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.003			0.004		
	(1.12)			(1.23)		
$Machine \times Item2.02$	-0.000			0.000		
	(-0.15)			(0.34)		
$CloudMachine$		-0.002***			-0.002**	
		(-3.18)			(-2.39)	
$CloudMachine \times Item2.02$		-0.000			-0.000	
		(-1.17)			(-0.20)	
$Human$			0.003***			0.004***
			(5.15)			(5.20)
$Human \times Item2.02$			-0.001***			-0.001**
			(-3.03)			(-2.48)
$Item2.02$	-0.003	-0.002*	-0.002**	-0.006*	-0.006**	-0.005**
	(-1.53)	(-1.96)	(-2.80)	(-1.82)	(-2.43)	(-2.84)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.70)	(6.80)	(6.81)	(5.64)	(5.63)	(5.60)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.009***
	(-5.29)	(-5.18)	(-5.38)	(-8.96)	(-8.52)	(-9.13)
ROA	-0.027***	-0.027***	-0.026***	-0.037***	-0.037***	-0.036***
	(-6.40)	(-6.42)	(-6.41)	(-7.01)	(-6.97)	(-7.00)
LEV	0.014***	0.014***	0.013***	0.023***	0.024***	0.022***
	(3.36)	(3.33)	(3.20)	(5.20)	(5.12)	(5.01)
STD_{RET}	0.027***	0.028***	0.027***	0.046***	0.046***	0.045***
	(3.14)	(3.16)	(3.07)	(3.29)	(3.30)	(3.23)
$InstOwn$	-0.007**	-0.007**	-0.006*	-0.014***	-0.014***	-0.013***
	(-2.18)	(-2.26)	(-2.11)	(-3.67)	(-3.76)	(-3.59)
$Analysts$	0.001	0.001	0.001	-0.000	-0.000	-0.000
	(0.61)	(0.63)	(0.65)	(-0.10)	(-0.07)	(-0.08)
$Spread$	0.620***	0.619***	0.627***	0.673***	0.674***	0.685***
	(4.59)	(4.55)	(4.56)	(5.19)	(5.22)	(5.23)
$Friday$	0.001***	0.001***	0.001***	0.002***	0.001***	0.002***
	(3.27)	(3.13)	(3.71)	(3.72)	(3.31)	(4.15)
$\#News$	-0.000	-0.000	-0.000*	-0.001***	-0.001***	-0.001***
	(-1.74)	(-1.56)	(-1.97)	(-3.58)	(-3.48)	(-3.65)
$\|CAR(0,1)\|$	0.107***	0.107***	0.104***	0.111***	0.112***	0.107***
	(9.68)	(9.37)	(9.41)	(9.27)	(8.74)	(8.90)
$FinNeg$	0.009	0.019	-0.004	0.035	0.048	0.017
	(0.23)	(0.47)	(-0.09)	(0.59)	(0.74)	(0.28)
$\#Items$	0.000	0.000*	0.000	-0.000	0.000	-0.000
	(0.47)	(1.94)	(0.24)	(-1.23)	(0.33)	(-1.01)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.222	0.222	0.222	0.243	0.243	0.244

Appendix Table A: Variables, definitions, and sources

Variable	Definition	Source
	Panel A: *Viewership Variables*	
$Machine$	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
$Human$	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
$TotalVisit$	The natural logarithm of both machine and human visits to an 8-K filing in day 0 and 1.	SEC Log File
$Machine_{DRT}$	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
$Human_{DRT}$	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
$CloudMachine$	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
$Machine/Total$	The ratio of machine viewership to total visits of 8K filings on days $t \in \{0,1\}$ relative to 8K publication date.	Our estimates
$CloudMachine/Total$	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
$InstInvMachine$	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
$DataVendMachine$	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
$OtherMachine$	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
	Panel B: *Dependent Variables / Informational efficiency variables*	
$PI(2,T)$	The post-announcement drift or reversal, measured as an absolute cumulative abnormal return (absolute difference between $CAR(0,T)$ and $CAR(0,2)$, where cumulative abnormal returns (CAR) are from a Fama-French 3 Factor + Momentum model).	CRSP & Compustat
$ScaledPI(2,T)$	$\lvert CAR(0,T) - CAR(0,1)\rvert / \lvert CAR(0,T)\rvert$, for $T = 10, 20$.	CRSP
$Noise$	The share of stock return variance that is attributable to noise using the method in Brogaard et al. (2022).	CRSP & Compustat
$PIN(0,T)$	Average of daily probability of information-based trading (PIN) over window (0,T). PIN is calculated daily using 5-minute intervals	TAQ
Cancel-to-Trade	The average daily number of cancellations divided by the number of trades, for displayed orders over the day 0 and 1 relative to 8-K publication date.	SEC MIDAS
Trade-to-Order	The average daily total volume traded divided by the total volume from all orders placed over the day 0 and 1 relative to 8-K publication date.	SEC MIDAS
$OddLotRatio(0,1)$	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
$TradeSize(0,1)$	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 sources (Continued)

Variable	Definition	Source
	Panel C: *8-K variables*	
$\#Items$	Number of topics included in an 8-K filing	WRDS SEC Suite
$\#Words$	Number of words used in an 8-K filing	WRDS SEC Suite
$FinNeg$	The proportion of negative words defined by Loughran and Mcdonald (2011) in an 8-K filing.	WRDS SEC Suite
Gap	The number of days between 8-K event date to publication date.	WRDS SEC Suite
$Item\#.\#$	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*	
LEV	Total debt over total assets	Compustat
BM	The ratio of book value over the market value of common equity in the year end prior to an 8-K filing	CRSP & Compustat
$SIZE$	The natural logarithm of market value of common equity in the year end prior to an 8-K filing	CRSP
$InstOwn$	The institutional ownership percentage calculated in the quarter prior to an 8-K filing	Factset 13F
$Analysts$	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
STD_{RET}	Standard Deviation of Monthly Return over the year prior to an 8-K filing	CRSP
$InIndex$	A dummy variable equal to one if a stock has been added into S&P500 index, zero otherwise.	CRSP
$Spread$	Averaged percent effective spread over the week before the 8-K publication date.	TAQ
$\#News$	Number of news mentions within two weeks before the 8-K publication date.	RavenPack RPA
$\lvert CAR(0,1)\rvert$	Absolute value of $CAR(0,1)$.	CRSP
$Friday$	A dummy variable equal to one if the information is viewed on a Friday, zero otherwise.	CRSP
	Panel E: *Instrumental Variables*	
$InvSent$	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
$MacroNews$	A dummy variable equal to one if a major macroeconomic indicator update is announced the 8-K publication data, zero otherwise.	Bloomberg
$CloudOutage$	A day with a service disruption to one or more major cloud service providers.	Gunawi et al. (2016)
	Panel F: *Numerical content Measures*	
NER	Percentage of numerical entities, including quantities, dollar amounts, and cardinal values in the 8-K filing. We rely on a Python package, spaCy, to parse sentences in the 8-K and require the sentence must include at least a noun (or pronoun) and a verb.	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 using alternative measures of machine and cloud machine viewership of firm i's 8-K filings on days $t \in \{0, 1\}$ (relative to the 8-K publication date). $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). $Machine/Total$ and $CloudMachine/Total$ denote fractions of viewership by machines and cloud machines, and $TotalVisits$ is the natural log of total viewership. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$		$PI(2,20)$	
	(1)	(2)	(3)	(4)
$Machine/Total$	-0.005		-0.007	
	(-1.20)		(-1.15)	
$CloudMachine/Total$		-0.014**		-0.018*
		(-2.29)		(-2.03)
$TotalVisits$	0.003	0.003	0.005	0.005
	(1.20)	(1.42)	(1.40)	(1.66)
BM	0.003***	0.003***	0.004***	0.004***
	(7.38)	(7.66)	(6.40)	(6.39)
$SIZE$	-0.006***	-0.006***	-0.010***	-0.010***
	(-7.77)	(-7.95)	(-10.91)	(-11.34)
ROA	-0.024***	-0.024***	-0.037***	-0.037***
	(-6.50)	(-6.59)	(-6.91)	(-6.96)
LEV	0.011***	0.011***	0.022***	0.022***
	(3.25)	(3.19)	(4.96)	(4.89)
STD_{RET}	0.029***	0.029***	0.044***	0.044***
	(3.82)	(3.84)	(3.26)	(3.28)
$InstOwn$	-0.008**	-0.007**	-0.015***	-0.015***
	(-2.84)	(-2.84)	(-4.03)	(-4.04)
$Analysts$	-0.000	-0.000	-0.000	-0.000
	(-0.02)	(-0.00)	(-0.37)	(-0.36)
$Spread$	0.139**	0.137**	0.159**	0.157**
	(2.35)	(2.35)	(2.27)	(2.27)
$Friday$	0.002***	0.002***	0.002***	0.003***
	(4.47)	(4.28)	(5.28)	(5.02)
$\#News$	-0.001***	-0.001***	-0.002***	-0.002***
	(-3.88)	(-3.59)	(-5.37)	(-4.91)
$\|CAR(0,1)\|$	0.096***	0.096***	0.105***	0.105***
	(10.00)	(10.37)	(8.94)	(9.47)
$FinNeg$	0.023	0.020	0.075	0.071
	(0.53)	(0.47)	(1.00)	(0.98)
$\#Items$	-0.000	-0.000	-0.001***	-0.001***
	(-1.71)	(-1.70)	(-4.08)	(-4.18)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142
Adj. R^2	0.230	0.230	0.243	0.244

Appendix Table D: Additional evidence from cloud outages

The table presents additional estimates of how cloud outages affect the information assimilation of machines. *Cloud Outage Day* is a day with service disruption to one or more major cloud service providers. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement PI over different horizons, as described in Equation (2). $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$			$PI(2,20)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.003			0.005		
	(1.09)			(1.22)		
$Machine \times CloudOutage$	0.001*			0.002*		
	(1.82)			(1.87)		
$CloudMachine$		-0.002***			-0.002**	
		(-3.39)			(-2.40)	
$CloudMachine \times CloudOutage$		0.005***			0.006***	
		(6.91)			(3.08)	
$Human$			0.002***			0.003***
			(4.23)			(4.22)
$Human \times CloudOutage$			0.000			0.001
			(0.12)			(0.86)
$CloudOutage$	-0.002	-0.012***	0.002	-0.009*	-0.018**	-0.001
	(-0.86)	(-5.40)	(1.52)	(-1.82)	(-2.80)	(-0.55)
BM	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***
	(6.73)	(6.82)	(6.81)	(5.64)	(5.64)	(5.60)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.008***
	(-5.20)	(-5.05)	(-5.24)	(-8.65)	(-8.19)	(-8.68)
ROA	-0.027***	-0.027***	-0.027***	-0.037***	-0.037***	-0.037***
	(-6.42)	(-6.45)	(-6.42)	(-7.01)	(-6.98)	(-6.98)
LEV	0.014***	0.014***	0.013***	0.023***	0.024***	0.022***
	(3.37)	(3.34)	(3.21)	(5.20)	(5.11)	(5.02)
STD_{RET}	0.028***	0.028***	0.027***	0.046***	0.046***	0.045***
	(3.14)	(3.17)	(3.08)	(3.29)	(3.31)	(3.23)
$InstOwn$	-0.007**	-0.007**	-0.007*	-0.014***	-0.014***	-0.013***
	(-2.18)	(-2.26)	(-2.12)	(-3.69)	(-3.77)	(-3.60)
$Analysts$	0.000	0.001	0.001	-0.000	-0.000	-0.000
	(0.60)	(0.66)	(0.64)	(-0.13)	(-0.09)	(-0.09)
$Spread$	0.619***	0.618***	0.625***	0.672***	0.672***	0.682***
	(4.60)	(4.55)	(4.55)	(5.22)	(5.26)	(5.24)
$Friday$	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(3.88)	(3.94)	(4.38)	(4.78)	(4.97)	(5.42)
$\#News$	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***	-0.003***
	(-3.60)	(-3.57)	(-3.81)	(-5.17)	(-5.04)	(-5.14)
$\|CAR(0,1)\|$	0.105***	0.105***	0.102***	0.108***	0.109***	0.104***
	(9.53)	(9.25)	(9.30)	(9.18)	(8.71)	(8.88)
$FinNeg$	0.034	0.043	0.028	0.081	0.094	0.071
	(0.72)	(0.87)	(0.56)	(1.07)	(1.17)	(0.90)
$\#Items$	-0.000	-0.000	-0.000***	-0.001***	-0.001*	-0.001***
	(-1.76)	(-0.12)	(-3.20)	(-3.81)	(-1.80)	(-3.84)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.221	0.221	0.222	0.243	0.243	0.243

Appendix Table E: Impact of by-facility human viewership.

The table presents additional estimates of how by-facility human viewership of 8-K filings on days $t \in \{0,1\}$ through cloud machine, institutional investor machine and data vendor machine, affects the information assimilation of humans. $PI(2,10)$ and $PI(2,20)$ are measures of the post-announcement drift or reversal over different horizons, as described in Equation (2). $Human(CloudMachine)$ is the human viewership through cloud machines, $Human(InstInvMachine)$ is the human viewership through institutional investor machines, and $Human(DataVendMachine)$ is the human viewership through data vendor machines. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$PI(2,10)$				$PI(2,20)$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Human(CloudMachine)$	0.001***			0.001***	0.002***			0.002***
	(4.85)			(4.52)	(5.57)			(4.51)
$Human(InstInvMachine)$		0.001**		0.001**		0.002**		0.002**
		(2.56)		(2.21)		(2.96)		(2.51)
$Human(DataVendMachine)$			0.001*	-0.000			0.001**	0.000
			(2.00)	(-0.42)			(2.81)	(0.73)
BM	0.004***	0.004***	0.004***	0.004***	0.005***	0.005***	0.005***	0.005***
	(6.74)	(6.80)	(6.77)	(6.77)	(5.60)	(5.65)	(5.63)	(5.63)
$SIZE$	-0.005***	-0.005***	-0.005***	-0.005***	-0.008***	-0.008***	-0.008***	-0.008***
	(-5.02)	(-5.08)	(-5.01)	(-5.09)	(-8.16)	(-8.26)	(-8.13)	(-8.28)
ROA	-0.027***	-0.027***	-0.027***	-0.027***	-0.037***	-0.037***	-0.037***	-0.037***
	(-6.42)	(-6.44)	(-6.42)	(-6.44)	(-6.98)	(-7.00)	(-6.98)	(-7.01)
LEV	0.014***	0.013***	0.014***	0.013***	0.023***	0.023***	0.023***	0.023***
	(3.31)	(3.31)	(3.34)	(3.29)	(5.10)	(5.14)	(5.14)	(5.11)
STD_{RET}	0.028***	0.028***	0.028***	0.027***	0.046***	0.046***	0.046***	0.046***
	(3.14)	(3.13)	(3.15)	(3.13)	(3.29)	(3.29)	(3.29)	(3.28)
$InstOwn$	-0.007**	-0.007**	-0.007**	-0.007**	-0.014***	-0.014***	-0.014***	-0.014***
	(-2.20)	(-2.18)	(-2.21)	(-2.17)	(-3.69)	(-3.68)	(-3.72)	(-3.66)
$Analysts$	0.001	0.000	0.001	0.000	-0.000	-0.000	-0.000	-0.000
	(0.62)	(0.60)	(0.62)	(0.60)	(-0.11)	(-0.15)	(-0.11)	(-0.13)
$Spread$	0.621***	0.621***	0.621***	0.621***	0.676***	0.675***	0.675***	0.676***
	(4.54)	(4.53)	(4.54)	(4.53)	(5.22)	(5.20)	(5.22)	(5.21)
$Friday$	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(4.14)	(4.13)	(4.09)	(4.16)	(5.07)	(5.03)	(4.98)	(5.10)
$\#News$	-0.001***	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***	-0.002***	-0.002***
	(-3.60)	(-3.65)	(-3.56)	(-3.66)	(-5.07)	(-5.09)	(-5.06)	(-5.09)
$\|CAR(0,1)\|$	0.105***	0.105***	0.105***	0.104***	0.108***	0.108***	0.109***	0.107***
	(9.19)	(9.27)	(9.22)	(9.26)	(8.64)	(8.68)	(8.67)	(8.66)
$FinNeg$	0.035	0.036	0.038	0.034	0.085	0.086	0.089	0.083
	(0.70)	(0.72)	(0.76)	(0.68)	(1.05)	(1.07)	(1.09)	(1.02)
$\#Items$	-0.000	-0.000	-0.000	-0.000	-0.001**	-0.001**	-0.001**	-0.001**
	(-1.14)	(-1.47)	(-1.01)	(-1.55)	(-2.28)	(-2.52)	(-2.21)	(-2.55)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	295,142	295,142	295,142	295,142	295,142	295,142	295,142	295,142
Adj. R^2	0.221	0.221	0.221	0.221	0.242	0.242	0.242	0.243

Appendix F: Further evidence from alternative instrumental variables

As an additional instrumental variables test to address potential endogeneity, we use 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)) and macroeconomic news announcements (Peng and Xiong, 2006; Kacperczyk et al., 2016) are likely to disproportionately affect humans' decisions. 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 (Barbopoulos et al., 2020). 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:

$$V_{i,t}^{H} = \alpha_0 \quad + \quad \beta_1 InvSent + \beta_2 MacroNews + \sum_{j=1}^{k} \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \tag{6}$$

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. $InvSent$ and $MacroNews$ are our instruments. $\Gamma_{i,j,t}$ is a vector of controls j that include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. $\tilde{f}$ and $\tilde{\tau}$ are firm and year fixed effects.

Appendix Table F reports the results. Column (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). While all sentiment metrics suggested by previous research are significant, we have an over-identification 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) indicate all sentiment instrumental variables have a comparable impact on human viewership.

To formally test the relevant condition of (the) instrumental variables, we report Kleibergen-Paap F-statistics (for non-i.i.d. errors) in Table E. The F-statistic in Column (3) for the first stage is about 28.9 for human viewership, which is much greater than Stock and Yogo (2002)'s critical values of 5% maximal IV relative bias and 10% maximal IV size tests.[42] Those results imply that human viewership identification does not suffer from a weak instruments problem.

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 instrumental variables do not affect information efficiency through non-human viewership.[43] The second-stage regressions are:

$$PI(2,T)_{i,t} = \alpha_0 + \beta \widehat{V}_{i,t}^{H} + \sum_{j=1}^{k} \varphi_j \Gamma_{i,j,t} + \tilde{f} + \tilde{\tau} + \varepsilon_{i,t} \tag{7}$$

where $PI(2,T)_{i,t}$ 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 F, Columns (4 and 5). They are consistent with our baseline results discussed earlier. Namely, human viewership is significantly positively related to absolute price drift or reversal. Hence, the instrumental variables models reinforce the estimates reported in Table (4) and discussed in Section (*5.1*).

[42] The critical values for the 5% maximal relative bias and the 10% maximal relative size are, respectively, 16.9 and 24.6 in our first-stage sample.

[43] We do not report the second-stage results for those falsification tests, as the coefficients estimated using weak instruments in the second stage would be inefficient, biased, and misleading (Roberts and Whited, 2013).

Appendix Table F: Further evidence from alternative instrumental variables

The table presents estimates from 2-Stage Least Squares (2SLS) analysis using two market sentiment metrics as instrumental variables. $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0, 1\}$. $PI(2, 10)$ and $PI(2, 20)$ are measures of the post-announcement drift or reversal over different horizons, as described in Equation (2). The instruments are $InvSent$ and $MacroNews$. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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. 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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	First Stage			Second Stage	
	$Machine$	$CloudMachine$	$Human$	$PI(2,10)$	$PI(2,20)$
	(1)	(2)	(3)	(4)	(5)
$Human$				0.077**	0.096**
				(2.83)	(2.60)
$InvSent$	-0.306*	0.411***	-0.269***		
	(-1.70)	(3.19)	(-3.31)		
$MacroNews$	0.110**	0.028	0.112**		
	(2.34)	(1.10)	(2.08)		
BM	0.000	-0.007	0.068***	-0.001	-0.001
	(0.05)	(-0.92)	(7.32)	(-0.64)	(-0.55)
$SIZE$	0.010	-0.005	0.073***	-0.011***	-0.015***
	(1.34)	(-0.70)	(6.17)	(-5.22)	(-5.50)
ROA	0.009	0.013	-0.121***	-0.018***	-0.026***
	(0.42)	(0.57)	(-3.17)	(-3.60)	(-3.94)
LEV	0.008	0.014	0.329***	-0.011	-0.008
	(0.26)	(0.47)	(6.44)	(-1.17)	(-0.64)
STD_{RET}	0.032	0.080**	0.302***	0.004	0.017
	(1.14)	(2.37)	(5.70)	(0.33)	(0.89)
$InstOwn$	-0.023	-0.043	-0.129***	0.003	-0.002
	(-0.79)	(-1.10)	(-2.65)	(0.48)	(-0.21)
$Analysts$	0.007	0.011*	-0.005	0.001	0.001
	(0.87)	(1.66)	(-0.48)	(1.34)	(0.52)
$Spread$	0.099	-0.784	-2.511***	0.773***	0.866***
	(0.11)	(-1.28)	(-4.22)	(4.76)	(5.18)
$Friday$	-0.020*	-0.002	-0.097***	0.011***	0.014***
	(-1.67)	(-0.21)	(-6.87)	(3.40)	(3.18)
$\#News$	-0.003	-0.011**	0.055***	-0.005***	-0.008***
	(-0.56)	(-2.12)	(9.61)	(-3.42)	(-3.55)
$\|CAR(0,1)\|$	0.166***	-0.014	1.476***	-0.011	-0.037
	(2.72)	(-0.34)	(10.51)	(-0.34)	(-0.80)
$FinNeg$	2.413**	1.161	6.212***	-0.413**	-0.475**
	(1.99)	(0.93)	(5.56)	(-2.55)	(-2.30)
$\#Items$	0.088***	0.061***	0.153***	-0.012**	-0.016**
	(11.42)	(5.70)	(20.48)	(-2.97)	(-2.90)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
NObs	295142	295142	295142	295142	295142
Weak ID F-Value	4.304	5.179	28.883	-	-
Hansen J-Stats	-	-	-	0.339	0.310

Appendix Table G: Viewership and algorithmic trading

The table presents estimates of the impact of $Machine$, $Cloudmachine$, and $Human$ viewership on algorithmic trading (including the $OddLotRatio(0,1)$ (Columns 1-3) and the $TradeSize(0,1)$ (Columns 4-6)). $OddLotRatio(0,1)$ is the fraction of volume associated with abnormally small trades (less than 100 shares) over the days $t \in \{0,1\}$ relative to 8-K publication date. $TradeSize(0,1)$ is the number of shares traded divided by the number of trades over the same period. $Machine$, $CloudMachine$, and $Human$ denote machine, cloud machine, and human viewership of 8-K filings on days $t \in \{0,1\}$. Control variables in all regressions include BM, $SIZE$, ROA, LEV, STD_{RET}, $InstOwn$, $Analysts$, $Spread$, $Friday$, $\#News$, $|CAR(0,1)|$, $FinNeg$, and $\#Items$. All variables are defined in Appendix Table A. All continuous variables are winsorized at the 1st and 99th percentiles. *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 and year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$OddLotRatio(0,1)$			$TradeSize(0,1)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Machine$	0.004**			-0.003		
	(2.61)			(-1.91)		
$CloudMachine$		0.006*			-0.002*	
		(2.44)			(-2.26)	
$Human$			-0.005**			0.003**
			(-3.45)			(3.69)
BM	-0.006*	-0.006*	-0.006*	0.011**	0.011**	0.011**
	(-2.19)	(-2.20)	(-2.07)	(3.39)	(3.39)	(3.37)
$SIZE$	0.023***	0.023***	0.024***	-0.018***	-0.018***	-0.019***
	(6.69)	(6.75)	(6.57)	(-6.54)	(-6.57)	(-6.67)
ROA	0.004	0.004	0.004	-0.045**	-0.045**	-0.045**
	(0.61)	(0.62)	(0.52)	(-3.29)	(-3.29)	(-3.26)
LEV	-0.032**	-0.032**	-0.030**	0.050***	0.050***	0.049***
	(-2.85)	(-2.83)	(-2.71)	(4.97)	(4.93)	(4.87)
STD_{RET}	-0.005	-0.004	-0.003	-0.034	-0.035	-0.035
	(-0.36)	(-0.34)	(-0.24)	(-1.92)	(-1.93)	(-1.97)
$InstOwn$	0.026	0.026*	0.025	-0.089***	-0.090***	-0.089***
	(1.97)	(2.03)	(1.94)	(-6.86)	(-6.92)	(-6.87)
$Analysts$	0.002	0.002	0.002	-0.007	-0.007	-0.007
	(1.02)	(1.03)	(1.04)	(-1.37)	(-1.37)	(-1.37)
$Spread$	2.351**	2.363**	2.321**	2.890***	2.886***	2.906***
	(3.93)	(3.95)	(3.93)	(10.12)	(10.11)	(10.20)
$Friday$	0.000	0.000	-0.000	-0.002	-0.001	-0.001
	(0.12)	(0.06)	(-0.05)	(-1.19)	(-1.12)	(-0.92)
$\#News$	-0.008***	-0.008***	-0.008***	0.002**	0.001**	0.001**
	(-4.63)	(-4.60)	(-4.36)	(3.72)	(3.47)	(3.15)
$\|CAR(0,1)\|$	-0.123***	-0.122***	-0.112***	0.120***	0.120***	0.115***
	(-9.34)	(-9.33)	(-9.58)	(8.34)	(8.43)	(7.76)
$FinNeg$	-0.282	-0.289	-0.239	0.061	0.062	0.038
	(-1.46)	(-1.55)	(-1.22)	(1.73)	(1.81)	(1.09)
$\#Items$	0.000	0.000	0.002	0.000	0.000	-0.001
	(0.12)	(0.13)	(0.61)	(0.22)	(0.08)	(-0.74)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	101,368	101,368	101,368	101,368	101,368	101,368
Adj. R^2	0.203	0.203	0.203	0.586	0.586	0.586