Human Information Production in the Machine Age: Evidence from Machine Adoption in the Asset Management Industry

Kai Du
Smeal College of Business
Penn State University
kxd30@psu.edu

Miao Liu
Carroll School of Management
Boston College
miao.liu@bc.edu

Shuyang Wang
D'Amore-McKim School of Business
Northeastern University
s.wang@northeastern.edu

February 2022

We thank Jeff Cohen, Amy Hutton, Eddie Riedl, Phong Truong, Olga Usvyatsky, and participants of the Boston Empirical Accounting Conference for helpful suggestions and comments. We acknowledge financial support from Penn State University, Boston College, and Northeastern University. All errors are our own.

Human Information Production in the Machine Age: Evidence from Machine Adoption in the Asset Management Industry

Abstract

We study how machines change humans' role in information production activities of the asset management industry. Using the initiation of machine-based, systematic SEC filing downloading as a proxy of machine adoption, we find that machine adoption frees up the human workforce and allows humans to research a broader set of portfolio firms. After machine adoption, humans reallocate their information production activities towards portfolio firms on which they possess a comparative advantage over machines, including firms with more intangible assets, growth firms, and conglomerate firms. We also find that buy-side analysts affiliated with machine-adopting investment companies tend to participate more and be more inquisitive during conference calls held by such portfolio firms. Our findings suggest that the unique value of the human workforce is amplified once machines have relieved humans from tedious and mechanical activities.

Keywords: machine; asset management; information acquisition; conference call; SEC filings

JEL Classification: D83, G11, G14, G23

As early as 1965, Barron's magazine spoke of the "immeasurable" rewards computers could render investors, and how the machines were capable of relieving an analyst of "dreary labor, freeing him for more creative activity."

— Gregory Zuckerman (2019), "The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution"

1. Introduction

An essential function of the financial industry is to facilitate resource allocation by producing information (Levine, 2005). This function has been profoundly transformed by modern information technologies. Little is known, however, about the impact of information technologies on the role of humans in information production in the financial industry. Recent research has shown the machine's superior ability to process information faster and more accurately in portfolio management (Gu, Kelly, and Xiu, 2020), banking (Fuster et al., 2019; Liu, 2022), and investment recommendations (Coleman, Merkley, and Pacelli, 2021), which seems to attest to the machine's potential to replace humans in the financial industry.[1] Other research, however, underscores humans' unique ability to produce soft information by interacting with borrowers in banking (Costello, Down, and Mehta, 2021) and managers in conference calls (Matsumoto, Pronk, and Roelofsen, 2011). Indeed, financial firms that base their decisions on machines remain a small segment of the market today in asset management, banking, and equity research.[2] In this study, we

[1] See, for example, https://www.wsj.com/articles/the-quants-run-wall-street-now-1495389108.

[2] In asset management, although machine-decision-driven quantitative funds grow faster than human-judgment-driven discretionary funds, the total market share of the quantitative funds remains moderate today at less than 10% (Abis, 2021). In the mortgage market, Fintech lenders represent less than 10% of the market in the sample of Fuster et al. (2019) and Buchak et al. (2018). In equity research, Robo-analysts contribute to 14% of all outstanding analyst recommendations in the sample of Coleman et al. (2021).

investigate how machines reshape what humans do in information production in the asset management industry.

A challenge to studying the machines' impact on investment companies is that machine-facilitated activities at these companies are typically proprietary and thus unobservable to the public. We tackle this challenge by using the EDGAR Log File data to infer the time when an investment company started downloading SEC filings systematically using automated algorithms ("machines"). Specifically, we identify information acquisition by a machine when the acquiring investment company consistently downloads a large volume of EDGAR filings beyond human comprehension within a short period of time. If an investment company uses machines to download financial filings, it would likely use machines to automate other data tasks, such as collecting and cleaning data and performing data analysis.[3]

Specifically, we decode the partially obfuscated IP addresses in the Log File following Chen et al. (2020) and identify IP addresses that belong to investment companies covered by the Thomson Reuters Institutional Holdings (s34) database. We then classify IP addresses that accessed SEC filings in quick successions as machines. Among investment companies associated with machine-generated downloads at some point during the sample period from 2003 to 2017, we further require the investment company to continue to use machine downloads each year after initiating such downloading behavior. Based on this procedure, we classify 122 out of the 2,022 investment companies as machine adopters. We treat the month when an investment company starts to have machine-based SEC filing downloads every year as its inception of machine-based information processing, i.e., machine adoption. Using machine adoption as a treatment, we

[3] We confirm this reasoning by interviewing a group of practitioners in the asset management industry.

conduct a difference-in-differences analysis to study how machine adoption changes the role of humans in information production at the adopting investment company.

To set the stage, we start by testing if machine adoption frees up human capital and allows humans to research a broader set of firms. Prior research has found that the limited information processing capacity of humans is a profound driving force of various capital market outcomes, such as how fast new information transmits into stock prices and how firm managers time their disclosures (for a review of the literature, see Blankespoor, deHaan, and Marinovic, 2020). Directly related to our setting, Van Nieuwerburgh and Veldkamp (2010) model the joint determination of investment and information choices and show that, due to limited information processing capacity, investment companies rationally under-diversify and only acquire information on a constrained set of firms. Our first result shows that human EDGAR downloading activities increase by 19.4% after the machine adoption by an investment company. In addition, we find that humans research twice as many firms in their holding portfolio after their investment companies adopt machines.

Next, we examine whether humans reallocate their information production capacity towards portfolio firms on which they have a comparative advantage over machines. Firms with high intangible capital or market-to-book ratios have more value embedded in their R&D, patents, and growth opportunities (Govindarajan, Rajgopal, and Srivastava, 2018; Green, Louis, and Sani, 2021; Iqbal et al., 2021). Processing financial information for these companies would require nonstandard valuation models and industry-specific knowledge and expertise that machine algorithms do not possess. We thus predict that humans shift their attention to produce more information on such firms after machine adoption. Prior research also suggests that humans might have an advantage in processing complex firms operating in multiple industries because any information about such firms, even if machine-readable, needs to be put into context to infer its implication for the entire conglomerate (Cohen and Lou, 2012). We thus predict that humans shift their attention to these firms after machine adoption. Consistent with our prediction, we find that the number of human-generated downloading activities related to high-intangible firms or conglomerate firms is significantly larger than those related to other firms after machine adoption.

In addition to the acquisition of regulatory filings, another vital channel of information production where humans have an advantage is their interactions with firm managers during conference calls (e.g., Bowen, Davis, and Matsumoto, 2002; Matsumoto et al., 2011; Mayew, Sharp, and Venkatachalam, 2013). We manually identify buy-side analysts from investment companies who participate in earnings conference calls from 2007 to 2017 and examine if machine adoption by their investment companies affects how buy-side analysts communicate with the portfolio firms' management. We find that investment companies are significantly more likely to participate in calls held by high-intangible firms, high-growth firms, and conglomerate firms. In addition, conditional on participation, buy-side analysts ask more questions during the conference calls held by conglomerate firms.

If humans benefit from machine adoption by reallocating effort toward information production tasks on which they possess an advantage, investment companies' portfolio allocation decisions should be more responsive to human information production after machine adoption. We examine this prediction in our final tests. The results suggest that investment companies are more likely to trade based on human information acquisition after machine adoption. In addition, we find that investment companies hold more firms in their portfolio after machine adoption, consistent with the expansion in the scope of human research manifests in the scope of portfolio firms held by the investment companies.

We address two key endogeneity concerns. First, machine adoption is not a random decision. Some omitted factors at the investment company level might explain machine adoption and subsequent changes in human information acquisition behavior. For example, an investment company undergoing a shift in strategy might adopt machines and change how humans research investment opportunities. In this case, the strategy shift is the omitted factor driving both our treatment and outcome variables. We address this concern in two steps.

First, we conduct a parallel-trend analysis and find no difference in pre-trend for treated vs. control investment companies regarding their human information production behavior. Second, we supplement our OLS regressions with an instrumental variable approach. We instrument the focal investment company's machine adoption decision using the (log) number of machine-based downloads by peer investment companies in the same geographic area. The intuition for the instrument is that machine adoption requires the existence of a labor force that possesses machine-related skills, and the labor market is often geographically segmented due to the relative immobility of labor as a production factor. To the extent that different investment companies in the same area may rely on the same pool of skilled labor to empower their machine-related technical capabilities, local peer companies' pervasive use of machines should be related to the focal company's propensity of machine adoption. We find strong support for this intuition with the F-Stat in the first stage of the 2SLS regression above 50. Further, we provide evidence that this instrument is likely to satisfy the exclusion restriction using a falsification test. If the instrument affects the focal company's human information acquisition only through the focal company's own machine adoption, we expect no relation between the instrument and human information acquisition from non-adopting local investment companies. This is indeed what we find.

The second endogeneity concern is that investment companies' portfolio selection is an endogenous decision. Machine-adopting investment companies might invest more in firms with a richer business environment that warrants more information inputs in the valuation analysis. If so, it could be firms' business fundamentals, instead of the machine adoption, that induce changes in human information acquisition at machine-adopting investment companies. We address this concern by leveraging the granularity of our data and including firm-by-year fixed effects in our model. We purge any time-varying firm characteristics and news that induce information acquisition from both treated and control investment companies with firm-by-year fixed effects. As a result, our identification comes from comparing, for the same firm-year, the difference in human information production between a treated vs. a control investment company.

Our findings shed light on the impact of machine adoption on human information production in the asset management industry. Informational choices are made in several stages of the investment process. Investment professionals and their data analytics colleagues start by deciding on a pool of investee firms to research, followed by the actual process of data collection, data cleaning, data processing, and decisions on potentially additional information to acquire. We find that when machines automate data collection, cleaning, and some aspects of data processing, they do not replace humans entirely. Instead, machines relieve humans from such tasks and manifest humans' strengths in identifying emerging investment opportunities and the pool of potential investees firms, choosing an investment strategy, and processing unconventional or soft financial information.

We contribute to three strands of literature. First, our paper provides new insights into the nascent literature examining how technology advancements reshape the financial industry. Prior studies mainly focus on comparing the performance of Fintech with traditional financial services in banking (Fuster et al., 2019; Tang, 2019; Vallee and Zeng, 2019), financial advising (D'Acunto, Prabhala, and Rossi, 2019), and sell-side investment research (Coleman et al., 2021). Our study is the first to examine how the role of humans changes after financial firms adopt modern information technologies, focusing on information production in the asset management industry.

By separating information-related tasks accomplished by automated algorithms from tasks completed by humans, our study sheds light on the comparative advantages of humans and machines. There are two dimensions on which machines and humans may differ in terms of their relative strengths in information acquisition. First, humans are subject to information processing constraints and other cognitive limitations (Blankespoor et al., 2020), while machines do not face cognitive constraints. As a result, machines are in a better position to process large-scale investment-related information. Second, information could be hard and soft (Liberti and Petersen, 2018).[4] Machines are arguably better able to process hard information, while humans possess the communication and social skills to generate and process soft information. We show that machine adoption, while improving investment companies' hard information processing efficiency, also encourages soft information production by reallocating human efforts to tasks where they have an advantage (e.g., interacting with firm managers).

[4] Most research on hard and soft information focuses on the banking industry, in which hard and soft information are often substitutes (Liberti and Petersen, 2018). As banks grow in size, they tend to substitute hard for soft information because it is difficult to transmit unverifiable soft information across hierarchies (Stein, 2002). As a result, the industrial organization of the banking industry bifurcates endogenously: large banks serve large borrowers with high-quality hard information (e.g., audited financials), and small banks serve opaque or small borrowers using their access to soft information (e.g., personal communications).

In a related working paper, Grennan and Michaely (2020) study how the emergence of AI algorithm-based stock recommendation affects the role of human equity analysts. Our paper differs in two aspects. First, while Grennan and Michaely focus on AI, a predictive technology that a small group of Fintech companies has only recently applied, we examine the impact of machine-based data collection and processing, a technology adopted by traditional investment companies since the early 2000s. Second, Grennan and Michaely study how AI-based stock recommendation directly competes with human analysts at traditional brokerage houses. In contrast, we identify a more subtle interaction between machines and humans: machine adoption may liberate humans, helping them prioritize tasks on which they excel.

Second, our study is related to the labor economics literature that examines the effects of automation. Prior literature suggests the potential for machines to replace humans in many routine tasks (e.g., Autor, 2015; Chui, Manyik, and Miremadi, 2016; Acemoglu and Restrepo, 2019, 2020). Our study focuses on information production, a critical activity in many professions. Information production is a complex process involving a series of tasks. Empirical evidence on individual tasks is currently scarce in the labor literature due to the limited availability of archival data on the task-level productive activities of the corporate labor force. By exploiting a granular dataset containing the digital footprints of investment company employees, our study provides large-sample empirical evidence on the substitution and reinforcement roles of machines. Our findings suggest that, when machines automate some tasks, such as data collection and machine-readable data processing, they allow humans to research a broader set of firms and reallocate effort to tasks where human strengths locate, such as interacting with firm managers and researching firms whose value cannot be easily inferred from machine-readable data.

Finally, our findings add to the empirical evidence on how institutional investors acquire and process financial information. An emerging literature based on internet traffic data (especially the EDGAR Log data as used in our study) has provided important evidence on how institutional investors' information choices support their trading behavior. These studies examine the information choices of institutional investors from a "rational" perspective but without considering the human factor. For example, studies show that institutional investors benefit from acquiring information about public companies (Crane, Crotty, and Umar, 2020; Dyer, 2021), insider trades (Chen et al., 2020), and peer investment companies' portfolio disclosures (Cao et al., 2021). A common feature of the empirical design of these studies is that institutional investors are treated as an atomic entity. In contrast, we provide a more granular picture of institutional investors' information choices by separating the role of machines and humans.

As is the case with other studies that attempt to decode the "black box" of financial institutions' operations from public data sources, our study faces a critical caveat. We infer machine adoption from investment companies' machine-based systematic EDGAR filing downloading activities. Thus, we may fail to capture the machine adoption of an investment company if it accesses regulatory filings through other venues or uses machines for purposes other than downloading SEC filings. Such misclassification would inadvertently place some machine-adopting companies into our control group and potentially reduce our ability to find a treatment effect. In Section 3, we also discuss two alternative methods of identifying machines and why we opt to use our current approach.

The remainder of the paper is organized as follows. Section 2 develops hypotheses. Section 3 describes the data and reports descriptive statistics. Section 4 presents the empirical results. Section 5 presents additional analysis. Section 6 concludes.

2. Hypothesis Development

It has long been recognized that limited information processing capacity is a crucial constraint behind human decisions (Simon, 1955). Sims (2003) formalizes this notion into a bounded rationality framework that features two central predictions. The first prediction is that humans tend to underreact to new information due to limited information processing capacity. This prediction successfully explains a wide range of market outcomes, including investors' trading (Blankespoor, 2019), consumer and manager's use of tax rates (Chetty, Looney, and Kroft, 2009; Graham et al., 2017), stock price responsiveness to disclosure (Hirshleifer, Lim, and Teoh, 2009, 2011; DellaVigna and Pollet, 2009; Lawrence et al., 2018), and firms' disclosure choices in response to investors' information processing costs (deHaan, Shevlin, and Thornock, 2015; Blankespoor, 2019; Abramova, Core, and Sutherland, 2020).

In our setting, investment professionals face similar cognitive constraints and are unable to sift through seas of financial information effectively. With the aid of modern information technologies, such as using machines to perform systematic data collection and processing, investment professionals are liberated from these laborious tasks, reducing their cost of acquiring information on average. As a result, human professionals may use the freed-up capacity to cover a broader spectrum of firms in the search for trading opportunities. Our first hypothesis is summarized as follows.

H1: Machine adoption liberates humans from routine tasks and allows humans to research a broader set of firms.

The second main prediction of the bounded rationality framework is that humans rationally allocate attention within their processing capacity by paying more attention to information sources that reduce uncertainty. We expect that, once machines relax humans' information processing constraints, humans allocate the additional processing capacity to produce information for firms on which they have a comparative advantage over machines. Prior research has identified characteristics of firms that fall into this category, such as firms that have high growth opportunities coming from intangible assets and complicated firms that operate in multiple industries because of costly information processing (Chan, Lakonishok, and Sougiannis, 2001; Cohen and Lou, 2012; Cohen, Diether, and Malloy, 2013). Since assessing the potential for investment of such firms requires more nonstandard valuation models and more industry-specific expertise, we expect humans to have a comparative advantage over machines.

H2a: Humans at investment companies allocate more attention to studying firms with high intangible assets, firms with high growth rates, and conglomerate firms after machine adoption.

An important channel through which humans develop an advantage over machines in information production is through private communications with firm managers in conference calls (e.g., Bowen et al., 2002; Matsumoto et al., 2011; Mayew et al., 2013). We, therefore, expect humans to participate more and engage more with managers in conference calls held by firms with high intangible assets, firms with high growth rates, and conglomerate firms after machine adoption.

H2b: Humans at investment companies participate more in conference calls held by firms with high intangible assets, firms with high growth rates, and conglomerate firms after machine adoption. In addition, humans tend to be more inquisitive during such conference calls.

Our final hypothesis is about the impact of machine adoption on investment companies' portfolio allocation decisions. Investment companies produce information to facilitate portfolio allocation. Investment companies should place greater weight on more precise information to guide their trading decisions (Van Nieuwerburgh and Veldkamp, 2010). To the extent that machine adoption enables humans to reallocate effort toward information production tasks on which they possess an advantage, we expect these information production activities to generate more valuable investment research. As a result, the trading decisions should rely more heavily upon human-generated information. Therefore, as the first part of this hypothesis, we conjecture that investment companies' trading decisions are more likely to respond to human information acquisition after machine adoption.

H3a: Investment companies' trading decisions are more likely to respond to human information acquisition after machine adoption.

Moreover, if humans can cover a wider range of potential investee firms in their research activities with the aid of machines, we expect the expansion in the scope of human research to manifest in the scope of portfolio firms held by the investment companies. The second part of this hypothesis states as follows:

H3b: Investment companies hold a larger number of portfolio firms after machine adoption.

3. Data and Descriptive Statistics

3.1. Data and sample

We use four types of data in this study: (i) the EDGAR Log File data, which contain information on downloads (referred to interchangeably as "viewing activities") of SEC filings; (ii) CapitalIQ Conference Call data, which include information on whether an investment company participates in a firm's conference call, how many questions are asked by buy-side analysts at the investment company during the call, and the number of words for each question; (iii) Thomson Reuters Institutional Holdings (s34) data, which provide information on investment company characteristics and portfolio holdings; and (iv) the Compustat Fundamental annual tape, which contains firm characteristics.

The EDGAR Log File data contain the downloads of firms' filings from January 1, 2003 through June 30, 2017. This data set has been extensively used in prior research (e.g., Lee, Ma, and Wang, 2015; Cao et al., 2021; Crane, Crotty, and Umar, 2021). The fourth octet of the IP address in the Log File data is obfuscated by a three-character string. This data set has been extensively used in prior research. Following Chen et al. (2020), we decipher the fourth octect and infer the organization to which the IP address is registered, based on the Whois database of the American Registry for Internet Numbers (ARIN). Next, we match the organizations associated with the IP addresses to investment companies covered by the Thomson Reuters s34 database. Information on organizational IP addresses comes from the Whois database of the American Registry for Internet Numbers (ARIN). We provide additional details about the IP matching process in Appendix B.

To mitigate selection biases, omissions, or errors, we exclude investment companies that did not download any filings during our sample period. We further require that an investment company downloads the firm's filings at least once during the entire sample period for the investment company-firm pair to be included in the sample.

Our earnings conference call data comes from the CapitalIQ conference call transcripts database. We require the sample to be after 2007 when the data are more populated, retain firm-years when the firms hold conference calls, and require the investment company participated in at least one conference call in the sample period to be included in the analysis. To identify conference call participants from investment companies, we undertake a three-step process. First, we exclude all participants from sell-side brokerage firms or investment banks based on data from I/B/E/S. In the second step, we match the remaining participants to the names of institutional investors in the Thomson Reuters database of 13F filings. In the third step, for all unmatched conference call participants in the CapitalIQ database, we manually correct misspellings and conduct Internet searches when necessary to match with 13F investment companies. Through this extensive procedure, we identify 1,394 unique 13F investment companies participating in 147,376 conference calls held by 6,318 unique companies in our sample.

We classify downloading requests as machine-generated if they are associated with self-identified web crawlers. In particular, for each day, we classify IP addresses that accessed more than 50 unique firms' filings as machine-based algorithms, a criterion also used by Lee et al. (2015). Based on this criterion, 531 out of the 2,022 investment companies had machine-generated downloads at some point during the sample period. For an investment company to be classified as a machine user, we further require it to continue to use machine downloads each year after initiating such downloading behavior. This requirement reduces the number of machine users to 122 investment companies, constituting the final sample of adopting companies used in our analysis. In addition, we identify each investment company's starting date to have machine-generated downloads of SEC filings. As discussed in prior studies (e.g., Chen et al., 2020; Cao et al., 2021), investment companies may access filings via channels other than EDGAR, such as the filers' websites or through a data vendor (e.g., Chen et al., 2020; Cao et al., 2021). Thus, the number of downloads from EDGAR likely understates the actual number of cases in which investment companies access filings.

There are conceivably two other methods for identifying investment companies' exposure to information technologies. First, researchers may infer such exposure from how intensive a company hires labor with such expertise (e.g., Guo and Shi, 2020; Abis and Veldkamp, 2021). This approach may identify affected areas of operations unrelated to information acquisition, such as model construction and trading decisions, which may also involve machine-related expertise (Lopez de Prado, 2020). Therefore, this approach would be overly inclusive for our purpose as we only focus on information production activities. Second, investment companies may disclose in their regulatory filings whether they utilize machine learning and natural language processing (NLP) as part of their investment strategies. In Appendix C, we compile a data set of machine adoption based on a keyword search in investment companies' SEC filings. Comparing this data with the Log File-based data we use for our analysis, we find that disclosures by investment companies reveal machine adoption in a timely fashion. By relying on machine-like patterns of the information acquisition activities, our method can better pinpoint when investment companies begin to use machines to execute informational tasks.

3.2. Descriptive statistics and determinants of machine adoption

We present descriptive statistics in Table 1. Panel A reports the statistics at the investment company-quarter level. Our sample contains 17,526 investment company-years, 4.5% of which are in the post period of machine adoption. The median dollar value of the portfolio is $535 million (natural log = 20.098). The portfolio has a weighted average return of 17.0% and volatility of 0.037. A median investment company holds 103 portfolio firms (natural log = 4.635).

We present the characteristics of 39,173 firm-years in panel B. A median firm has total assets of $555 million (natural log = 6.319, in millions). The firm, on average, has two segments and has intangible assets equal to 15.7% of total assets. We calculate the book-to-market ratio of all firms in Compustat and classify the bottom three deciles as high-growth firms. 30.4% of firms in our sample are classified as high-growth firms.

In panel C, we present the summary statistics of 10,386,045 investment company-firm-year observations.[5] 20.2% of these investment company-firm-years contain human-generated viewing activities. The median distance between the headquarter of a firm and the downloading IP of an investment company is 871 miles (natural log = 6.771). The average relationship length between a firm and an investment company is 0.23 years, suggesting the relatively short-term nature of the investment decisions.

Panel D reports the conference call characteristics of 2,093,253 investment company-firm-years. On average, an investment company participates in 3.6% or 0.09 times of the firms'

[5] Because each firm can potentially be viewed by every investment company in any year, this comprehensive sample includes all potential viewing activities between any investment company and firm in any year. The construction of sample is comprehensive but could inflate the number of non-viewing activities (i.e., 0s in *#HumanView* or *HumanView*) because each investment company may not be interested in viewing all firms' filings. Nevertheless, it should dilute the effect that we could find.

conference calls each year. The average number of questions asked is 0.05.[6] The questions have a median of 55 words (natural log = 4.013).

In Table 2, we present the results of the regressions that test the determinants of machine adoption. We use a linear probability model in column (1) and find that investment companies that manage portfolios with higher dollar value (*PortSize* = 0.012, *t* = 3.59), more portfolio firms (*#PortFirm* = 0.045, *t* = 6.75), more concentrated holding (*PortHHI* = 0.256, *t* = 5.82), lower portfolio returns (*PortRet* = -0.082, *t* = -4.08), or higher portfolio return volatility (*PortVol* = 0.502, *t* = 4.75) are more likely to adopt machines. We use a logistic regression model in column (2) and find similar results.

4. Empirical Results

4.1. Human information acquisition – Main effects

To test the effect of machine adoption on human information acquisition behaviors (H1), we conduct the following regressions:

$$\#HumanView_{ijt} = Post_{it} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (1)$$

$$\#HumanView_{ijt} = Post1_{it} + Post2_{it} + Post3_{it} + Post4_{it} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (2)$$

In both regressions, the dependent variable is the number of times investment company *I* has human-generated downloads of firm *j*'s filings during year *t*. For regression (1), the independent variable of interest is $Post_{it}$, which captures whether investment company *i* has started to use machines for filings retrieval on EDGAR in year *t*. For regression (2), we replace *Post* with *Post1*

---

[6] If an investment company does not participate in the firm's conference call, we assign the number of questions asked as zero.

through *Post4*, where *Post1*, *Post2*, and *Post3* refer to the first, second, and third years after the month when an investment company adopted machines, respectively; *Post4* refers to all years after *Post3*. We include control variables at the investment company-firm-year level that may affect the relationship between machine adoption and human information acquisition, including the distance between the headquarter of the firm to the location of the IP address of the investment company (*Distance*), the duration of the relationship between the investment company and the firm since the first holding date (*RelationDuration*), and the percentage of ownership of the firm by the investment company (*Shares*). We also control for investment company-year level variables that may affect the machine adoption decisions, including portfolio size, holding concentration, annual return, return volatility, and the number of portfolio firms. To mitigate endogeneity concerns, we control for the firm-by-year fixed effects, which helps rule out any effects related to time-variant characteristics of the firms. We further control for investment company fixed effects to remove variations driven by the characteristics of the investment company.

The results are presented in Table 3. In panel A, we use the intensity of the human-generated downloads as a dependent variable. Columns (1) and (2) show the results without control variables. We find that *Post* is significantly positive in column (1) (0.189, t = 4.18), suggesting that investment companies have significantly more human-generated downloads after machine adoption. This finding is consistent with H1 that machine adoption relaxes human information processing constraints. Furthermore, we document that such effect persists beyond the first year of machine adoption: *Post1* through *Post4* are all positive and significant (*Post1* = 0.119, t = 3.14; *Post2* = 0.184, t = 4.19; *Post3* = 0.190, t = 3.79; *Post4* = 0.251, t = 4.24). Columns (3) and (4) present the results with control variables. We find similar results that after machine adoption, investment companies initiate more human-generated information retrievals (column (3): *Post* = 0.177, *t* = 4.01; column (4): *Post1* = 0.109, *t* = 2.91; *Post2* = 0.174, *t* = 4.01; *Post3* = 0.178, *t* = 3.459; *Post4* = 0.236, *t* = 4.10). The economic magnitude of the effects is also large: human EDGAR downloading activities increase by 19.4% after machine adoption of an investment company.[7]

In panel B of Table 3, we replace the dependent variable with an indicator variable of whether the investment company had human-generated downloads of the firm's filings during the year (*HumanView*). We find similar results to those in panel A. *Post* is significantly positive in column (1) (0.101, *t* = 5.13) and column (3) (0.096, *t* = 4.94). *Post1* through *Post4* are also positive and significant in column (2) (*Post1* = 0.078, *t* = 4.29; *Post2* = 0.107, *t* = 5.16; *Post3* = 0.117, *t* = 5.17; *Post4* = 0.107, *t* = 4.63) and column (4) (*Post1* = 0.074, *t* = 4.12; *Post2* = 0.103, *t* = 4.98; *Post3* = 0.113, *t* = 4.95; *Post4* = 0.100, *t* = 4.42). These results indicate that the effects identified in panel A are unlikely to be driven by investment companies with large numbers of human downloads.

Investment companies may elect to adopt machines for endogenous reasons. One might be concerned that adopters differ from non-adopters in some unobservable dimension which, in turn, explains their difference in human information acquisition in the post period. To alleviate this concern, we test for parallel trends between the treatment and control groups. Specifically, we run the following regression:

[7] For example, 0.177 in column (3) translates into a 19.4% (=exp(0.177)-1) change in the dependent variable.

$$HumanView_{ijt} = Pre4_{it} + Pre3_{it} + Pre2_{it} + Post1_{it} + Post2_{it} + Post3_{it} + Post4_{it} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (3)$$

We use *Pre1* (the year before machine adoption) as the baseline. *Pre2*, *Pre3*, and *Pre4* refer to the second, third, and fourth year or beyond before the machine adoption month, respectively. We plot the coefficients in Figure 2. The coefficients of *Pre2*, *Pre3*, and *Pre4* are not significantly different from zero, whereas *Post1*, *Post2*, *Post3*, and *Post4* are significantly positive, indicating that adopters and non-adopters have no discernible differences in human information acquisition activities up till the adoption event.

Investment companies acquire information on not only portfolio firms but also prospective investee firms. In panel C of Table 3, we examine whether machine adoption allows humans to research more on firms currently held by the investment company. We use a sample of 17,526 investment company-years and replace the dependent variable with the log number of portfolio firms (*HumanViewPort*) or the percentage of portfolio firms (*HumanViewPortPct*) viewed by the investment company by humans. We control for investment company characteristics, investment company fixed effects, and year fixed effects. We find that investment companies have more human-generated views of portfolio firms after machine adoption (*HumanViewPort*: *Post* = 0.840, *t* = 5.44; *HumanViewPortPct*: *Post* = 0.040, *t* = 3.72), suggesting a broader research coverage by humans. In economic magnitude, humans from an average investment company research 132% more firms in their holding portfolio after machine adoption, or 111% more in terms of the percentage of the portfolio firms.[8]

[8] The coefficient in column (1) translates into a 132% (=exp(0.840)-1) change in the number of portfolio firms whose filings are viewed. The coefficient in column (2) suggests that the percentage of portfolio firms whose filings are viewed more than doubled (0.040/0.036=111%) after machine adoption.

We further test whether machine adoption impacts human information acquisition of different SEC filing types, and find that information acquisition increases after machine adoption across all filing types. We include the results in the Supplementary Appendix.[9] Overall, our results support H1, which states that machine adoption relaxes human information processing constraints and allows humans to research and acquire information on a broader set of firms.

4.2. Human information acquisition – Cross-sectional tests

Next, we test the effect of machine adoption on human information acquisition of firms where humans should have a comparative advantage (H2a). We run the following group of regressions:

$$View_{ijt} = Post_{it} + Post_{it} \times X_{jt} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (4)$$

The dependent variable $View_{ijt}$ refers to how many times (*#HumanView*) or whether (*HumanView*) investment company *i* has human-generated views of firm *j*'s filings during year *t*. For each regression, we interact $Post_{it}$, which captures whether investment company *i* has started to use machines to retrieve filings in year *t*, with firm *j*'s characteristics in year *t*, X_{jt}, including the share of intangible assets (*Intangible*), whether the firm is a high-growth firm (*HiGrowth*), and the number of industry segments (*#Segment*). Detailed definitions are in Appendix A.

In columns (1) through (3) of Table 4, we find that the positive effect of machine adoption on the intensity of human information acquisition of firm *j* is more pronounced when the firm has

[9] Panel A of Table S.2 presents the results with different filing types viewed by investment companies by humans as dependent variables, including the scheduled filings, text-heavy filings, or trading-related filings. We find that after adopting machine, investment companies have significantly more views of scheduled filings (*Post* = 0.120, *t* = 3.59), text-heavy filings (*Post* = 0.154, *t* = 3.78), and trading-related filings (*Post* = 0.011, *t* = 3.55). Panel B of Table S.2 changes the dependent variables to be human viewing of 10-Ks or 8-Ks by investment companies. We find that investment companies have more human-generated views of 10-Ks and 8-Ks after machine adoption (10-Ks: *Post* = 0.083, *t* = 3.44; 8-Ks: *Post* = 0.058, *t* = 3.17).

more intangible assets (*Intangible* × *Post* = 0.249, *t* = 6.21) or when the firm has more segments (*#Segment* × *Post* = 0.069, *t* = 7.30). Similarly, in columns (4) through (6), we find that the positive impact of machine adoption on the likelihood of human information acquisition for firm *j* is more pronounced when the firm has more intangible assets (*Intangible* × *Post* = 0.066, *t* = 6.37) or when the firm has more segments (*#Segment* × *Post* = 0.021, *t* = 8.84). The results support H2a, which states that humans at investment companies allocate more attention to firms with high intangibles and conglomerate firms.

4.3. Conference call participation

To test the effect of machine adoption on the participation of conference calls held by firms on which humans should have a comparative advantage (H2b), we run the following group of regressions:

$$Participation_{ijt} = Post_{it} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (5)$$

$$Participation_{ijt} = Post_{it} + Post_{it} \times X_{jt} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt} \quad (6)$$

The dependent variable is one of the following four variables: 1) whether investment company *i* participates in firm *j*'s conference calls during year *t* (*Participate*); 2) the number of times investment company *i* participates in firm *j*'s conference calls during year *t* (*#Participate*); 3) the number of questions participants from investment company *i* ask during firm *j*'s conference calls in year *t* (*#Question*); and 4) the average number of words per question asked by participants from investment company *i* during firm *j*'s conference calls in year *t* (*QuestionLength*). The independent variable of interest is $Post_{it}$, which captures whether investment company *i* has started to use machines to retrieve filings in year *t*, and the interaction between *Post* and firm *j*'s characteristics in year t, X_{jt}, including the share of intangible assets (*Intangible*), whether the firm is a high-growth firm (*HiGrowth*), and the number of industry segments (*#Segment*).

The results are presented in Table 5. In panel A, investment companies do not have a different likelihood of participating in conference calls after machine adoption. However, cross-sectional results show that investment companies shift their focus to firms with higher intangible assets (*Intangible* × *Post* = 0.027, t = 2.07), growth firms (*HiGrowth* × *Post* = 0.015, t = 2.05), and firms with more segments (*#Segment* × *Post* = 0.009, t = 2.63). In panel B, we document similar results for the number of times that investment companies participate in firms' conference calls (*Intangible* × *Post* = 0.096, t = 2.25; *HiGrowth* × *Post* = 0.051, t = 2.19; *#Segment* × *Post* = 0.024, t = 2.44). In panel C, we find that participants ask more questions during conference calls held by firms with more business segments after machine adoption by their investment companies (*#Segment* × *Post* = 0.013, t = 2.59). In panel D, we find that participants generally raise longer questions after machine adoption by their investment companies (*Post* = 0.045, t = 3.52).

Overall, the results support H2b that humans at investment companies are more likely to participate and become more inquisitive in conference calls held by firms with high intangible assets, growth firms, and conglomerate firms after machine adoption.

4.4. Trading decisions

In this section, we test the effect of machine adoption on investment companies' trading behavior (H3a). We run the following regressions:

$$AbsChHolding_{i,j,q+1} \; or \; ChHolding_{i,j,q+1}$$
$$= Post_{it} + Post_{it} \times HumanView_{ijq} + Controls_{ijq} + \mu_{jt} + \delta_i + \varepsilon_{ijq} \quad (7)$$

The dependent variable, $AbsChHolding_{i,j,q+1}$ ($ChHolding_{i,j,q+1}$), is the absolute (signed) change in holding of firm *j* by investment company *i* during quarter $q+1$. The independent variable of interest is the interaction of $Post_{it}$ and whether investment company *i* has human-generated downloads of firm *j*'s filings during quarter *q* ($HumanView_{ijq}$). Control variables include the distance between the headquarter of the firm to the location of the IP address of the investment company, the duration of the relationship between the investment company and the firm, the ownership percentage of the firm by the investment company, portfolio size, holding concentration, return, return volatility, and the number of portfolio firms measured at quarter *q*. We also control for firm-by-year fixed effects and investment company fixed effects. Column (1) shows that investment companies change more of their holdings if they have human-generated downloads during the post-adoption period (*HumanView* × *Post* = 0.015, *t* = 2.88). Column (2) suggests that investment companies increase more of their holdings if they have human-generated downloads after machine adoption (*HumanView* × *Post* = 0.010, *t* = 2.51). Collectively, these results support H3a and suggest that the trading behaviors of investment companies are more responsive to human information acquisition behaviors after investment companies adopt machines.

4.5. The scope of portfolio allocation

Finally, we test the effect of machine adoption on investment companies' portfolios (H3b). Specifically, we are interested in the number of portfolio firms held by investment companies. We run the following regressions:

$$\#PortFirm_{i,q+1} \text{ } or \text{ } \#PortFirm_{i,q} = Post_{iq} + Controls_{iq} + \mu_q + \delta_i + \varepsilon_{iq} \quad (8)$$

The dependent variable, $\#PortFirm_{i,q+1}$ ($\#PortFirm_{i,q}$), is the number of portfolio firms held by investment company *i* during quarter $q+1$ (quarter *q*). The independent variable of interest is $Post_{iq}$, whether investment company i adopts machine during quarter q. Control variables include portfolio size, return, and return volatility measured at quarter q. Year-by-quarter fixed effects and investment company fixed effects are controlled. Column (1) shows that investment companies expand their portfolio by holding more firms during the post-machine adoption period (*Post* = 0.151, t = 2.65). Column (2) suggests that investment companies also increase the number of firms in the next quarter after the machine adoption (*Post* = 0.126, t = 2.20). Overall, the results support H3b that machine adoption increases the number of firms in investment companies' portfolios. It is worth noting that this result is different from the results on the number of firms analyzed by humans (Table 3, panel C). The earlier results suggest that humans expand the scope of their research activities (but these activities may not necessarily lead to investment decisions). In contrast, the results in this section suggest that investment companies indeed increase the broadness of their investment decisions.

5. Additional Analysis

5.1. Instrumental variable approach

The decision to adopt machines is arguably endogenous, and we may not have sufficiently controlled for the factors that drive such a decision. To address this concern and more convincingly establish connections between machine adoption and changes in human-based information acquisition activities, we employ an instrumental variable (IV) approach. Our IV is the natural logarithm of the number of machine-based downloads by peer investment companies in the same geographic area (zip code), denoted *#PeerMachineView*.

Machine adoption requires the existence of a labor force that possesses machine-related skills, and the labor market is often segmented geographic areas. To the extent that different investment companies in the same area may rely on the same pool of skilled labor to empower their machine-related technical capabilities, the pervasive use of machines by local peer companies should be related to a company's propensity of machine adoption (i.e., the relevance criterion of IV). Additionally, this IV is not only relevant to machine adoption but also arguably satisfies the exclusion restriction—the use of machines by local peer companies should only affect the information acquisition activities of a company's human labor force through the company's own adoption of machines.

Table 8 reports the results of the IV approach. In the first stage, we estimate a linear probability model of how various determinants, including *#PeerMachineView*, affect *Post*. This regression is conducted over a sample of investment-year observations. As reported in column (1), we find that the IV has a positive and significant effect (*#PeerMachineView*: 0.008, t = 7.99). In columns (2) and (3) of panel A, we regress *#HumanView* and *HumanView* on the IV over the sample of investment companies that have never adopted machine downloads. We do not find any significant association, further attesting to the exclusion restriction of the IV.

In the second stage, we use the instrumented value of *Post* (*Post_IV*) to examine the effects of machine adoption on human activities, using the sample of investment company-firm-years. As reported in columns (1) and (2), we find that the coefficient of *Post_IV* is positive and significant (*#HumanView*: 0.384, t = 1.94; *HumanView*: 0.241, t = 2.33). Overall, the results of the IV approach are consistent with the results from our main analysis.

5.2. Alternative definitions of machine downloads

Because the identification of machine-generated downloads is pivotal for our proxy of investment company's machine adoption, we also adopt two alternative definitions of machine downloads: (i) daily IP addresses that searched more than 1,000 filings or more than five filings per minute (Drake, Roulstone, and Thornock, 2015); and (ii) daily IP addresses that searched more than 500 filings, more than 25 filings per minute, or more than three unique registrants' filings per minute (Dechow, Lawrence, and Ryans, 2015; Ryans, 2017). We replicate our empirical analyses using these two definitions of machine downloads. Untabulated results suggest that our empirical findings are qualitatively the same using these alternative definitions of machine downloads.

6. Concluding Remarks

In this study, we examine the advancement of machines and their implications for humans. Our analyses focus on the asset management industry, where changes in human information production activities are observable upon machine adoption. We find that (i) machine adoption frees up the human workforce and allows humans to analyze a broader set of (prospective) portfolio firms; (ii) machine adoption reallocates human information production towards firms where humans may have a comparative advantage, including high-intangible firms, high-growth firms, and conglomerate firms; and (iii) buy-side analysts from investment companies that adopt machines tend to participate more and be more inquisitive during conference calls held by the firms on which humans have expertise.

Our findings underscore the notion that information production is a complex activity consisting of a series of tasks. As machines automate a portion of these tasks, such as data collection, cleaning, and processing, they liberate human information processing capacity. They allow humans to shift attention to tasks where humans possess an apparent advantage, including determining investment scope and strategy, studying opaque and complex firms, and communicating with firm managers.

Our study has implications for the role of machines in reshaping the landscape of investors. Even though conventional wisdom (from the atomic-entity perspective) holds that institutional investors are more sophisticated than retail investors, investment research teams at institutions are ultimately staffed by humans who face cognitive limitations. To the extent that machines help humans circumvent cognitive constraints, our findings suggest that the machine-adoption trend may exacerbate the skill gap between institutional and retail investors.

Future research may extend our study in several directions. First, our study focuses on the human-machine interactions in the asset management industry. Additional research can focus on other industries in which information plays a different role in decision making. Second, by leveraging publicly available data to identify machine adoption, we do not sufficiently probe into the causes and nature of such adoption. A deeper understanding of these issues would likely require scrutiny from inside the black box of investment companies' operations, such as survey evidence. Third, our research question is limited to machine adoption in the context of information acquisition. Machine adoption may take other forms, and in some cases, may be viewed as a prerequisite for AI/ML-aided process transformation. More research is called for to provide a more granular understanding of the implications of machines and algorithms for various economic decisions.

References

Abis, S. 2021. Man vs. machine: Quantitative and discretionary equity management. Working Paper.

Abis, S. and L. Veldkamp. 2021. The changing economics of knowledge production. Working Paper.

Abramova, I., J. E. Core, and A. Sutherland. 2020. Institutional investor attention and firm disclosure. *The Accounting Review* 95(6): 1–21.

Acemoglu, D., and P. Restrepo. 2019. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives* 33(2): 3–30.

Acemoglu, D., and P. Restrepo. 2020. Robots and jobs: Evidence from US labor markets. *Journal of Political Economy* 128(6): 2188–2244.

Autor, D. H. 2015. Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29(3): 3–30.

Bartram, S., Barnke, J., and M. Motahari. 2020. Artificial intelligence in asset management. CFA Institute Research Foundation Literature Review.

Blankespoor, E. 2019. The impact of information processing costs on firm disclosure choice: Evidence from the XBRL mandate. *Journal of Accounting Research* 57(4): 919–967.

Blankespoor, E., E. deHaan, and I. Marinovic. 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics* 70(2-3): 101344.

Bowen, R., A., Davis, and D. Matsumoto. 2002. Do conference calls affect analysts' forecasts? *The Accounting Review* 77(2): 285–316.

Buchak, G., Matvos G., Piskorski T., and Seru A. 2018. Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics* 130: 453–483.

Cao, S., K. Du, B. Yang, and L. Zhang. 2021. Copycat skills and disclosure costs: Evidence from peer companies' digital footprints. *Journal of Accounting Research* 59(4): 1261–1302.

Chan, L., J. Lakonishok, and T. Sougiannis. 2001. The stock market valuation of research and development expenditures. *Journal of Finance* 56(6): 2431–2456.

Chen, H., L. Cohen, U. Gurun, D. Lou, and C. Malloy. 2020. I.Q. from I.P.: Simplifying search in portfolio choice. *Journal of Financial Economics* 138(1): 118–137.

Chetty, R., A. Looney, and K. Kroft. 2009. Salience and taxation: Theory and evidence. *The American Economic Review* 99(4): 1145–1177.

Chui, M., Manyika, J. and Miremadi, M. 2016. Where machines could replace humans – and where they can't (yet). *McKinsey Quarterly*. Available at https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/where-machines-could-replace-humans-and-where-they-cant-yet.

Cohen, L., K. Diether, and C. Malloy. 2013. Misvaluing innovation. *Review of Financial Studies* 26(3): 635–666.

Cohen, L., and D. Lou. 2012. Complicated firms. *Journal of Financial Economics* 104(2): 383–400.

Coleman, B., K. J. Merkley, and J. Pacelli. 2021. Man versus machine: A comparison of robo-analyst and traditional research analyst investment recommendations. Working Paper.

Costello, A.M., Down, A.K. and Mehta, M.N. 2020. Machine+ man: A field experiment on the role of discretion in augmenting AI-based lending models. *Journal of Accounting and Economics* 70(2-3): 101360.

Crane, A., K. Crotty, and T. Umar. 2021. Hedge funds and public information acquisition. Working paper.

D'Acunto, F., N. Prabhala, and A. G. Rossi. 2019. The promises and pitfalls of robo-advising. *Review of Financial Studies* 32(5): 1983–2020.

Dechow, P. M., Lawrence, A., Ryans, J.P., 2015. SEC Comment Letters and Insider Sales. *The Accounting Review* 91, 401-439.

deHaan, E., T. Shevlin, and J. Thornock. 2015. Market (in)attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics* 60(1): 36–55.

DellaVigna, S., and J. M. Pollet. 2009. Investor inattention and Friday earnings announcements. *Journal of Finance* 64(2): 709–749.

Drake, M. S., D. T. Roulstone, and J. R. Thornock. 2015. The determinants and consequences of information acquisition via EDGAR. *Contemporary Accounting Research* 32(3): 1128–1161.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery. 2019. The role of technology in mortgage lending. *Review of Financial Studies* 32(5): 1854–1899.

Graham, J., M. Hanlon, T. Shevlin, and N. Shroff. 2017. Tax rates and corporate decision-making. *Review of Financial Studies* 30(9): 3128–3175.

Green, J., Louis, H. and Sani, J. 2021. Intangible investments, scaling, and the trend in the accrual-cash flow association. *Journal of Accounting Research*, forthcoming.

Grennan, J., and R. Michaely. 2020. Artificial intelligence and high-skilled work: Evidence from analysts. Working Paper.

Govindarajan, V., Rajgopal, S., and Srivastava, A. 2018. Why we need to update financial reporting for the digital era. *Harvard Business Review*. Available at https://hbr.org/2018/06/why-we-need-to-update-financial-reporting-for-the-digital-era.

Guo, X., and Z. Shi. 2020. The impact of A.I. talents on hedge fund performance. Working paper.

Gu, S., B. T. Kelly, and D. Xiu. 2020. Empirical Asset Pricing via Machine Learning. *Review of Financial Studies* 33(5): 2223–2273.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64(5): 2289–2325.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2011. Limited investor attention and stock market misreactions to accounting information. *Review of Asset Pricing Studies* 1(1): 35–73.

Iqbal, A., Rajgopal, S., Srivastava, A. and Zhao, R. 2021. Value of internally generated intangible capital. Working paper.

Lawrence, A., J. P. Ryans, E. Sun, and N. Laptev. 2018. Earnings announcement promotions: A Yahoo Finance field experiment. *Journal of Accounting and Economics* 66(2-3): 399–414.

Lee, C., P. Ma, and C. Wang. 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116(2): 410–431.

Levine, R. 2005. Finance and growth: Theory and evidence. In *Handbook of Economic Growth* (Eds), P. Aghion & S. Durlauf, Amsterdam: North-Holland Elsevier Publishers.

Liberti, J. M. and M. A. Petersen. 2018. Information: Hard and soft. *Review of Corporate Finance Studies* 8(1): 1–41.

Liu, M. 2022. Assessing human information processing in lending decisions: A machine learning approach. *Journal of Accounting Research*, forthcoming.

Lopez de Prado, M. 2020. *Machine Learning for Asset Managers*. Cambridge University Press.

Matsumoto, D., M. Pronk, and E. Roelofsen. 2011. What makes conference calls useful? The information content of managers' presentations and analysts' discussion sessions. *The Accounting Review* 86(4): 1383–1414.

Mayew, W., N. Sharp, and M. Venkatachalam. 2013. Using earnings conference calls to identify analysts with superior private information. *Review of Accounting Studies* 18(2): 386–413.

Ryans, J. 2017. Using the EDGAR log file data set. Working Paper, London Business School.

Simon, H. A. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69(1): 99–118.

Sims, C.A. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50(3): 665–690.
Stein, J. 2002. Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance* 57: 1891–1921.
Tang, H. 2019. Peer-to-peer lenders versus banks: Substitutes or complements?. *Review of Financial Studies* 32(5): 1900–1938.
Vallee, B., and Y. Zeng. 2019. Marketplace lending: A new banking paradigm?. *Review of Financial Studies* 32(5): 1939–1982.
Van Nieuwerburgh, S., and L. Veldkamp. 2010. Information acquisition and underdiversification. *Review of Economic Studies* 77(2): 779–805.
Zuckerman, G. 2019. *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*. Penguin Publishing Group.

Appendix A: Definitions of Variables

Variable	Definitions
Investment company level	
Post	An indicator variable which equals 1 if year *t* is after the month when an investment company is identified as starting to use machine to view filings on EDGAR, and 0 otherwise.
PortHHI	The Herfindahl index of the portfolio, defined based on the market value of each component stock in the portfolio of an investment company in year *t*.
PortRet	Portfolio return, defined as the average return of an investment company's portfolio in year *t*.
PortSize	Portfolio size, defined as the natural logarithm of the market value of an investment company's holdings at the end of year *t*.
PortVol	Portfolio return volatility, defined as the return volatility of an investment company's portfolio in year *t*.
#PortFirm	The natural logarithm of the number of portfolio firms of an investment company in year *t*.
Post1	An indicator variable which equals 1 if the year *t* is within one year after the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Post2	An indicator variable which equals 1 if the year *t* is between one year and two years after the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Post3	An indicator variable which equals 1 if the year *t* is between two years and three years after the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Post4	An indicator variable which equals 1 if the year *t* is more than three years after the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Pre1	An indicator variable which equals 1 if the year *t* is within one year before the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Pre2	An indicator variable which equals 1 if the year *t* is between one year and two years before the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Pre3	An indicator variable which equals 1 if the year *t* is between two years and three years before the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.
Pre4	An indicator variable which equals 1 if the year *t* is more than three years before the month when an investment company is identified as starting to use machines to view filings on EDGAR, and 0 otherwise.

HumanViewPort	The natural logarithm of the number of portfolio firms whose filings are viewed (downloaded) by an investment company during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B.
HumanViewPortPct	The percentage of portfolio firms whose filings are viewed (downloaded) by an investment company during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B.
#PeerMachineView	The natural logarithm of one plus the number of times the investment companies in the same zip code (excluding the focal investment company) view (i.e., download) the filings of the firm during year *t*, and the viewing activity is classified as machine-generated based on a procedure detailed in Appendix B.
Firm level	
Size	The natural logarithm of total assets (Compustat item *at*) in year *t*.
Intangible	Intangible assets, calculated using intangible assets (*intan*) scaled by total assets (*at*) in year *t*.
HiGrowth	An indicator variable which equals 1 if the firm has book-to-market ratio in the bottom three deciles, and 0 otherwise. Book-to-market ratio is defined as the difference between total assets (*at*) and long-term debt (*dltt*), scaled by market value of common stock (*prcc_f*×*csho*) in year *t*.
#Segment	The number of business or operating segments in year *t*.
Investment company-firm level	
#HumanView$_{ijt}$	The natural logarithm of one plus the number of times an investment company *i* views (i.e., downloads) the filings of the firm *j* during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B.
HumanView$_{ijt}$	An indicator which equals 1 if an investment company *i* views (i.e., downloads) the filings of the firm *j* during year *t*, and 0 otherwise. The viewing activity is classified as human-generated based on a procedure detailed in Appendix B.
Distance	The natural logarithm of one plus the distance between the headquarter of the firm to the location of the IP address of the investment company.
RelationDuration	The number of years of the relationship between the investment company and the firm since the first holding date.
Shares	Percentage of ownership of a firm by an investment company in year *t*.
Participate	An indicator variable which equals 1 if an investment company participates in the firm's conference call during year *t*, and 0 otherwise.
#Participate	The number of times an investment company participates in the firm's conference call during year *t*.
#Question	The natural logarithm of the number of questions an investment company asks in the firm's conference call during year *t*.

QuestionLength	The natural logarithm of the average length of the questions an investment company asks in the firm's conference call during year t.
$AbsChHolding_{ijq}$	The absolute value of the percentage change in holding of the firm j by an investment company i from quarter $q-1$ to quarter q.
$ChHolding_{ijq}$	The percentage change in holding of the firm j by an investment company i from quarter $q-1$ to quarter q.

Appendix B: EDGAR Viewing Activities by Investment Companies

We obtain records of the retrieval of filings from the EDGAR Log File data, which cover the period between January 1, 2003 and June 30, 2017.[10] Each record from the EDGAR Log File data contains the IP address of the requesting user with the fourth octet obfuscated. It also includes the timestamp of the request and the accession number of the filing requested. Investment companies may access firms' filings via channels other than EDGAR, such as the filers' websites or through a data vendor. Thus, the number of downloads from EDGAR likely understates the actual number of cases in which institutions access filings.

We exclude unsuccessful requests and requests that land on index pages. We merge the Log File data with EDGAR index files by accession number to gather information on the form type, filing date and time, and name of the filing entity.

We match the organizations associated with the IP addresses to investment companies covered by the Thomson Reuters 13F database. Information on organizational IP addresses comes from the Whois database of the American Registry for Internet Numbers (ARIN). We follow Chen et al. (2020) to decipher the fourth octet, an obfuscated IPv4 address from the EDGAR Log File. The matching results in a mapping file between IP addresses and *mgrno* (Thomson's identifier of investment managers).

We classify viewing activities as machine-generated if they are associated with self-identified web crawlers or with daily IP addresses that searched more than 50 unique firms' filings, a criterion also used by Lee et al. (2015). We classify the rest viewing activities as human-generated.

[10] Available at https://www.sec.gov/dera/data/edgar-log-file-data-set.html.

Appendix C: Identifying Machines from Investment Company Disclosures

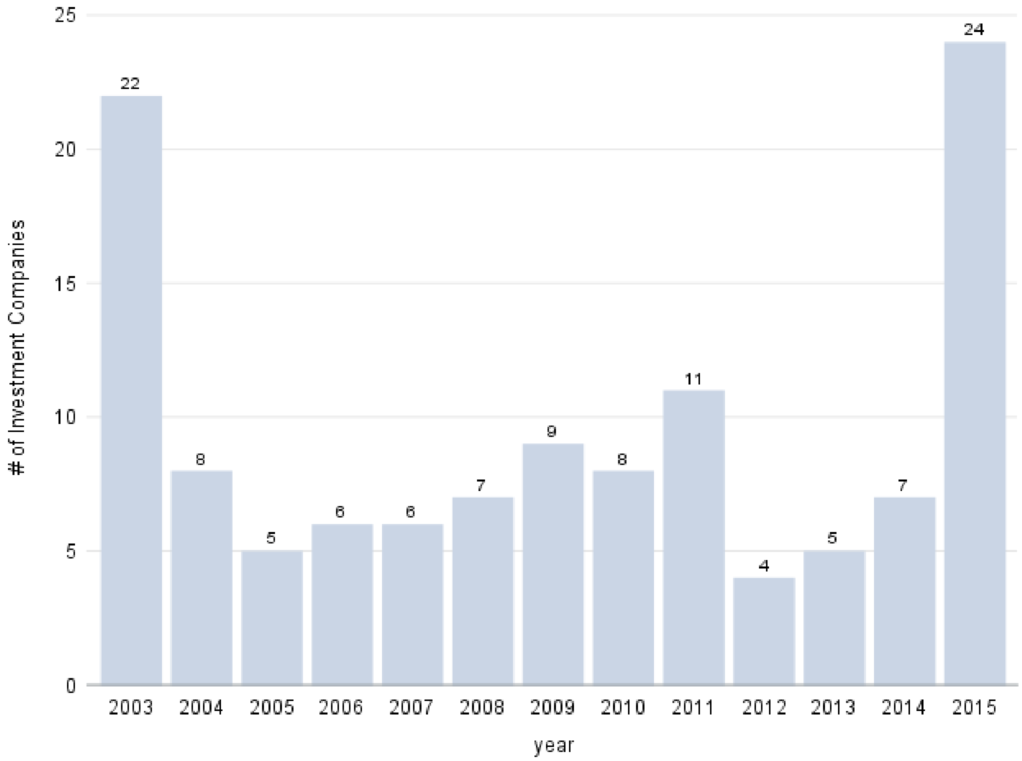
As discussed in Section 3.1, investment companies may disclose in regulatory filings whether they utilize machine learning and natural language processing (NLP) as part of their investment strategies. In this appendix, we describe a procedure we use to compile a data set of machine adoption based on a keyword search in investment companies' SEC filings, and show that this data does not serve our research question.

We start by searching the following keywords in all SEC filings by the investment companies in our sample: *machine learning*, *natural language processing*, and *NLP*. We then identify when an investment company starts disclosing at least one of these keywords, and label the date as disclosure-based machine adoption date.

Our textual analyses were able to compile 152 disclosure-based machine adoption dates of investment companies. Comparing this disclosure-based data with our Log File-based data, we find that most machine adopting investment companies in our sample do not show up in the disclosure-based data. Specifically, only 20 out of 122 machine adopters are covered by the disclosure-based data.

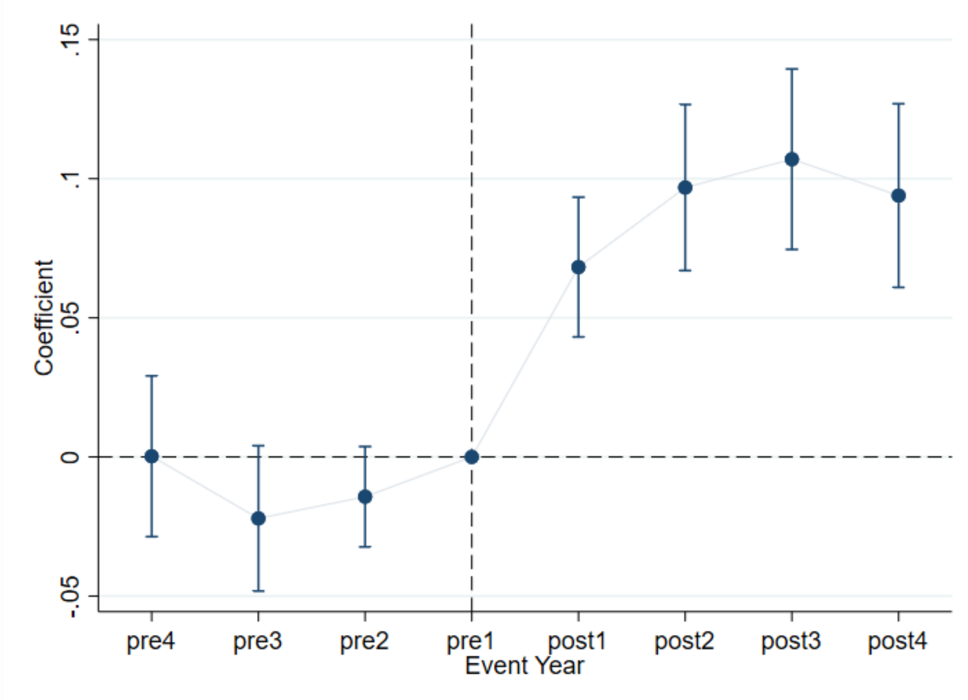
What is more, the years of adoption, according to investment company disclosures, are concentrated from 2016 to 2020. This is inconsistent with the conventional wisdom that information technologies had already been widely adopted by the asset management industry before this period (e.g., Bartram, Branke, and Motahari, 2020). Overall, we view this method as unlikely to serve our purpose.

Figure 1. Number of Investment Companies that Adopted Machine by Year



This figure presents the time-series trends of the number of investment companies adopting machines to download filings.

Figure 2. Parallel Trends of Human Downloads



This figure reports the parallel trends of human-generated downloads for our main regression. We run the following regression with baseline as *Pre1*: $HumanView_{ijt} = Pre4_{it} + Pre3_{it} + Pre2_{it} + Post1_{it} + Post2_{it} + Post3_{it} + Post4_{it} + Controls_{ijt} + \mu_{jt} + \delta_i + \varepsilon_{ijt}$. The coefficients of *Pre4*, *Pre3*, *Pre2*, *Post1*, *Post2*, *Post3*, and *Post4* are shown. The spikes represent 90% confidence intervals. All variables are defined in Appendix A.

Table 1. Descriptive Statistics

Panel A. Characteristics of investment company-years

	N	Mean	SD	P25	P50	P75
Post	17,526	0.045	0.208	0.000	0.000	0.000
PortSize	17,526	20.427	1.856	19.072	20.098	21.569
PortHHI	17,526	0.068	0.115	0.014	0.028	0.062
PortRet	17,526	0.170	0.199	0.074	0.156	0.261
PortVol	17,526	0.037	0.042	0.014	0.023	0.042
#PortFirm	17,526	4.715	1.446	3.829	4.635	5.576
HumanViewPort	17,526	0.766	1.423	0.000	0.000	1.099
HumanViewPortPct	17,526	0.036	0.088	0.000	0.000	0.015

Panel B. Characteristics of firm-years

	N	Mean	SD	P25	P50	P75
Size	39,173	6.391	2.104	4.880	6.319	7.795
Intangible	39,173	0.157	0.190	0.003	0.073	0.253
HiGrowth	39,173	0.304	0.460	0.000	0.000	1.000
#Segment	39,173	2.052	1.363	1.000	2.000	3.000

Panel C. Characteristics of investment company-firm-years

	N	Mean	SD	P25	P50	P75
#HumanView	10,386,045	0.327	0.779	0.000	0.000	0.000
HumanView	10,386,045	0.202	0.401	0.000	0.000	0.000
Distance	10,386,045	6.320	1.577	5.785	6.771	7.367
RelationDuration	10,386,045	0.227	0.685	0.000	0.000	0.000
Shares	10,386,045	0.000	0.001	0.000	0.000	0.000

Panel D. Characteristics of conference calls

	N	Mean	SD	P25	P50	P75
Participate	2,093,253	0.036	0.186	0.000	0.000	0.000
#Participate	2,093,253	0.091	0.520	0.000	0.000	0.000
#Question	2,093,253	0.050	0.272	0.000	0.000	0.000
QuestionLength	75,077	4.013	0.383	3.762	4.013	4.263

This table reports the descriptive statistics of our sample. All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A.

Table 2. Determinants of Machine Adoption

Dependent Variable:	*Post*	
	(1)	(2)
	LPM	Logit
PortSize	0.012***	0.211***
	(3.59)	(3.04)
PortHHI	0.256***	4.528***
	(5.82)	(6.16)
PortRet	-0.082***	-1.314**
	(-4.08)	(-2.26)
PortVol	0.502***	8.046***
	(4.75)	(3.00)
#PortFirm	0.045***	0.806***
	(6.75)	(7.48)
Year FE	Yes	Yes
N	17,526	17,526
Adjusted R^2	0.090	
Pseudo R^2		0.199

This table reports the determinants of machine adoption by investment companies. The sample consists of 17,526 investment company-years. Column (1) uses a linear probability model. Column (2) uses a logistic regression model. All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics or *z*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 3. The Effect of Machine Adoption on Human Information Acquisition

Panel A. Number of downloads

Dependent Variable:	*#HumanView*			
	(1)	(2)	(3)	(4)
Post	0.189***		0.177***	
	(4.18)		(4.01)	
Post1		0.119***		0.109***
		(3.14)		(2.91)
Post2		0.184***		0.174***
		(4.19)		(4.01)
Post3		0.190***		0.178***
		(3.79)		(3.59)
Post4		0.251***		0.236***
		(4.24)		(4.10)
Distance			-0.023***	-0.023***
			(-10.24)	(-10.24)
RelationDuration			0.071***	0.071***
			(5.88)	(5.90)
Shares			36.235***	36.481***
			(5.45)	(5.50)
PortSize			0.015**	0.014*
			(1.97)	(1.94)
PortHHI			-0.030	-0.043
			(-0.39)	(-0.56)
PortRet			0.009	0.005
			(0.32)	(0.16)
PortVol			0.124	0.097
			(0.93)	(0.72)
#PortFirm			0.001	0.000
			(0.10)	(0.03)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	10,386,045	10,386,045	10,386,045	10,386,045
Adjusted R^2	0.277	0.278	0.287	0.288

Panel B. Downloading indicator

Dependent Variable:	*HumanView*			
	(1)	(2)	(3)	(4)
Post	0.101***		0.096***	
	(5.13)		(4.94)	
Post1		0.078***		0.074***
		(4.29)		(4.12)
Post2		0.107***		0.103***
		(5.16)		(4.98)
Post3		0.117***		0.113***
		(5.17)		(4.95)
Post4		0.107***		0.100***
		(4.63)		(4.42)
Distance			-0.010***	-0.010***
			(-11.89)	(-11.89)
RelationDuration			0.029***	0.029***
			(7.83)	(7.86)
Shares			15.345***	15.350***
			(7.71)	(7.72)
PortSize			0.011***	0.012***
			(3.11)	(3.14)
PortHHI			-0.009	-0.011
			(-0.28)	(-0.36)
PortRet			0.004	0.006
			(0.26)	(0.37)
PortVol			0.066	0.052
			(0.93)	(0.73)
#PortFirm			-0.005	-0.005
			(-0.83)	(-0.90)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	10,386,045	10,386,045	10,386,045	10,386,045
Adjusted R^2	0.235	0.235	0.242	0.242

Panel C. Portfolio firms

Dependent Variable:	*HumanViewPort*	*HumanViewPortPct*
	(1)	(2)
Post	0.840***	0.040***
	(5.44)	(3.72)
PortSize	0.094***	0.007***
	(5.85)	(4.95)
PortHHI	-0.049	-0.008
	(-0.34)	(-0.42)
PortRet	0.023	-0.004
	(0.41)	(-0.48)
PortVol	0.102	0.049
	(0.41)	(1.31)
#PortFirm	0.199***	-0.013***
	(6.46)	(-5.37)
Mgrno FE	Yes	Yes
Year FE	Yes	Yes
N	17,526	17,526
Adjusted R^2	0.722	0.553

This table reports how machine adoption by investment companies affects human-generated downloading activities. In panel A and panel B, the sample consists of 10,386,045 investment company-firm-years. The dependent variable in panel A is the intensity of human-generated downloading activities (*#HumanView*). The dependent variable in panel B is an indicator variable that measures whether the investment company has human-generated downloads of the firms' filings (*HumanView*). In panel C, the sample contains 17,526 investment company-years. The dependent variables are the log number of portfolio firms covered by human download (*HumanViewPort*) and the percentage of portfolio firms covered by human download (*HumanViewPortPct*). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 4. The Effect of Machine Adoption on Human Information Acquisition – Cross-Sectional Analyses

Dependent Variable:	*#HumanView*			*HumanView*		
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.135***	0.175***	0.036	0.085***	0.094***	0.053***
	(3.10)	(3.92)	(0.81)	(4.41)	(4.79)	(2.80)
Intangible × *Post*	0.249***			0.066***		
	(6.21)			(6.37)		
HiGrowth × *Post*		0.009			0.007	
		(0.77)			(1.28)	
#Segment × *Post*			0.069***			0.021***
			(7.30)			(8.84)
Distance	-0.023***	-0.023***	-0.023***	-0.010***	-0.010***	-0.010***
	(-10.15)	(-10.25)	(-9.95)	(-11.82)	(-11.91)	(-11.65)
RelationDuration	0.071***	0.071***	0.068***	0.029***	0.029***	0.028***
	(5.89)	(5.88)	(5.94)	(7.84)	(7.83)	(7.84)
Shares	36.176***	36.231***	36.050***	15.329***	15.342***	15.289***
	(5.46)	(5.45)	(5.54)	(7.71)	(7.71)	(7.76)
PortSize	0.015**	0.015**	0.015**	0.011***	0.011***	0.012***
	(1.99)	(1.97)	(2.02)	(3.12)	(3.11)	(3.14)
PortHHI	-0.032	-0.030	-0.031	-0.009	-0.009	-0.009
	(-0.41)	(-0.39)	(-0.40)	(-0.29)	(-0.28)	(-0.29)
PortRet	0.009	0.009	0.009	0.004	0.004	0.004
	(0.33)	(0.32)	(0.33)	(0.26)	(0.26)	(0.26)
PortVol	0.123	0.124	0.125	0.065	0.066	0.066
	(0.92)	(0.93)	(0.94)	(0.93)	(0.93)	(0.94)
#PortFirm	0.001	0.001	0.001	-0.005	-0.005	-0.005
	(0.10)	(0.11)	(0.13)	(-0.84)	(-0.83)	(-0.82)
Gvkey × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes	Yes	Yes
N	10,386,045	10,386,045	10,386,045	10,386,045	10,386,045	10,386,045
Adjusted R^2	0.288	0.287	0.290	0.242	0.242	0.243

This table reports the cross-sectional regression results of how machine adoption and firm characteristics affect human-generated downloading activities by investment companies. The sample consists of 10,386,045 investment company-firm-years during 2003–2017. The dependent variable in columns (1)–(3) is the intensity of human-generated downloading activities (*#HumanView*). The dependent variable in columns (4)–(6) is an indicator variable that measures whether the investment company has human-generated downloads of the firms' filings (*HumanView*). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 5. The Effect of Machine Adoption on Conference Call Participation

Panel A. Participation indicator

Dependent Variable:	*Participate*			
	(1)	(2)	(3)	(4)
Post	0.000	-0.005	-0.005	-0.021**
	(0.02)	(-1.02)	(-1.03)	(-2.21)
Intangible × *Post*		0.027**		
		(2.07)		
HiGrowth × *Post*			0.015**	
			(2.05)	
#Segment × *Post*				0.009***
				(2.63)
Distance	-0.002***	-0.002***	-0.002***	-0.002***
	(-3.04)	(-3.04)	(-3.06)	(-3.03)
RelationDuration	0.006**	0.006**	0.006**	0.006**
	(2.48)	(2.48)	(2.46)	(2.36)
Shares	3.616**	3.612**	3.617**	3.627**
	(2.07)	(2.07)	(2.07)	(2.10)
PortSize	0.000	0.000	0.000	0.000
	(0.25)	(0.24)	(0.25)	(0.26)
PortHHI	-0.000	-0.000	-0.000	0.000
	(-0.01)	(-0.01)	(-0.01)	(0.02)
PortRet	-0.009*	-0.009*	-0.009*	-0.009*
	(-1.72)	(-1.71)	(-1.72)	(-1.72)
PortVol	0.021	0.021	0.021	0.021
	(1.05)	(1.03)	(1.05)	(1.05)
#PortFirm	0.004	0.004	0.004	0.004
	(1.56)	(1.57)	(1.56)	(1.58)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	2,093,253	2,093,253	2,093,253	2,093,253
Adjusted R^2	0.232	0.232	0.233	0.233

Panel B. Participation times

Dependent Variable:	*#Participate*			
	(1)	(2)	(3)	(4)
Post	-0.012	-0.031*	-0.028**	-0.071**
	(-1.06)	(-1.94)	(-1.98)	(-2.57)
Intangible × *Post*		0.096**		
		(2.25)		
HiGrowth × *Post*			0.051**	
			(2.19)	
#Segment × *Post*				0.025***
				(2.68)
Distance	-0.004***	-0.004***	-0.004***	-0.004***
	(-2.66)	(-2.66)	(-2.69)	(-2.65)
RelationDuration	0.017**	0.017**	0.017**	0.016**
	(2.49)	(2.49)	(2.47)	(2.38)
Shares	8.778*	8.764*	8.783*	8.809*
	(1.82)	(1.82)	(1.82)	(1.84)
PortSize	-0.001	-0.001	-0.001	-0.001
	(-0.15)	(-0.15)	(-0.15)	(-0.14)
PortHHI	-0.007	-0.007	-0.007	-0.006
	(-0.26)	(-0.26)	(-0.27)	(-0.23)
PortRet	-0.031**	-0.031**	-0.031**	-0.031**
	(-1.98)	(-1.98)	(-1.98)	(-1.98)
PortVol	0.097	0.096	0.097	0.097
	(1.40)	(1.38)	(1.40)	(1.40)
#PortFirm	0.010	0.010	0.010	0.010
	(1.46)	(1.47)	(1.46)	(1.48)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	2,093,253	2,093,253	2,093,253	2,093,253
Adjusted R^2	0.204	0.204	0.204	0.205

Panel C. Number of questions

Dependent Variable:	*#Question*			
	(1)	(2)	(3)	(4)
Post	0.006	0.003	0.003	-0.024**
	(1.20)	(0.64)	(0.71)	(-1.99)
Intangible × *Post*		0.013		
		(0.89)		
HiGrowth × *Post*			0.008	
			(0.90)	
#Segment × *Post*				0.013***
				(2.59)
Distance	-0.002***	-0.002***	-0.003***	-0.002***
	(-2.68)	(-2.67)	(-2.69)	(-2.66)
RelationDuration	0.008***	0.008***	0.008***	0.008**
	(2.64)	(2.63)	(2.62)	(2.52)
Shares	5.713**	5.711**	5.714**	5.729**
	(2.31)	(2.31)	(2.31)	(2.34)
PortSize	0.002	0.002	0.002	0.002
	(1.00)	(0.99)	(1.00)	(1.01)
PortHHI	0.001	0.001	0.001	0.002
	(0.12)	(0.12)	(0.11)	(0.15)
PortRet	-0.007	-0.007	-0.007	-0.007
	(-1.12)	(-1.11)	(-1.12)	(-1.12)
PortVol	0.012	0.012	0.012	0.012
	(0.50)	(0.49)	(0.50)	(0.50)
#PortFirm	0.004	0.004	0.004	0.004
	(1.27)	(1.28)	(1.27)	(1.30)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	2,093,253	2,093,253	2,093,253	2,093,253
Adjusted R^2	0.206	0.206	0.206	0.207

Panel D. Question length

Dependent Variable:	*QuestionLength*			
	(1)	(2)	(3)	(4)
Post	0.045***	0.048***	0.040***	0.047***
	(3.52)	(3.15)	(2.85)	(3.31)
Intangible × *Post*		-0.014		
		(-0.42)		
HiGrowth × *Post*			0.013	
			(1.18)	
#Segment × *Post*				-0.001
				(-0.28)
Distance	-0.003**	-0.003**	-0.003**	-0.003**
	(-2.21)	(-2.24)	(-2.25)	(-2.22)
RelationDuration	-0.008***	-0.008***	-0.008***	-0.008***
	(-2.91)	(-2.92)	(-2.90)	(-2.90)
Shares	1.859	1.859	1.870	1.864
	(1.45)	(1.45)	(1.45)	(1.45)
PortSize	0.010**	0.010**	0.010**	0.010**
	(2.40)	(2.41)	(2.40)	(2.40)
PortHHI	-0.107**	-0.107**	-0.107**	-0.107**
	(-2.45)	(-2.45)	(-2.44)	(-2.45)
PortRet	-0.024	-0.024	-0.024	-0.024
	(-0.44)	(-0.44)	(-0.43)	(-0.44)
PortVol	-0.046	-0.047	-0.047	-0.046
	(-0.29)	(-0.30)	(-0.30)	(-0.30)
#PortFirm	-0.022**	-0.023**	-0.022**	-0.022**
	(-2.49)	(-2.50)	(-2.49)	(-2.49)
Gvkey × Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	75,077	75,077	75,077	75,077
Adjusted R^2	0.281	0.281	0.281	0.281

This table reports how machine adoption by investment companies affects conference call participation. The sample in panels A, B, and C (panel D) consists of 2,093,253 (75,077) investment company-firm-years during 2007–2017. The dependent variables are whether the investment company participates in the firm's conference calls (*Participate*), the number of times the investment company participates in the firm's conference calls (*#Participate*), the number of questions participants from the investment company ask during the firm's conference calls (*#Question*), and the average number of words per question asked by participants from the investment company during the firm's conference calls (*QuestionLength*). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 6. The Effect of Machine Adoption on Trading Decisions

Dependent Variable:	$AbsChHolding_{i,j,q+1}$	$ChHolding_{i,j,q+1}$
	(1)	(2)
Post	-0.000	0.001
	(-0.04)	(0.52)
HumanView	0.004**	0.004***
	(2.29)	(3.16)
HumanView× Post	0.015***	0.010**
	(2.88)	(2.51)
Distance	-0.000	-0.000
	(-0.18)	(-0.43)
RelationDuration	0.059***	0.021***
	(15.63)	(7.42)
Shares	-3.985***	-8.044***
	(-3.39)	(-8.42)
PortSize	0.000	0.000
	(0.08)	(0.07)
PortHHI	0.019***	0.013***
	(4.62)	(3.82)
PortRet	0.002	0.001
	(0.88)	(0.67)
PortVol	0.046	0.025
	(1.46)	(1.06)
#PortFirm	0.007***	0.005***
	(5.74)	(5.08)
Gvkey × Year FE	Yes	Yes
Mgrno FE	Yes	Yes
N	78,646,299	78,646,299
Adjusted R^2	0.152	0.054

This table reports how machine adoption by investment companies affects the relationship between human information acquisition and subsequent trading behaviors. The sample consists of 78,646,299 investment company-firm-quarters during 2003–2017. The dependent variables is the absolute value of change in holding (*AbsChHolding*) in the next quarter in column (1), and the change in holding (*ChHolding)* in the next quarter in column (2). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 7. The Effect of Machine Adoption on the Scope of Portfolio Allocation

Dependent Variable:	*#PortFirm*$_{i,q}$	*#PortFirm*$_{i,q+1}$
	(1)	(2)
Post	0.151***	0.126**
	(2.65)	(2.20)
PortSize	0.285***	0.299***
	(18.17)	(19.56)
PortRet	0.506***	0.538***
	(9.53)	(10.33)
PortVol	-3.406***	-3.603***
	(-12.24)	(-12.40)
Mgrno FE	Yes	Yes
Year × Qtr FE	Yes	Yes
N	62,211	62,211
Adjusted R^2	0.932	0.932

This table reports how machine adoption affects the number of portfolio firms held by investment companies. The sample contains 62,211 investment company-quarters. The dependent variables are the log number of portfolio firms held by investment companies (*#PortFirm*) in the current or the next quarter. All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Table 8. The Effect of Machine Adoption on Human Information Acquisition (IV)

Panel A: First-stage regression and falsification test

Dependent Variable:	*Post*	*#HumanView*	*HumanView*
	(1)	(2)	(3)
#PeerMachineView	0.008***	0.005	0.003
	(7.99)	(1.19)	(1.34)
Distance		-0.033***	-0.013***
		(-5.88)	(-7.22)
RelationDuration		0.099***	0.033***
		(4.16)	(4.93)
Shares		38.528***	11.602***
		(2.94)	(3.67)
PortSize	0.008**	0.003	0.009
	(2.49)	(0.18)	(1.36)
PortHHI	0.247***	-0.130	-0.055
	(6.06)	(-0.79)	(-0.88)
PortRet	-0.079***	0.087	0.018
	(-4.07)	(1.00)	(0.50)
PortVol	0.322***	-0.215	0.095
	(3.40)	(-0.49)	(0.48)
#PortFirm	0.045***	0.000	-0.012
	(7.13)	(0.01)	(-0.89)
Year FE	Yes	No	No
Gvkey × Year FE	No	Yes	Yes
Mgrno FE	No	Yes	Yes
N	17,526	3,845,698	3,845,698
Adjusted R^2	0.140	0.397	0.335

Panel B: Second-stage regression

Dependent Variable:	*#HumanView*	*HumanView*
	(2)	(3)
Post_IV	0.384*	0.241**
	(1.94)	(2.33)
Distance	-0.023***	-0.010***
	(-10.21)	(-11.88)
RelationDuration	0.072***	0.029***
	(5.85)	(7.83)
Shares	37.031***	15.629***
	(5.59)	(8.15)
PortSize	0.011	0.010**
	(1.46)	(2.36)
PortHHI	-0.139	-0.075*
	(-1.51)	(-1.78)
PortRet	0.036	0.021
	(1.07)	(1.14)
PortVol	0.027	0.003
	(0.18)	(0.04)
#PortFirm	-0.012	-0.013*
	(-0.91)	(-1.90)
Year FE	No	No
Gvkey × Year FE	Yes	Yes
Mgrno FE	Yes	Yes
N	10,386,045	10,386,045
Adjusted R^2	0.286	0.240

This table reports how machine adoption by investment companies affects human-generated downloading activities using an instrumental variable. In panel A, the sample contains 17,526 investment company-years in column (1), and 3,845,698 investment company-firm-years in columns (2) and (3). The sample in panel B consists of 10,386,045 investment company-firm-years. The dependent variable in panel A column (1) is whether the investment company has adopted machine (*Post*). The instrumental variable is the log number of machine downloads from peer investment companies in the same zip code (*#PeerMachineView*). The dependent variables in panel A columns (2) and (3) and panel B are the intensity of human-generated downloading activities (*#HumanView*) and an indicator variable that measures whether the investment company has human-generated downloads of the firms' filings (*HumanView*). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Supplementary Appendix

Table S.A: Definitions of Variables

Variable	Definitions
#HumanViewScheduled	The natural logarithm of one plus the number of times an investment company views (i.e., downloads) the scheduled filings of the firm during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B. Scheduled filings include Forms 10-Q, 10-K, 13F-HR, 6-K, DEF 14A, 10KSB, 10QSB, 10-K405, and 20-F.
#HumanViewText	The natural logarithm of one plus the number of times an investment company views (i.e., downloads) the text-heavy filings of the firm during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B. Text-heavy filings include Forms 8-K, 10-Q, 10-K, 6-K, DEF 14A, 424B2, S-1, 10KSB, 10QSB, DEFA14A, 10-K405, 20-F, UPLOAD, and CORRESP.
#HumanViewTrading	The natural logarithm of one plus the number of times an investment company views (i.e., downloads) the trading-related filings of the firm during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B. Trading-related filings include Forms 4, 3, 13F-HR, SC 13G/A, SC 13G, SC 13D/A, 4/A, SC 13D, 13F-HR/A, 5, and 13F-NT.
#HumanView10K	The natural logarithm of one plus the number of times an investment company views (i.e., downloads) form 10-K of the firm during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B.
#HumanView8K	The natural logarithm of one plus the number of times an investment company views (i.e., downloads) form 8-K of the firm during year *t*, and the viewing activity is classified as human-generated based on a procedure detailed in Appendix B.

Table S.1. Descriptive Statistics

	N	Mean	SD	P25	P50	P75
#HumanViewScheduled	10,386,045	0.222	0.620	0.000	0.000	0.000
#HumanViewText	10,386,045	0.285	0.725	0.000	0.000	0.000
#HumanViewTrading	10,386,045	0.024	0.148	0.000	0.000	0.000
#HumanView10K	10,386,045	0.139	0.459	0.000	0.000	0.000
#HumanView8K	10,386,045	0.096	0.387	0.000	0.000	0.000

This table reports the descriptive statistics of the sample. All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Table S.A.

Table S.2. The Effect of Machine Adoption on Human Information Acquisition by Filing Types

Panel A. Scheduled, text-heavy, and trading-related filings

Dependent Variable:	*#HumanViewScheduled*		*#HumanViewText*		*#HumanViewTrading*	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.120***		0.154***		0.011***	
	(3.59)		(3.78)		(3.55)	
Post1		0.081***		0.096***		0.005
		(2.81)		(2.77)		(1.59)
Post2		0.111***		0.147***		0.013***
		(3.44)		(3.76)		(3.01)
Post3		0.111***		0.150***		0.008**
		(2.98)		(3.33)		(2.18)
Post4		0.164***		0.208***		0.016***
		(3.72)		(3.89)		(4.11)
Distance	-0.018***	-0.018***	-0.021***	-0.021***	-0.003***	-0.003***
	(-9.33)	(-9.33)	(-9.80)	(-9.80)	(-10.44)	(-10.46)
RelationDuration	0.067***	0.067***	0.071***	0.070***	0.006***	0.006***
	(5.55)	(5.56)	(5.69)	(5.71)	(4.37)	(4.34)
Shares	26.501***	26.694***	32.629***	32.864***	5.546***	5.567***
	(4.53)	(4.58)	(5.06)	(5.11)	(4.78)	(4.81)
PortSize	0.008	0.007	0.011	0.011	0.002**	0.001**
	(1.37)	(1.33)	(1.63)	(1.59)	(2.02)	(2.03)
PortHHI	-0.026	-0.034	-0.030	-0.041	-0.002	-0.003
	(-0.43)	(-0.57)	(-0.42)	(-0.58)	(-0.21)	(-0.32)
PortRet	0.010	0.005	0.013	0.008	-0.002	-0.002
	(0.47)	(0.25)	(0.51)	(0.32)	(-0.59)	(-0.67)
PortVol	0.073	0.061	0.110	0.088	0.002	-0.000
	(0.74)	(0.60)	(0.90)	(0.71)	(0.10)	(-0.02)
#PortFirm	0.001	0.000	0.001	0.000	-0.001	-0.001
	(0.08)	(0.03)	(0.11)	(0.04)	(-0.88)	(-0.90)
Gvkey×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes	Yes	Yes
N	10,386,045	10,386,045	10,386,045	10,386,045	10,386,045	10,386,045
Adjusted R^2	0.251	0.252	0.271	0.271	0.093	0.093

Panel B. 10-K and 8-K filings

Dependent Variable:	*#HumanView10K*		*#HumanView8K*	
	(1)	(2)	(3)	(4)
Post	0.083***		0.058***	
	(3.44)		(3.17)	
Post1		0.049**		0.018
		(2.36)		(1.19)
Post2		0.073***		0.051***
		(3.23)		(2.99)
Post3		0.074***		0.055***
		(2.79)		(2.78)
Post4		0.121***		0.098***
		(3.83)		(3.80)
Distance	-0.011***	-0.011***	-0.008***	-0.007***
	(-8.44)	(-8.44)	(-9.00)	(-8.99)
RelationDuration	0.050***	0.049***	0.029***	0.028***
	(4.85)	(4.85)	(4.46)	(4.47)
Shares	16.187***	16.360***	14.827***	15.001***
	(3.38)	(3.42)	(4.01)	(4.06)
PortSize	0.004	0.004	0.002	0.002
	(1.00)	(0.93)	(0.78)	(0.67)
PortHHI	-0.026	-0.033	-0.016	-0.024
	(-0.62)	(-0.81)	(-0.53)	(-0.80)
PortRet	0.003	-0.002	0.011	0.007
	(0.18)	(-0.11)	(0.91)	(0.60)
PortVol	0.057	0.047	0.036	0.022
	(0.82)	(0.66)	(0.71)	(0.42)
#PortFirm	0.001	0.000	0.003	0.003
	(0.09)	(0.03)	(0.66)	(0.56)
Gvkey×Year FE	Yes	Yes	Yes	Yes
Mgrno FE	Yes	Yes	Yes	Yes
N	10,386,045	10,386,045	10,386,045	10,386,045
Adjusted R^2	0.220	0.220	0.166	0.167

This table reports how machine adoption by investment companies affects human-generated downloading activities by different filing types. The sample consists of 10,386,045 investment company-firm-years. The dependent variables in panel A are the intensity of human-generated downloading activities of scheduled filings (*#HumanViewScheduled*), text-heavy filings (*#HumanViewText*), and trading-related filings (*#HumanViewTrading*). The dependent variables in panel B are the intensity of human-generated downloading activities of 10-Ks (*#HumanView10K*) and 8-Ks (*#HumanView8K*). All continuous variables are winsorized at 1 and 99 percentiles. All variables are defined in Appendix A and Table S.A. *t*-statistics, in parentheses, are based on standard errors clustered by investment company. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.