From Man vs. Machine to Man + Machine:

The Art and AI of Stock Analyses*

Sean Cao† Wei Jiang‡ Junbo Wang§ Baozhong Yang¶

This Draft: November 2022

ABSTRACT

We train an AI analyst that digests corporate disclosures, industry trends, and macroeconomic indicators to the extent it beats most analysts. Human wins the "Man vs. Machine" contest when a firm is complex with intangible assets, and AI wins when information is transparent but voluminous. Analysts catch up with machines over time, especially after firms are covered by alternative data and their institutions build AI capabilities. AI power and human wisdom are complementary in generating accurate forecasts and mitigating extreme errors, portraying a future of "Man + Machine" (instead of human displacement) in financial analyses, and likely other high-skill professions.

*The authors have benefited from discussions with Simona Abis (discussant), Francesco D'Acunto (discussant), Svetlana Bryzgalova (discussant), Si Cheng (discussant), Jungho Choi (discussant), Will Cong, Jin-chuan Duan, Michael Gofman (discussant), Paul Goldsmith-Pinkham (discussant), Jillian Grennan, Bing Han, Jingyu He (discussant), Gerard Hoberg, Byoung-Hyoun Hwang, Mark Liu (discussant), Bin Li (discussant), Roxana Mihet (discussant), Markus Pelger, Uday Rajan, Siew Hong Teoh, Paul Tetlock, Eric Yeung (discussant), Hong Zhang, Xiaoyan Zhang, Guofu Zhou, Luofeng Zhou, and Christina Zhu (discussant), and comments and suggestions from seminar/conference participants at the NBER Economics of AI Conference, the Stanford Engineering AI & Big Data in Finance Research Forum (ABFR) webinar, AllianceBernstein, Baruch, Case Western, CICF, CFRC, CFTRC, CKGSB, Copenhagen Business School, CUHK-Shenzhen, Econometric Society China Meetings, ESSEC, FARS, Fordham, Frankfurt School of Finance & Management, Fintech: Innovation, Inclusion and Risks Conference, HARC, ITAM, Louisiana State University, MIT Sloan, National University of Singapore, Northern Illinois University, Stockholm School of Economics AI conference, Tsinghua PBC, University of Florida Machine learning Conference, University of Mississippi, University of New Orleans, the Annual Conference on Digital Economics, the Quantitative Work Alliance at Boston (QWAFAFEW), the 28th Finance Forum at the Nova School of Business and Economics, NEOMA, the Midwest Finance Annual Conference, the Northern Finance Association conference, 3rd Shanghai Financial Forefront Symposium, and the UT Dallas Finance Conference. We appreciate the computing resources offered by the Advanced Research Computing Technology and Innovation Core at GSU and excellent research assistance by Soyun Nam and Meng Wang.

†Robert H. Smith School of Business, University of Maryland. Email: scao824@umd.edu

‡Goizueta Business School, Emory University. Email: wj2006@columbia.edu

§E.J. Ourso College of Business, Louisiana State University Email: junbowang@lsu.edu

¶J. Mack Robinson College of Business, Georgia State University. Email: bzyang@gsu.edu

Key Words: Artificial Intelligence; Machine Learning; FinTech; Stock Analyst; Alternative Data; Disruptive Innovation

1. Introduction

Since the inception of artificial intelligence (AI) and as it continues to rise, AI has constantly made human beings rethink their own roles. While AI is meant to augment human intelligence, concerns abound that it could replace humans in increasingly skilled tasks and thus displace jobs currently performed by the better-paid and better-educated workers (Muro, Maxim, and Whiton, 2019). Such concern and the associated debates have motivated a quickly growing literature. Recent works by Webb (2020), Acemoglu, Autor, Hazell, and Restrepo (2022), Babina, Fedyk, He, and Hodson (2020), and Jiang, Tang, Xiao, and Yao (2022) have all conducted large-sample analyses on the extent of job exposure and vulnerability to AI-related technology as well as the consequences for employment and productivity.

The existing literature has been mostly focusing on characterizing the type of jobs that are vulnerable to disruption by AI's evolution, as well as those it could create. In other words, the sentiment of the existent studies mostly involves a theme of "Man versus Machine," which characterizes the contest between humans and AI, explores ways humans adapt, and predicts the resulting job redeployments. In such settings, human beings are often rendered passive or reactive—dealing with disruptions and looking for new opportunities defined by the AI landscape. There has been relatively little research devoted to prescribing how skilled human workers could tap into a higher potential with enhancement from AI technology, which is presumably the primary goal for humans to design and develop AI in the first place. This study aims to connect the contest of "Man versus Machine" ("Man vs. Machine" hereafter) to a potential equilibrium of "Man plus Machine" ("Man + Machine" hereafter).

Our study could be motivated by the experience of chess grand master Garry Kasparov. The story that IBM's Deep Blue beat the then reigning grand master in 1997 is well-known. Afterwards, multiple contests repeated in a similar setting killed any remaining suspense for the outcome of Man vs. Machine in chess playing. What is far less known is that humans, despite having lost interest in Man vs. Machine chess contests, have not lost interest in either the game or the machine. In fact, the encounter with Deep Blue was a catalyst for people like Kasparov to pioneer the concept of Man + Machine matches, in which a chess player equipped with AI assistance (a "centaur" player) competes against AI. Up to today, the centaur has kept an upper hand against machines; and even more encouragingly, there have been more and better human chess players with the advent of affordable AI-based chess programs.[1]

If AI can help more humans become better chess players, it stands to reason that it can help more of us become better at many skilled jobs, including pilots, medical doctors, and investment advisors. In this study, we zoom into the profession of stock analysis, whose data availability allows us to calibrate both Man vs. Machine and Man + Machine. Stock analysts are among the most important information intermediaries in the market place (e.g., Brav and Lehavy, 2003; Jegadeesh, Kim, Krische, and Lee, 2004; Crane and Crotty, 2020). Their job, which requires both institutional knowledge and data analytics, has not been spared by AI, as making powerful and fast predictions at a relatively low cost is at the technology's heart (Agrawal, Gans, and Goldfarb, 2018). More and more investors have begun to heed AI-powered recommendations about stock picking and portfolio formation.[2]

To trace out the path from "Man vs. Machine" to "Man + Machine," we decided to build our own AI model for 12-month stock predictions so that we have a consistent and time-adapted benchmark for AI performance that we understand and are able to explain. Target prices and earnings are the two primary subjects of analyst forecasts; we choose the former because the latter are subject to managerial discretion, as made manifest by a large body of accounting literature on earnings management. Our "AI analyst" is built on training a combination of current machine-learning (ML) tool kits using timely, publicly available data and information. More specifically, we collect firm-level, industry-level, and macroeconomic

[1]Source of information: *The Inevitable*, by Kevin Kelly, Penguin Publishing Group, 2016. See also "Defeated chess champ Garry Kasparov has made peace with AI," *Wired*, February 2020.

[2]Sources: "What machine learning will mean for asset managers," Robert C. Pozen and Jonathan Ruane, *Harvard Business Review*, December 3, 2019. "How startup investors can utilize AI to make smarter investments," Jia Wertz, *Forbes*, January 18, 2019.

variables, as well as textual information from firms' disclosures (updated to right before the time of an analyst forecast), as inputs or predictors, but we deliberately exclude information from analyst forecasts (past and current) themselves. We resort to machine learning models instead of traditional economics models (such as regressions) due to the advantages of the former in managing high-dimensional unstructured data and in their flexibility in optimizing and fitting unspecified functional forms. More recent development in the area also allows us to mitigate overfitting and improve out-of-sample performance.

We select a set of state-of-the-art machine learning models and build our AI analyst based on an ensemble model. Our AI analyst is able to beat human analysts as a whole: the AI analyst outperforms 55.9% of the target price predictions made by all IBES analysts during the sample period of 2001 to 2018. The machine's advantage could arise from its superior ability to process information, or its immunity from predictable human biases due to incentives or psychological traits (e.g., Abarbanell, 1991; Stickel, 1990). To separate the two, we compare AI forecasts with "debiased" analyst forecasts where biases are predicted and then removed using machine learning (henceforth, "Machine-debiased Man" or "MDM" forecasts). Such an improved version of human analyst still trails the machine (MDM only outperforms AI in 48.2% of forecasts), suggesting that "correctable" biases explain around 69% of the Man-Machine gap.[3]

Despite the power of AI model, we are interested in knowing the circumstances under which human analysts retain their advantage, in that a forecast made by an analyst beats the concurrent AI forecast in terms of lower squared forecast error relative to the ex post realization (i.e., the actual 12-month stock price). We find human analysts perform better for smaller and more illiquid firms and those with asset-light business models (i.e., higher intangible assets), consistent with the notion that such firms are subject to higher information asymmetry and require better institutional knowledge or industry experience to decipher. Analysts affiliated with large brokerage houses also stand a higher chance of beating the

[3]Since analysts outperform AI in 44.2% of the cases, the percentage attributable to bias correction is (48.2 - 44.2)/(50 - 44.2) = 69%.

AI, thanks to a combination of their abilities and the research resources available to them. Analysts are more likely to have the upper hand when the firm is in a dynamically competitive environment or is subject to higher distress risk, revealing again AI's limitation in analyzing uncertain and/or unfamiliar scenarios.[4] As expected, AI enjoys a clear advantage in its capacity to process information and is more likely to outsmart analysts when the volume of public information is larger.

Just like the centaur chess player which Kasparov pioneered, the superior performance of an AI analyst does not rule out the value of human inputs. If humans and machines have relative advantages in information processing and decision making, then human analysts may still contribute critically to a centaur analyst: a human analyst who makes forecasts that combine their own knowledge and the outputs and recommendations from AI models. After we add analyst forecasts to the information set of the machine learning models underlying our AI analyst, the resulting "Man + Machine" model outperforms 55.8% of the forecasts of the AI-only model. Thus, the AI analyst does not displace human analysts yet, and in fact, an investor or analyst who combines AI's computational power and the human art of understanding soft information can still outpower AI itself.

We are thus interested in knowing when the incremental value of humans to a Man + Machine model is the highest, as manifested in the relative performance of that model versus the pure AI model. Similar to previous findings, we find inputs from analysts are more valuable when covering firms that are more illiquid and those with more intangible assets. Moreover, analyst inputs have more incremental value when a firm faces higher distress risk. Importantly, the incremental value of humans does not decrease as the volume of information (hence the demand for processing capacity) increases, though this constitutes a disadvantage for humans working alone. Similarly, analysts from small brokerage houses make a similar level of contribution to the Man + Machine model compared to their counterparts from

[4] This is consistent with the limitation of current machine learning and AI models which still lack reasoning functions to handle unfamiliar situations well. Source: "What AI still can't do," Brian Bergstein, *MIT Technology Review*, February 19, 2020.

larger banks, suggesting that AI could potentially help level the disparity along professional hierarchy. Importantly, the Man + Machine model avoids 76.3% of extreme errors[5] made by analysts and 32.2% of those by the AI (while with minimal creation of its own large errors). To the extent that large errors are calamitous in many skilled professions, there is substantial benefit in combining human and AI capabilities.

Finally, we resort to an event study to sharpen the inference of the impact of integrating humans and machines in stock analyses. In recent years, the infrastructure of "big data" has created a new class of information about companies that is collected and published outside of the firms, and such information provides unique and timely clues into investment opportunities. An important and popular type of alternative data captures "consumer footprints," often in the literal sense such as satellite images of retail parking lots. Such data, which have to be processed by machine learning models, have been shown to contain incremental information for stock prices (Zhu, 2019; Katona, Painter, Patatoukas, and Zeng, 2022). We build on data from Katona, Painter, Patatoukas, and Zeng (2022) on the staggered introduction of several important alternative databases and conduct a difference-in-differences test of analysts' performance versus our own AI model before and after the availability of the alternative data. The underlying premise is that analysts who cover firms while using this alternative data could be in the situation of Man + Machine, as they have the opportunity to use the additional AI-processed information. Indeed, we find that post alternative data, analysts covering affected firms improve their performance relative to the AI model we build which serves as a benchmark. Furthermore, such improvement concentrates in the subset of analysts who are affiliated with brokerage firms with strong AI capabilities, measured by AI-related hiring using the Burning Glass U.S. job posting data[6] and the classification algorithm developed in Babina, Fedyk, He, and Hodson (2020).

[5]Extreme errors are forecast errors those above the 90th percentile of all analyst forecast errors on the same firm in the past three years.

[6]Burning Glass is currently the leading data vendor of U.S. job postings. The postings are scraped from websites, newsletters, and agency reports. They cover the period of 2007 and then 2010–2019. Acemoglu, Autor, Hazell, and Restrepo (2022) show that Burning Glass data cover 60–80% of all U.S. job vacancies.

Overall, this study supports the hypothesis that analyst capabilities could be augmented by AI, and more importantly, that analysts' work possesses incremental value such that they, with the assistance of AI, can still beat a machine model that does not include human inputs, analogous to centaur chess players' outperforming machines. If there is some external validity from chess and stock analysis to skilled workers in general, the inference from our study is encouraging news for humans in the age of AI.

2. Literature, Data Construction, and Machine Learning Models

2.1. Relation to the Literature

Our work is related to the rapidly growing literature on the competition and threat to human workers posed by new technology including robots and AI.[7] This literature overall finds that when low- or intermediate-skill jobs are replaced by machines, humans tend to move to high-skill jobs that are more difficult to replace (Autor, Levy, and Murnane, 2003). However, the most recent wave of AI innovations disrupt many high-skill jobs. Our study focuses on humans' relative advantage over machines and, more importantly, the potential synergies between humans and machines.[8] We envision a future in which AI and machines can assist humans with the more tedious and quantitative tasks and democratize access to information, allowing humans to be more creative and productive.[9]

A few recent and contemporaneous papers also study the impact of big data and AI in

[7] An incomplete list of recent papers includes Aghion, Jones, and Jones (2017), Acemoglu and Restrepo (2018), Acemoglu and Restrepo (2019), Brynjolfsson, Mitchell, and Rock (2018), Webb (2020), Ray and Mookherjee (2021), Cao, Cong, and Yang (2019), Acemoglu, Autor, Hazell, and Restrepo (2022), and Jiang, Tang, Xiao, and Yao (2022).

[8] An example in a non-finance setting comes from sports. After trying video assistant referee (VAR) technology for a few seasons, the English Premier League decided not to let VAR overpower referee judgment. One main reason is that players will reverse-engineer the rules underlying the VAR decisions and play to their advantage, such as committing more "low-degree" (to the machine) but atrocious (to humans) fouls. This can be remedied by giving a human referee the final say. See "Why has the introduction of video technology gone so badly in soccer?" James Reade, *Forbes*, December 10, 2020.

[9] Due to the complementary nature of AI and humans, the advent of AI technologies can potentially create more jobs than they destroy. See "Artificial intelligence to create 58 million new jobs by 2022, says report," Amit Chowdry, *Forbes*, September 18, 2018.

the financial industry. Abis (2022) studies how quantitative investment strategies influence mutual fund performance. Abis and Veldkamp (2022) examine the change in labor shares in the financial industry driven by the new data management and AI jobs. Coleman, Merkley, and Pacelli (2022) compare the performance of robot analysts from fintech companies with that of human analysts. Grennan and Michaely (2020, 2021) study how analysts perform and adjust in response to the advent of AI-processed recommendations in the markets. Rossi and Utkus (2021) compare human asset managers with robot advisors. Agrawal, Gans, and Goldfarb (2019b) discuss the ambiguous impact of AI on labor given the elements of AI that tend toward automating decisions versus enhancing human decisions. Jansen, Nguyen, and Sham (2021) analyze human and machine decisions in loan underwriting. Cao, Jiang, Yang, and Zhang (2022) study the impact of AI readership on corporate disclosure policies. Finally, Pagliaro, Ramadorai, Rossi, Utkus, and Walther (2022) consider human interactions with algorithmic wealth management advisors. Our paper differs from the existing literature in that we explore the internal mechanism of the AI process we constructed ourselves instead of market-level proxies,[10] and aim to identify their relative advantages to, as well as synergies with, humans using model inputs and outputs in our own hands.

We also contribute to the literature of building and assessing the performance of machine learning models in financial applications, such as in predicting asset prices (Gu, Kelly, and Xiu, 2020, Brogaard and Zareei, 2022), robo-advising (D'Acunto, Prabhala, and Rossi, 2019), managing portfolios (Chen, Pelger, and Zhu, 2022; Cong, Tang, Wang, and Zhang, 2022), estimating values of artwork (Aubry, Kraeussl, Manso, and Spaenjers, 2022), forecasting earnings (van Binsbergen, Han, and Lopez-Lira, 2022; Cao and You, 2021, Silva and Thesmar, 2021), making lending decisions (Liu, 2022), classifying and evaluating innovations (Chen, Wu, and Yang, 2019; Zheng, 2022) and estimating bank risk (Hanley and

[10]For example, Grennan and Michaely (2020) resort to the amount of social media information as a proxy for the AI research intensity for a stock, and focus on analysts' response to the AI shock. This study, in contrast, aims at decipher the nature of the AI shock.

Hoberg, 2019).[11] While the structure of our analysis shares similarity with some of the papers, notably Silva and Thesmar (2021) and van Binsbergen, Han, and Lopez-Lira (2022) which calibrate biases in analyst expectations regarding earnings using author-developed AI model, the primary research questions of our paper are different from theirs. Silva and Thesmar (2021) and van Binsbergen, Han, and Lopez-Lira (2022) focus on the term structure of analyst biases, and link them to corporate actions such as security issuance, while our ultimate goal is to explore the complementary value humans can offer in the age of AI once we have a good understanding of their relative advantage.

In summary, our study contributes to the emerging literature studying the implications of combining humans and machines in the financial markets.[12] Given the increasing presence of machines and AI beyond finance, we also hope that this case study contributes to a better understanding of how technology can complement and improve humans, bringing to fruition the original mission of AI development.[13]

2.2. Sample of Forecasts

Our sample of analysts' 12-month target price forecasts builds on the Thomson Reuters I/B/E/S analyst database using data from 1996 to 2018.[14] We choose 12-month target price forecasts because the target prices for other horizons are less than 1% of the volume of 12-month forecasts after combining with our predictor data described below. After merging I/B/E/S with CRSP and Compustat, the final sample consists of 948,054 12-month target price forecasts on 6,190 firms, issued by 11,341 analysts from 820 brokerage firms.

[11] See also Cong, Liang, Yang, and Zhang (2020), Martin and Nagel (2022) and Goldstein, Spatt, and Ye (2021) for surveys and discussions of methodologies.

[12] In different settings, Armour, Parnham, and Sato (2020) study the impact of AI and the associated digital technologies on the law profession. They find that AI-enabled services will augment the capabilities of human lawyers and also generate new roles for legal experts to produce such services. Brogaard, Ringgenberg, and Rösch (2021) find that human floor traders can complement algorithmic traders in providing information to the market in complicated environments.

[13] This echoes the mission of the Stanford Human-Centered AI Institute, "to advance AI research, education, policy and practice to improve the human condition." See https://hai.stanford.edu/about.

[14] The I/B/E/S coverage prior to 1996 was limited, with fewer than 2,000 target price forecasts in total.

We choose to analyze analysts' price forecasts instead of earning forecasts because earnings are subject to managerial discretion or even manipulation. Earnings could also be endogenous to analyst forecasts due to the well-documented feedback loop caused by the managerial incentive to meet and beat analyst consensus (Abarbanell and Lehavy, 2003; Doyle, Jennings, and Soliman, 2013). Prices are more difficult to control by insiders and thus provide a more objective benchmark to assess performance of human analysts and machines.[15] Unlike earnings forecasts which could target a fixed firm-quarter in the future by the same analyst, forecasts in our sample are for 12-month moving windows. We thus consider all such forecasts to be newly initiated instead of being revisions of some previous forecasts.

2.3. Building the Information Set for the AI Analyst

Given our goal to build an AI analyst to compete with professional analysts, we need to define the information set available to such a professional whenever a price forecast is made. The unit of analysis in our main setup is a forecast on the 12-month stock price for firm i by human analyst k on date t (in year u). The information set, $\mathcal{I}_t$, would, in an ideal setting, include all publicly available data and information up to $t-$. We assume that professional analysts do not have access to material nonpublic information, which is essentially the requirement of Regulation FD.[16] We approximate $\mathcal{I}_t$ with firm and industry information from CRSP and Compustat; textual information from firms' SEC filings, including annual reports (10-K), quarterly reports (10-Q), ad hoc disclosure of material corporate news and developments (8-K), and other reports; and macroeconomic data from the Federal Reserve Economic Data at the Federal Reserve Bank of St. Louis.

[15]While some studies document that analyst forecast revisions lead to stock price reactions (e.g., Stickel, 1991), such effects typically last much shorter than the 12-month horizon we examine in this study, and have substantially diminished since 2003 due to markets that more efficiently incorporate information (Altinkilic, Hansen, and Ye, 2016).

[16]Regulation FD ("fair disclosure"), implemented in 2000, generally prohibits public companies from disclosing previously nonpublic, material information to certain parties unless the information is distributed to the public first or simultaneously.

To operationalize time adaptation, we adopt the following rolling-window approach. For a given forecast made by a human analyst on date t in year u, all forecasts in the previous three years $u-3, u-2, u-1$ form the training sample.[17] That is, data up to the dates of those forecasts (but excluding the forecasts themselves) and the corresponding realized prices were used to train our machine learning models. Moreover, if the past three years include a "distress" year (defined by negative market return), we trace back to the first year of the distress s and expand the training window to years $s-3, s-2, s-1, \cdots, u-1$. The benefit of this approach is that human analysts in a recession likely predict future prices based on information over a full business cycle. Including the years before distress can mimic the information used by human analysts. Moving to the estimation sample, we then feed data available up to date $t-1$ in year u into the trained model to make the 12-month-price prediction at time t. Our AI analyst makes its first prediction in 2001. Though we allow (public) information to be updated till $t-1$, most of the information inputs came from disclosed quarterly data from the previous eight quarters.[18]

2.4. Information and Variables as Inputs to Machine Learning

Firm Characteristics The firm characteristics fed into machine learning models are retrieved or processed based on information from standard databases accessed via WRDS, especially CRSP/Compustat and Thomson Reuters Ownership databases. The first set of predictors include stock prices at the end of the previous month as well as the stock prices one to four years before the end of the previous month. The 12-month returns over the past 5 years are also included, together with the realized earnings within the past 3, 6, 9, 12, 24, and 36 months. We also include a number of firm characteristics known to predict cross-sectional differences of the stock prices. In particular, we include anomalies from each of the

[17]We have also trained the model with a five-year rolling window and obtained similar results. Downsides from using a much longer training windows include the shrinkage of estimation period and added noise from more distant historical data.

[18]Data that come in different frequencies, such as corporate news releases (8-K), are aggregated at the quarterly level.

six broad categories considered in Hou, Xue, and Zhang (2020): the momentum, value versus growth, investment, profitability, intangibles, and trading frictions anomalies.[19] Variables in this group are constructed quarterly using information available at the previous quarter-end.

Industry Variables We compose a set of industry-level variables that capture competition, industry dynamics, and other factors relevant for firm valuation based on the existent literature. These variables include (i) The competition measure from 10-Ks following Li, Lundholm, and Minnis (2013), which captures the degree of competition resulting from rivalry within and across industries as perceived by the management; (ii) the product market fluidity measure following Hoberg, Phillips, and Prabhala (2014), which quantifies the product market poaching threat posed by the movement of competitors toward the focal firm; (iii) industry affiliation with the Fama-French 48 industries (48 indicator variables); (iv) industry size, measured by the number of firms in the Fama-French 48 industry within the past 3, 6, 9, 12, 24, and 36 months; and (v) equally weighted industry average earnings per share realized within the past 3, 6, 9, 12, 24, and 36 months.

Macro Variables Macroeconomic and stock market development are common factors to all firms' valuation and returns (e.g., Fama and French, 1989; Chen, Roll, and Ross, 1986). For this category, we include the following variables: (i) Industrial Production Index; (ii) Consumer Price Index; (iii) Crude oil price (WTI); (iv) three-month treasury bill rate; (v) ten-year treasury constant-maturity rate; and (vi) The BAA–AAA yield spread. These macro variables are obtained from the Federal Reserve Economic Data at the Federal Reserve Bank of St. Louis on a monthly frequency.

Textual Sentiment Information One leading strength of AI over human beings is the former's ability to digest large volume of information. One new edge that machine learning models boast over traditional statistical methods is the capacity to process unstructured

[19] We list all variables serving as inputs into the machine learning models, their definitions, and sources in Table A1 in Appendix A.

textual data based on firms' SEC filings, including annual reports (10-K), quarterly reports (10-Q), corporate news (8-K), and other reports. The new developments allow researchers to quantify information which was considered qualitative or "soft," commonly termed "sentiments."

Two different sets of sentiment variables from textual data serve as inputs to our AI analyst. The first is based on the Loughran and McDonald (2011) sentiment, which has been widely used in the academic literature. We calculate the frequency of positive and negative sentiment words from the firm-issued SEC filings following Loughran and McDonald (2011). The second set of machine-learning-based sentiment variables follow Cao, Kim, Wang, and Xiao (2020), who trained a deep-learning neural network model to incorporate contextual information and syntactic relations between performance-related words. The second approach aims to isolate managerial sentiment related to the firm's future performance from sentiment regarding other issues (such as location and weather).

2.5. Potential Factors for the Relative Performance of AI and Human Analysts

A main objective of the study is to assess the factors that contribute to the relative performance of AI vs. humans as well as the synergy between the two. We hypothesize that these factors are related to the information environment of the firm, industry, and analysts. Needless to say, equity analysts are often evaluated along dimensions other than forecast accuracy, such as promoting investment banking or trading businesses, and intermediating between firms and large investors. We focus on forecast accuracy, not only because it is objective and quantifiable, but also because it represents a primary quality in analyst evaluation (Stickel, 1992; Desai, Liang, and Singh, 2000).

We first consider the following firm-level variables: the *Amihud Illiquidity* measure (Amihud, 2002), which is the ratio of absolute daily stock return to the daily trading volume (in dollars); *Log Market Cap*, the logarithm of market capitalization; *Standard Deviation of*

Earnings; *# 8K Reports*, which is the number of 8-K reports filed each year and represents the volume of available information about the firm; and *Intangible Assets*, defined as the first principal component of four proxies: intangible assets minus goodwill divided by total assets, one minus the ratio of PP&E to total assets, organization capital scaled by assets, and knowledge capital scaled by assets. The last two measures, derived from the accumulation of SG&A and R&D expenditures, are constructed following Ewens, Peters, and Wang (2022) (see e.g., Eisfeldt and Papanikolaou, 2013, 2014; Peters and Taylor, 2017; Falato, Kadyrzhanova, and Sim, 2020 for the modeling and development of these and related measures of intangible capital).

We further include a number of variables that characterize the information environment and resources for analysts: *Star Analyst* represents an "all-star" status awarded by the *Institutional Investor* magazine at the beginning of the year;[20] *# Analysts in Brokerage Firm* is the number of analysts and proxies for the size and resources of the brokerage firm; *% Institutional Holdings* is the 13F institutional holdings as a percentage of shares outstanding, which can reflect the prevalence of informative investors; *Book Leverage* is constructed following Fama and French (1992) and proxies for firms' financial exposure to distress shocks (Babina, 2020); *Fluidity* represents the competition firms face in the product markets by tracing changes in rival firms' products relative to the firm's products (Hoberg, Phillips, and Prabhala, 2014); and *Time Trend* equals the number of years elapsed from the beginning of the sample (2001). The final set of variables are related to analysts' access to alternative data and AI resources, which will be introduced in Section 5.3. Table 1 presents the summary statistics of variables.

[Insert Table 1 here]

[20] See more details at `https://www.institutionalinvestor.com/research`.

2.6. Machine Learning Models

There are a number of candidate machine learning models developed in recent decades to build our AI analyst, including lasso, elastic-net, support vector machines, random forest, gradient boosting, and long short-term memory neural networks. Because the machine learning models are essential tools but not the ultimate objectives of this study, we provide an overview in Appendix B of the models referenced where we focus on the economic intuition of each methodology's mechanism and strength without going into technical details. For further details, we refer the reader to the most representative references in this field; for example, Hastie, Tibshirani, and Friedman (2009) and Goodfellow, Bengio, and Courville (2016). Of the models considered, random forest, gradient boosting, and long short-term memory neural networks are the state-of-art nonlinear models that have been increasingly used and proved of their advantage over the other methods in the existent literature of computer sciences, finance, and other disciplines. Our main AI model thus is built as an ensemble of these three models, i.e., adopts the mean prediction of the three models.[21]

Our candidate machine learning models strive to be at the leading edge of AI practice in investment management. They are similar to those covered in two prominent industry reports: The JP Morgan Big Dada and AI Strategies report and the report on Artificial Intelligence in Asset Management by the CFA institute, and are also favored in the current industry practice.[22]

[21]While the model selection was based on the literature review and reasoning, we nevertheless evaluate the performance of all candidate models for each year on a rolling basis. For each given year, we take the prior three years and split it into a training sample (first two years) and a testing sample (the third year). We train the model parameters over the training sample and then evaluate the model performance on the testing sample. Results are shown in Internet Appendix Table IA1. Perhaps not surprisingly, the three models turn out to be the best performing ones.

[22]These reports can be found at the following links: `https://www.cognitivefinance.ai/single-post/big-data-and-ai-strategies` and `https://zonavalue.com/wp-content/uploads/2020/09/CFA-Institute-artificial-intelligence-in-asset-management.pdf`. We have presented and discussed our paper and models with about half a dozen teams who are leaders in AI-directed investments. Most importantly, we confirmed with these teams that the rates at which AI models beat human analysts are on par with the current state-of-art.

3. Construction and Performance of the AI Analyst

Our ultimate goal is not to build an AI analyst per se, but to analyze the relative strength as well as the synergy between machines and humans. Such a goal, however, sets the premise that the AI analyst at hand needs to be strong enough to compete with, or even beat, human analysts. This section describes how we build an AI analyst of that calibre.

3.1. The Predictive Models

For each stock i at time t, where t is the day when an analyst makes a forecast, $F^{Man}_{i,j,t}$ (wherein i, j, t are indices for the stock, the analyst, and the date, and the superscript Man indicates human as opposed to AI), of the 12-month target price, our model makes an attempt at predicting the same target; that is, the stock price $P_{i,T(t)}$, wherein $T(t)$ is the last trading day of the 12-month forward price from time t. Included in the predictive information set is all public information (as described in Section 2 and Appendix A) up to $t-$. We summarize the prediction model as

$$log(P_{i,T}) = F^{AI}_{i,t} + \epsilon_{i,t},\ F^{AI}_{i,t} = f_{t-}(X_{i,t-}). \tag{1}$$

Here, f_{t-} is the prediction function for all stocks at time $t-$. This is consistent with asset pricing models with conditioning information; that is, we assume there is a uniform prediction model for every stock at a given time while allowing the model to be time-varying. In (1), stock price is expressed in natural logarithm in order to rule out explosive variables and ensure the usual regularity conditions for the model estimate to converge.

Next, we compare the AI forecast $F^{AI}_{i,t}$ and the analyst forecast $F^{Man}_{i,j,t}$ in terms of their accuracy relative to the ex post realized price $P_{i,T}$. AI beats human if $|F^{AI}_{i,t} - log(P_{i,T})| < |F^{Man}_{i,j,t} - log(P_{i,T})|$, and vice versa. We define *Beat* to be an indicator variable for AI winning. Moreover, we define $F^{AI}_{i,t} > F^{Man}_{i,j,t}$ to be a "buy" signal and the opposite condition to be a "sell" signal. Figure 1 shows the relative performance of AI vs. human analyst forecasts over time.

[Insert Figure 1 here]

Out of 820,099 forecasts made from 2001 to 2018, the average human analyst is only able to beat the AI 44.1% of the time. The p-value for the percentage to be drawn from a distribution with the neutral probability of half for this sample size is less than 0.01%. However, the human analyst disadvantage is volatile from year from year, ranging from 25.9% in 2001 to 51.8% in 2010 with an overall positive drift, as shown by the trend-fitted line. We conjecture (which will be tested) that the waning advantage of the AI analyst is precisely because human analysts have been adapting to technology. Not only analysts increasingly have access to and are assisted by improving technologies in data collection, statistical packages, and machine learning tools, but also they could specialize and move their coverage to firms that are less followed by machine (Grennan and Michaely, 2020).

3.2. Contribution of Variables to the AI Prediction

In this section, we examine the contribution of different groups of input variables to the predictions of the AI model. We divide the features into six groups: returns (in the past five years), firm characteristics, earnings (past firm and industry earnings), industry information, macroeconomic variables, and variables using textual information. The contribution from each group is the difference in forecast performance between the full information model and one that omits the given group. Specifically, we compute the percentage of times that the AI model with complete information beats the model that is otherwise the same except excluding that given set of information. A probability of 50% reflects neutrality in relative performance. We then scale the percentages representing the incremental effect of each group (i.e., above 50%) by the total sum so that they sum up to unit. Figure 2 presents the results.

[Insert Figure 2 here]

Each group of features contributes substantially to the AI prowess. Firm-specific information from past earning variables and firm characteristics contribute the most (23.7% and 22.4%, respectively), followed by the industry (17.6%) and macro (16.5%) indicators. The 10.4% contribution from textual information highlights the importance of qualitative information. It is perhaps not surprising that information from the stock market (past returns) claims the lowest share (9.5%), as their contribution should be minimal if the stock market is close to being informationally efficient.

3.3. Debiased Analysts vs. AI

It has been well documented that analysts exhibit biases in their forecasts (e.g., Abarbanell, 1991; Stickel, 1990). There are a multitude of explanations of such biases, including the incentive to issue more favorable forecasts for corporate clients of the analysts' affiliated brokerage firms (Michaely and Womack, 1999), the need to obtain access to information from the management (Lim, 2002), and human psychological traits (e.g., DeBondt and Thaler, 1990; Hirshleifer, Levi, Lourie, and Teoh, 2019). A natural question thus arises: Could human underperformance relative to AI be remedied simply by "debiasing" analyst forecasts with a machine learning model (henceforth, "Machine-debiased Man" forecasts or "MDM"), or will the human shortfall remain after such a procedure in which case it would be due to the limitation in human ability to acquire and process information? A comparison of MDM forecasts with the AI analyst would reveal the nuance regarding the innate predictive ability of analysts after filtering out their predictable biases.

In constructing the MDM forecasts, we first predict the analyst forecast errors in the next period with all current information, analogous to Equation (1).

$$F_{i,j,t}^{Man} - log(P_{i,T(t)}) = g(X_{i,t-}, Z_{i,t-}) + \epsilon_{i,t}, \tag{2}$$

where we include all variables $X_{i,t-}$ that we have employed to predict target prices, and a set of analyst and brokerage-firm characteristics $Z_{i,t-}$, including the mean and standard error of analysts' past prediction biases, analysts' experiences (number of years covering the firm, the industry, or any public firm), analysts' efforts (whether the analysts provide forecasts of additional information such as sales or cash flows), and brokerage firm size proxied by the number of analysts. We use the same procedure as in Section 3.1 to train the same machine learning model and then estimate (2). The MDM prediction is then the analyst prediction $F_{i,j,t}^{Man}$ minus the corresponding bias as predicted by the machine learning model. To compare the MDM with AI, we plot the MDM beat ratio, or the frequency of MDM forecasts beating AI forecasts, in each year from 2001 to 2018 in Figure 3. As expected, MDM exhibits better performance than the raw forecasts and beats the AI more frequently than human analysts alone in most years. Over the entire sample, MDM beat AI analysts in 48.2% of the cases, an over four percentage-point improvement over humans without debiasing. Since pre-debiased analysts outperform AI in 44.2% of the cases, we estimate that correcting human biases (without additional information about firm fundamentals) could eliminate around 69% $(= (48.2 - 44.2)/(50 - 44.2))$ of the Man-Machine gap.

[Insert Figure 3 here]

3.4. AI vs. Analysts with Persistent Performance

Analysts are a large group with heterogeneous skill levels such that forecast performance would be persistent if skills were innate. Moreover, the market recognizes, at least partially, the relatively more skilled analysts by responding more strongly to their forecasts or recommendations (Chen, Francis, and Jiang, 2005; Li, 2005; Mikhail, Walther, and Willis, 2007). Thus, a higher hurdle is for our AI analyst to beat the subset of skilled analysts. We assess the relative performance with respect to the higher hurdle with two tests. First, we sort all analysts into the top and bottom halves based on their average prediction error (normalized by stock prices) over a past period with length ranging one, two, three, four, and five years. We then track the percentage of their future forecasts that beat our AI analyst during each time period. In the second test, we repeat the same procedure except selecting the analysts that are among top and bottom quantiles each year during the past one, two, three, four and five years. The second specification is more demanding on persistent skills as only about 7.3% of the analysts are able to stay at the top half in each year for a five-year period. Table 2 reports the results.

[Insert Table 2 here]

Results in Table 2 show that the AI comfortably beats the analysts in the low-skill quantiles. It is basically neck and neck to the more successful analysts and is almost even with analysts (analyst beat ratio of 49.9%) who demonstrated superior performance in each of the past five years, an excellence only achieved by less than one tenth of all analysts.

3.5. Performance of Portfolio Following AI Recommendations

Analysts make forecasts as a way to advise portfolio formation or turnover. The performance of a portfolio following the analyst advice is thus a natural metric for analyst skill. For the same reason, we can form portfolios based on the different opinions between the AI and human analysts. The performance of the resulting portfolio is testament of their relative proficiency. Our approach is different from the usual one which follows analyst directional recommendations as our model requires a clear investment horizon that is lined up with the horizon of the signal (i.e., 12-month price target).

In each month, we gather all predictions made by all analysts and the corresponding AI forecasts in past 30, 60, 90, and 360 days. For each pair of predictions, if the AI predicts a higher (lower) price, it is considered as a buy (sell) signal. During the given time horizon, the portfolio will long the stock if there are more buy than sell signals, and short the stock otherwise. The portfolio is equal weighted. In a monthly (or six-monthly) rebalanced portfolio, we hold the position for one month (six months), or until the signals reverse, whichever is earlier. The portfolio contains 620 to 1,150 stocks with signals from past 30 days to 360 days.[23] Table 3 Panel A reports the performance of the monthly rebalanced long-short portfolio in terms of average return and alpha estimated using Fama-French three-factor, Carhart four-factor, Fama-French five-factor and Fama-French six-factor models.

[Insert Table 3 here]

Results in Table 3 are highly encouraging in that the AI is able to generate superior returns/alpha, relative to analysts, on the order of 75 to 166 basis points monthly, significant at less than the 5% level in all cases. To the extent that our portfolio approach compares the AI with all human analysts, our result implies that the AI forecast is superior to the analysts' consensus. When we separately examine the long and short portions of the portfolio, we discover that the superior returns are significant (all at the 1% level) only on the long side (for which transaction costs are lower). Such an asymmetry could be driven by the well-documented positive bias in analyst forecasts (Lim, 2002). While AI forecasts contain no average directional bias, we confirm that the median analyst price forecast in our sample contains an 8.0% positive bias. In such a scenario, signals are not as informative when analysts are more optimistic than the AI.

Then a natural question arises as whether analyst bias is the key reason making AI superior to humans. Indeed, van Binsbergen, Han, and Lopez-Lira (2022) show that analyst expectations are significantly upward biased relative to a statistically optimal unbiased machine-learning benchmark. While we confirm that a bias similar to that in van Binsbergen, Han, and Lopez-Lira (2022) exists in price target forecasts, the objective of our analysis is to show that the machine out-performance goes beyond correcting human directional bias. With the earlier build-up (Section 3.3), we can readily compare the portfolio performance between the MDM (which corrects analyst bias using a machine learning model) and the AI. We observe that the AI still generates a superior (monthly) return or alpha of 39 to 75 basis points relative to the MDM. Therefore, around half of the superior performance of the

[23] The average monthly turnover rate of the monthly rebalanced portfolios ranges from 8% to 52%; and that for the six-monthly rebalanced portfolio is 5% to 10%.

AI can be attributed to its true capability in processing data, and the other half to the AI's absence of predictable biases.[24]

3.6. Combined Wisdom of Man + Machine

Results from the previous sections suggest that the analyst profession could be seriously disrupted by AI technology. However, the superior performance of the AI analyst does not rule out the possibility that analyst forecasts contain information that is incremental to AI-produced forecasts. In other words, if analysts possess information that is not picked up by the AI, then the AI forecast is not sufficient to replace the analyst forecast, even though analysts lose to AI in forecasting accuracy. An investor who combines the wisdom of both should attain even better performance.

To assess the performance of the combined analytical power, we consider adding the analyst forecasts to the information set for our machine learning model. That is, the information set $\mathcal{I}_t$ now includes the analyst forecasts, $F_{i,j,t}$, made on the same firm i during the 90-day window ending on date t. In particular, we obtain analyst and brokerage-firm characteristics (including analysts' experiences, analysts' efforts, and the number of analysts in the brokerage firm), consensus and mean square error of the forecasts by analysts in the previous 90 days, current analysts' predictions, MDM predictions, and the consensus predictions from analysts with the lowest 50% errors over the last five years and build a "Man + Machine" hybrid analyst using the ensemble model. We find that the hybrid analyst outperforms human analysts 57.8% of the time and AI-alone forecasts 55.8% of the time.

Figure 4 plots the relative performance of the hybrid analyst (Man + Machine) vs. the plain AI (Machine). Interestingly, Man + Machine outperforms plain Machine in 16 out of 18 years (the beat ratios for two other years are 48.3% and 49.5%, which are fairly close to 50%). Such outperformance, which captures the incremental value of human analysts

[24] For details about MDM, please see the Internet Appendix Table IA2. We also show that the results are robust to the frequency of rebalancing and trading cost considerations. When the portfolio is rebalanced semi-annually (Table IA3), the above findings remain.

to machines, increases with time. The combined results portray a bright future for Man + Machine: not only does the combination attain better performance than either side alone, but also, the incremental value of humans does not weaken with the technology. Moreover, the beat ratio of the hybrid analyst is always higher than that of human analysts for all 18 years, suggesting the power of the human and AI combination over humans alone.

[Insert Figure 4 here]

4. Man vs. Machine: Relative Advantages

4.1. Determinants of Relative Performance

In this section, we strive to understand when human analysts perform better than the AI and when otherwise. Such understanding will help "unbox" the black box associated with AI or machine learning and provide intuition and guidance on the applicability of AI for researchers and investors.

We consider a number of variables at the analyst, firm, and industry levels that are potentially relevant for the performance of human analysts and AI. These are defined in Section 2.5. We group these variables into several classes. First, we consider a number of proxies for information asymmetry or opacity, including *Amihud Illiquidity*, *Log Market Cap*, *Standard Deviation of Earnings*, and *% Institutional Holdings*. Second, we include variables representing the volume of information (*# 8K Reports*) and the tangibility of information (*Intangible Assets* and *Fluidity*). Third, we examine several variables that affect the information and resources available to the analyst, such as *# Analysts in Brokerage Firm* and *Star Analyst*. Finally, we consider *Book Leverage*, highlighting financial exposure of firms to shocks (Babina, 2020), and *Time Trend*, which can help capture temporal patterns.

For each target price forecast, we define two variables that measure the relative performance of humans vs AI. First, the indicator variable *Analyst Beats AI* equals one if the absolute value of forecast error of the analyst is smaller than that of the AI, and zero otherwise. Second, the continuous measure *Forecast Error Difference* is the difference between the squared prediction error (of log price as defined in Equation (1)) of the AI and that of the analyst, scaled by the maximum of these two prediction errors if the difference is non-zero. A positive and large value of *Forecast Error Difference* is in favor of analyst accuracy.

We estimate the following regression on the panel data of firm i, analyst j, and date t to understand the determinants of the relative strengths of humans and AI,

$$Relative\ Performance_{i,j,t} = X'_{i,j,t}\beta + \alpha_i + \alpha_j + \alpha_{year} + \epsilon_{i,j,t}, \tag{3}$$

wherein the dependent variable *Relative Performance* is either *Analyst Beats AI* or *Forecast Error Difference*. The vector of independent variables, $X_{i,j,t}$, includes those discussed in Section 2.5, and α_i/α_j and α_{year} represent firm/analyst and year fixed effects, respectively. The results are reported in Table 4.

[Insert Table 4 here]

Table 4 shows that, controlling for year and firm fixed effects, humans are more likely to outperform when covering illiquid and small firms and those with higher intangible assets, consistent with the notion that such firms are subject to higher information asymmetry and require deeper institutional knowledge to understand. On the other hand, equipped with vast processing power, AI performs better for firms with a larger volume of disclosed information, as proxied by *# 8K Reports* each year. Star analysts are more likely to have a lower forecast error relative to AI, compared with non-star analysts, justifying their stardom. Analysts working for larger brokerage firms perform better, potentially because of the more abundant resources and research capacity at such places as well as a positive match between analyst skill and brokerage house prestige. Humans perform better when the focal firm is subject to higher financial distress risk, captured by book leverage, and when the firm faces higher dynamically competitive pressure, measured by fluidity, suggesting that the AI has more difficulty handling more uncertain scenarios. Analysts also perform better for firms with higher institutional holdings, possibly because analysts are immersed with information produced and processed by institutional investors, including brokerage houses. Finally, when year fixed effects are not included, we are able to uncover the time trend of the comparative performance, showing that human advantage increases with time. This is probably due to the fact that human analysts are increasingly assisted by AI and big data technologies.

4.2. Disagreement between Man and Machine

An equally important question is whom we should trust more if and when humans and machines disagree. To start with, Figure 5 plots the annual time series of the average squared differences in (logged) forecasts between analysts and AI. Interestingly, we find that the man-machine disagreement has been on a downward trend, possibly because analyst forecasts increasingly incorporate insights from big data and AI tools. Further, the disagreement tends to be high before economic crises, when high investor sentiment may exert a disproportional influence on analysts.

[Insert Figure 5 here]

We next examine the relative performance of Man vs. Machine precisely when they disagree to a large degree. Gaining an understanding into such situations has significant implications for AI-guided decision making including investment. For each pair of forecasts, we define an indicator variable, *Disagreement*, to be one if the magnitude of the disagreement between the analyst and our AI model is above the 90th percentile among all forecasts on the same firm over the past three years. Such benchmarking ensures that the disagreement could be measured on a similar scale. Conditional on the existence of a *Disagreement*, we further define two sub-indicators, *Machine wins* and *Man wins*, depending on which side has a lower absolute prediction error. We then relate these outcome variables to the set of regressors, with results reported in Table 5. Because the regressions involve high-dimensional fixed effects, we apply the linear probability model.

[Insert Table 5 here]

The first two columns of Table 5 examine the relation between the occurrence of *Disagreement* and the underlying firm and analyst attributes and economic conditions, with firm fixed effects (column (1)) or double firm/analyst fixed effects (column (2)). Overall results indicate that large disagreement is more likely when either the human or machine is expected to have a clear advantage (as discussed in Table 4). For example, the analyst and AI disagree when AI is expected to outperform: firm information is voluminous (higher number of 8-K reports), or analysts do not enjoy the resources of large brokerage houses. They also disagree more strongly when the human is expected to do better: high intangible assets, or more fluid product market competition. While these cases are pooled in columns (1) and (2), they become separated in columns (3) to (6) where all disagreement cases are bifurcated into "Machine wins" and "Man wins" in columns (3) to (6),[25] and we observe some additional interesting results. For example, more 8-K reports are associated with significantly more "Machine wins" but no fewer "Man wins" in disagreement (both are relative to the base state of no disagreement). In contrast, institutional ownership significantly boosts the chance of "Man wins" but does not lower "Machine wins." The asymmetric effects echo the sources of relative human or machine advantages examined in Table 4.

5. Man + Machine: Combining Strengths and Incremental Contributions

5.1. Incremental Value of Analysts in Forecasts Made by Man + Machine

Acknowledging that Man + Machine is superior to either the human or machine alone, it is still instructive to understand the respective incremental values of the human and

[25]In the "Machine wins" regressions, the observations of "Man wins" in disagreement are excluded because they are not part of the control sample. The same for the "Man wins" regressions. A nested multinomial regression could test all parallel states in a full sample, but such a regression is not able to incorporate high-dimensional fixed effects.

the machine in the combination. Analogous to the previous section, we define relative performance measures of the hybrid analyst vs the AI to capture the incremental value of humans. We then reestimate Equation (3) with these relative performance measures as dependent variables. Table 6 presents the results.

[Insert Table 6 here]

Similar to the previous findings, we find inputs from analysts are more valuable when covering firms that are more illiquid and firms with more intangible assets and earnings volatility. Moreover, analyst inputs have more incremental value when they are stars and when firms have higher distress-risk exposure (book leverage). The institutional holding percentage also helps the hybrid model beat AI analyst.

We note that while analysts in larger brokerage firms perform better than other analysts when pitted against the AI (Table 4), that advantage does not hold versus the Man + Machine model. Such a contrast in the results highlight that democratizing AI technology levels the playground: when we let all analysts (from large and small brokerage houses) be equipped with AI assistance in the Man + Machine model, disparity in institutional resources does not significantly affect the incremental value of human inputs.

5.2. Can Man + Machine Avoid Extreme Error?

As in many other skilled professions, extreme forecast errors could be calamitous to the reputation of the forecasters and to the welfare of the recipients of investment advice. However, as the common saying “to err is human; to forgive is divine” goes, machine errors are far less tolerated than human mistakes (Prahl and Swol, 2017). We are thus interested in the resilience of Man + Machine against extreme errors, a quality which would be crucial for the future of the combination, in addition to its superior average forecast accuracy.

To set the stage, we benchmark the forecast error of each forecast to the 90th (or 75th, as a sensitivity check) percentile of squared prediction errors from all analysts on the same firm over past three years. Such a setup leads to four outcomes with regard to who commit(s) an extreme error: (1) both the analyst and the AI model ("Both"); (2) Analyst; (3) AI; and (4) neither commits an extreme error ("Neither"). We examine these four scenarios and compute their empirical frequencies.[26] We then compute the unconditional and conditional probabilities that the Man + Machine model can avoid the extreme error committed in the first three scenarios and, equally importantly, the probability that Man + Machine creates an extreme error in the fourth scenario. All probabilities are reported in Table 7.

[Insert Table 7 here]

We discover that the analyst and the AI are about equally likely to make extreme errors (10.8% and 12.4% using the 90th percentile threshold).[27] There is a further probability of 2.5% that both make lousy forecasts. It turns out that the Man + Machine model can help avoid 76.3% of extreme errors made by human and 32.2% of those by AI. Even when both analysts and AI seem to be out of the ballpark, their combination still manages to bring 15.1% of such cases back to a reasonable range. Furthermore, Man + Machine only creates its own extreme error in 1.3% of the "Neither" scenario. The overall results present a significant complementary benefit of combining human and AI capabilities.

5.3. Impact of Man + Machine: An Event Study

In this section, we resort to an event study to sharpen the inference of the impact of integrating man and machine in stock analyses. In recent years, the infrastructure of "big data" has created a new class of information about companies that is collected and published outside of the firms and which can provide unique and timely clues into market demand,

[26]These four cases are not mutually disjoint, as the "Analyst" (scenario 2) and "AI" (scenario 3) cases both include the "Both" cases (scenario 1). We adopt this convention to evaluate how the Man + Machine model performs in terms of avoiding extreme errors relative to Man/Machine, independent of the counterparty's performance. Untabulated, we also conduct the same analysis for four disjoint scenarios, i.e., "Both," "Analyst Only," "AI Only," and "Neither," and find qualitatively similar results; in fact, the Man + Machine model corrects an even greater fraction of extreme errors committed by analysts alone.

[27]A sensitivity analysis using the 75th percentile yields qualitatively similar results.

profit prospects, and investment opportunities. An important and popular type of such alternative data captures "consumer footprints," often times in the literal sense such as satellite images of retail parking lots. Such data, which have to be processed by machine learning models, have been shown to contain incremental information for earnings and stock prices conditional on corporate disclosure and news coverage (Zhu, 2019; Katona, Painter, Patatoukas, and Zeng, 2022). Chi, Hwang, and Zhang (2022) show that analysts who use alternative data more frequently have more precise forecasts.

We build on data from Katona, Painter, Patatoukas, and Zeng (2022) on the staggered introduction of several important alternative data bases, and conduct a difference-in-differences test of analysts' performance versus our AI model before and after the availability of the alternative data on specific firms. The underlying premise is that analysts who cover firms that are served by the alternative data are potentially in the situation of Man + Machine, as they have the opportunity to use the additional, AI-processed, information. We define two variables based on the staggered introduction of alternative data coverage. The first is *Alt Data Covered*, which is one if satellite imaging data are available for the firm at any point in our sample period (based on the list of covered firms and coverage start dates in Table A1 in Katona, Painter, Patatoukas, and Zeng, 2022), and if the firm is in an industry with a retail footprint,[28] and zero otherwise. The second variable is *Post*, which is an indicator variable that is one if satellite data are currently available (based on coverage start dates in Table A1 in Katona, Painter, Patatoukas, and Zeng, 2022), or if the firm is not listed in that table but the date is after 2014,[29] and zero otherwise. In our analysis, a firm is "treated" by the alternative data if it is an *Alt Data Covered* firm and the time is *Post* the coverage. In the panel, we define a firm(i)-analyst(j)-year(t) triple to be an observation in the "treated" status if alternative data about firm i became available prior to year t. The rest of the

[28]We define industries with retail footprints to be those that rely mainly on retail traffic, such as the entertainment, healthcare, personal services, retail, restaurant, and hotel industries. Specifically, these include industries 6, 7, 11, 33, 40, 42, 43, 44, 45, and 46 in the Fama-French 48-industry classification.

[29]Based on anecdotal evidence from news and discussion with industry experts, 2014 is the year most alternative data became widely available.

observations are in the control subsample. Moreover, we only include an observation if the brokerage house with which analyst j is affiliated is covered by the Burning Glass job posting data any time during $[t-3, t]$.[30]

Alternative data tend to be large in volume and unstructured. Such data are hard to process with traditional tool kits. Commercial data vendors may preprocess the alternative data; for example, by converting satellite imaging data into car counts for each business location. However, substantial additional analysis is still needed to render such data useful for stock analysis. Whether analysts covering the alternative data "treated" firms could capitalize on the novel information source depends on the AI resources in their workplace. We measure AI resources that analysts have access to by the variable *AI Hiring*, which is the ratio of the number of AI jobs to the total number of job postings using the Burning Glass U.S. job posting data and following the classification algorithm developed in Babina, Fedyk, He, and Hodson (2020).

We estimate the following difference-in-differences model,

$$\begin{aligned} Analyst\ Beats\ AI_{i,j,t} = &\ \beta_1 Treat_{i,t} \times AI\ Hiring_{j,t} \\ &+ \beta_2 AI\ Hiring_{j,t} + \beta_3 Alt\ Data\ Covered_i \\ &+ \beta_4 Treat + Controls_{i,j,t} + \alpha_i + \alpha_{year} + \epsilon_{i,j,t}. \end{aligned} \tag{4}$$

Here $Treat_{i,t} = Alt\ Data\ Covered_i \times Post_{i,t}$. Note that *Alt Data Covered* and *Post* are indexed by firm i and date t while *AI Hiring* is indexed by the analyst j (or the brokerage firm associated with the analyst) and date t. Table 8 reports the results. The sample here is smaller than those in Tables 4 and 6 due to the requirement that the *AI Hiring* be observable.

[Insert Table 8 here]

Columns (1) and (2) of Table 8 show that analysts associated with brokerage houses with greater AI capabilities generally perform better against our AI model, a piece of direct evi-

[30] The reason for this restriction is to ensure that the information about AI hiring is reasonably accurate, as we cannot infer AI hiring in case of missing data.

dence that humans complemented by AIs enjoy a step up in predictive capabilities. Columns (3) and (4) show that post alternative data, analysts covering affected firms improve their performance relative to the AI model, but only significantly so when interacting with *AI Hiring*.[31] In other words, the improvement of predictive performance post alternative data concentrates in the subset of analysts who are affiliated with brokerage firms with strong AI capabilities. Overall results suggest that augmenting humans with new technologies constitutes a promising direction for the analyst profession.

6. Concluding Remarks

In this paper, we build an AI analyst to digest corporate disclosure and other information (qualitative and quantitative), and to perform forecast tasks similar to those of stock analysts. Our AI analyst is able to beat the majority of human analysts in stock forecasts. A portfolio following the difference between AI and analyst forecasts generates a monthly abnormal return of more than 75 basis points. In the contest of "Man vs. Machine," we find that the relative advantage of such an AI analyst is stronger when information is more transparent and voluminous. Human analysts remain competitive when critical information requires institutional knowledge (such as the nature of intangible assets). The edge of the AI analyst over human analysts declines over time, especially when analysts gain access to alternative data and to in-house AI resources. Combining AI and the art of human experts produces the highest potential in generating accurate forecasts in settings wherein the two skills are complementary, suggesting a future of "Man + Machine" in high-skill professions.

The complementarity between humans and machines documented in this study also provides guidance about how humans can survive and thrive in the age of machines. For example, reforming education and professional training to strengthen soft skills and creativity can help human professionals to better prepare for the incoming future.

[31] In these specifications, we do not simultaneously control for firm and analyst fixed effects due to insufficient variation in the pairing during the few years around the event

References

Abarbanell, Jeffrey S., 1991, Do analysts' earnings forecasts incorporate in-formation in prior stock price changes?, *Journal of Accounting and Economics* 14, 147–165.

Abarbanell, Jeffery, and Reuven Lehavy, 2003, Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts earnings forecasts, *Journal of Accounting and Economics* 36, 105–146.

Abis, Simona, 2022, Man vs. machine: Quantitative and discretionary equity management, Working paper.

Abis, Simona, and Laura Veldkamp, 2022, The changing economics of knowledge production, Working paper.

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo, 2022, Artificial intelligence and jobs: Evidence from online vacancies, *Journal of Labor Economics*, 40, 293–340.

Acemoglu, Daron, and Pascual Restrepo, 2018, The race between man and machine: Implications of technology for growth, factor shares and employment, *American Economic Review* 108, 1488–1542.

Acemoglu, Daron, and Pascual Restrepo, 2019, Automation and new tasks: How technology displaces and reinstates labor, *Journal of Economic Perspectives* 33, 3–30.

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones, 2019, Artificial intelligence and economic growth, *The Economics of Artificial Intelligence: An Agenda* (University of Chicago Press, Chicago).

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, 2018, *Prediction Machines: The Simple Economics of Artificial Intelligence* (Harvard Business Press, Boston).

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, 2019a, *The Economics of Artificial Intelligence: An Agenda* (University of Chicago Press, Chicago).

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, 2019b, Artificial intelligence: The ambiguous labor market impact of automating prediction, *Journal of Economic Perspectives* 33, 31–50.

Altinkilic, Oya, Robert S. Hansen, and Liyu Ye, 2016 Can analysts pick stocks for the long-run?, *Journal of Financial Economics* 119, 371–398.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-Section and time series effects, *Journal of Financial Markets* 5, 31–56.

Armour, John, Richard Parnham, and Mari Sato, 2020, Augmented lawyering, Working paper.

Aubry, Mathieu, Roman Kraeussl, Gustavo Manso, and Christophe Spaenjers, 2022, Biased auctioneers, *Journal of Finance*, forthcoming.

Autor, David H., Frank Levy, and Richard J. Murnane, 2003, The Skill content of recent technological change: An empirical exploration, *The Quarterly Journal of Economics*, 118, 1279–1333.

Babina, Tania, 2020, Destructive creation at work: How financial distress spurs entrepreneurship, *Review of Financial Studies* 33, 4061–4101.

Babina, Tania, Anastassia Fedyk, Alex Xi He, and James Hodson, 2020, Artificial intelligence, firm growth, and industry concentration, Working paper.

Barbee, William C., Sandip Mukherji, and Gary A. Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size? *Financial Analysts Journal* 52, 56 60.

van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira, 2022, Man vs. machine learning: The term structure of earnings expectations and conditional biases, *Review of Financial Studies* forthcoming.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129–156.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507–528.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* 62, 877–915.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan, 2006, The relation between corporate financing activities, analysts' forecasts and stock returns, *Journal of Accounting and Economics* 42, 53–85.

Brav, Alon, and Reuven Lehavy, 2003, An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics, *Journal of Finance* 58, 1933–1967.

Breiman, Leo, Random forests, 2001, *Machine Learning* 45, 5–32.

Brogaard, Jonathan and Matthew C. Ringgenberg and Dominik Rösch, 2021, Does Floor Trading Matter? Working Paper.

Brogaard, Jonathan and Abalfazl Zareei, 2022, Machine learning and the stock market, *Journal of Financial and Quantitative Analysis*, forthcoming.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock, 2018, What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings* 108, 43–47.

Campbell, John Y., Jens Dietrich Hilscher, and Jan Szilagyi, 2008 In search of distress risk, *Journal of Finance* 63, 2899–2939.

Cao, Sean, Lin William Cong, and Baozhong Yang, 2019, Financial reporting and blockchains: Audit pricing, misstatements, and regulation, Working paper.

Cao, Sean, Wei Jiang, Baozhong Yang, and Alan L. Zhang, 2022, How to talk when a machine is listening: Corporate disclosure in the age of AI, NBER Working paper 27950.

Cao, Sean, Yongtae Kim, Angie Wang, and Houping Xiao, 2020, Power of deep learning: Quantifying language to explain cross-sectional returns, Working paper.

Cao, Kai, and Haifeng You, 2021, Fundamental analysis via machine learning, Working paper.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2001, The stock market valuation of research and development expenditures, *Journal of Finance* 56, 2431–2456.

Chen, Qi, Jennifer Francis, and Wei Jiang, 2005, Investor Learning about Analyst Predictive Ability, *Journal of Accounting and Economics* 39, 3–24.

Chen, Luyang, Markus Pelger, and Jason Zhu, 2022, Deep learning in asset pricing, *Management Science*, forthcoming.

Chen, Mark A., Qinxi Wu, and Baozhong Yang, 2019, How valuable is FinTech innovation? *Review of Financial Studies* 32, 2062–2106.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2014, An alternative three-factor model, working paper.

Chen, Nai-Fu, Richard Roll, and Stephen Ross, 1986, Economic forces and the stock market, *The Journal of Business* 59, 383–403.

Chi, Feng, Byoung-Hyoun Hwang, and Yaping Zheng, 2022, The use and usefulness of big data in finance: Evidence from financial analysts, Working paper.

Coleman, Braiden, Kenneth J. Merkley, and Joseph Pacelli, 2022, Human versus machine: A comparison of robo-analyst and traditional research analyst investment recommendations, The Accounting Review, Forthcoming.

Cong, Lin William, Tengyuan Liang, Baozhong Yang, and Xiao Zhang, 2020, Analyzing textual information at scale, in Kashi R. Balachandran, ed.: *Information to Facilitate Efficient Decision Making: Big Data, Blockchain and Relevance* (World Scientific Publishers, Singapore).

Cong, Lin William, Ke Tang, Jingyuan Wang, and Yang Zhang, 2022, AlphaPortfolio: Direct construction through reinforcement learning and interpretable AI, Working paper.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Crane, Alan, and Kevin Crotty, 2020, How skilled are security analysts, *Journal of Finance* 75, 1629–1675.

D'Acunto, Francesco Nagpurnanand Prabhala, and Alberto G Ross, 2019, The promises and pitfalls of robo-advising, *The Review of Financial Studies* 32, 1983–2020.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

De Bondt, Werner, and Richard Thaler, 1990, Do security analysts overreact?, *American Economic Review*, 80, 52–57.

Desai, Hemang, Bing Liang, and Ajai K. Singh, 2019, Do All-Stars Shine? Evaluation of Analyst Recommendations. *Financial Analysts Journal* 56, 20–29.

Doyle, Jeffrey T., Jared N. Jennings, and Mark T. Soliman, 2013, Do managers define non-GAAP earnings to meet or beat analyst forecasts? *Journal of Accounting and Economics* 56, 40–56.

Eisfeldt, Andrea L, and Dimitris Papanikolaou, 2013, Organization capital and the cross-section of expected returns, *Journal of Finance* 68, 1365–1406.

Eisfeldt, Andrea L, and Dimitris Papanikolaou, 2014, The value and ownership of intangible capital, *American Economic Review* 104, 189–194.

Ewens, Michael, Ryan H. Peters, and Sean Wang, 2022, Measuring intangible capital with market prices, NBER Working paper 25960.

Falato, Antonio, Dalida Kadyrzhanova, and Jae Sim, 2013, Rising intangible capital, shrinking debt capacity, and the US corporate savings glut, *Journal of Finance* forthcoming.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Goldstein, Itay, Chester S. Spatt, and Mao Ye, 2021, Big data in finance, *Review of Financial Studies* 34, 3213–3225.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville, 2016, *Deep Learning*, (MIT Press, Cambridge, MA).

Grennan, Jillian, and Roni Michaely, 2020, Artificial intelligence and high-skilled work: Evidence from analysts, Working paper.

Grennan, Jillian, and Roni Michaely, 2021, FinTechs and the market for financial analysis, *Journal of Financial and Quantitative Analysis* 56, 1877–1907.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *Review of Financial Studies* 33, 2223–2273.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2019, Dynamic interpretation of emerging risks in the financial sector, *Review of Financial Studies* 32, 4543–4603.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, 2009, *Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, New York).

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38, 297–331.

Hirshleifer, David, Yaron, Levi, Ben, Lourie, Siew, and Hong Teoh, 2019, Decision fatigue and heuristic analyst forecasts, *Journal of Financial Economics*, 133, 83–98.

Hoberg, Gerard, Gordon Phillips, and Nagpurnanand Prabhala, 2014, Product market threats, payouts, and financial flexibility, *Journal of Finance* 69, 293–324.

Hochreiter, Sepp, and Jürgen Schmidhuber, 1997, Long short-term memory, *Neural Computation* 9, 1735–1780.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019–2133.

James, Gareth, Daniela Witten, Trevor Hastie, and Rob Tibshirani, 2013, *An Introduction to Statistical Learning: with Applications in R* (Springer, New York).

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D. Krische, and Charles M. C. Lee, 2004, Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59, 1083–1124.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jansen, Mark, Hieu Nguyen, and Amin Sham, 2021, Rise of the Machines: The Impact of Automated Underwriting, Working paper.

Jiang, Wei, Yuehua Tang, Rachael Jiqiu Xiao, and Vincent Yao, 2022, Surviving the FinTech disruption, NBER Working paper 28668.

Katona, Zsolt, Marcus Painter, Panos N. Patatoukas, and Jean Zeng, 2022, On the capital market consequences of alternative data: Evidence from outer space, Working paper.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lamont, Owen, Christopher Polk, and Jesús Saá-Requejo, 2001, Financial constraints and stock returns, *Review of Financial Studies* 14, 529–554.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, 2015, Deep learning, *Nature* 521, 436–444.

Li, Xi, 2005, The persistence of relative performance in stock recommendations of sell-side financial analysts, *Journal of Accounting and Economics* 40, 129–152.

Li, Feng, Russell Lundholm, and Michael Minnis, 2013, A measure of competition based on 10-K filings, *Journal of Accounting Research* 51, 399–436.

Lim, Terence, 2002, Rationality and analyst forecast bias, *Journal of Finance* 56, 369–385.

Liu, Miao, 2022, Assessing human information processing in lending decisions: A machine learning approach, *Journal of Accounting Research* 60, 607–651.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Loughran, Tim, and Ritter, Jay R., 1991, The new issues puzzle, *Journal of Finance* 50, 23–51.

Loughran, Tim, and Jay W. Wellman, 2011, New evidence on the relation between the enterprise multiple and average stock returns, *Journal of Financial and Quantitative Analysis* 46, 1629–1650.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008, The new issues puzzle: Testing the investment-based explanation, *Review of Financial Studies* 21, 2825–2855.

Martin, Ian W. R., and Stefan Nagel, 2022, Market efficiency in the age of big data, *Journal of Financial Economics* 145, 154–177.

Michaely, Roni, and Kent L. Womack, 1999, Conflict of interest and the credibility of underwriter analyst recommendations, *The Review of Financial Studies*, 12, 653–686.

Mikhail, Michael, Beverly Walther, and Richard Willis, 2007, When security analysts talk, who listens? *The Accounting Review* 82, 1227–1253.

Novy-Marx, Robert, 2011, Operating leverage, *Review of Finance* 15, 103–134.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Muro, Mark, Robert Maxim, and Jacob Whiton, 2019, *Automation and Artificial Intelligence: How Machines are Affecting People and Places*, Brookings.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109–131.

Ortiz-Molina, Hernan, and Gordon Phillips, 2014, Real asset illiquidity and the cost of capital, *Journal of Financial and Quantitative Analysis* 49, 1–32.

Pagliaro, Cynthia, Tarun Ramadorai, Alberto G. Rossi, Stephen Utkus, and Ansgar Walther, 2022, Algorithm aversion: Theory and evidence from robo-advice, Working Paper.

Peters, Ryan H, and Lucian A Taylor, 2017, Intangible capital and the investment-q relation, *Journal of Financial Economics* 123, 251–272.

Prahl, Andrew, and Lyn Van Swol, 2017, Understanding algorithm aversion: when is advice from automation discounted? *Journal of Forecasting* 36, 691–702.

Ray, Debraj, and Dilip Mookherjee, 2021, Growth, automation and the long run share of labor, *Review of Economic Dynamics* forthcoming.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman, and Irem Tuna, 2005, Accrual reliability, earnings persistence and stock prices, *Journal of Accounting and Economics* 39, 437–485.

Ritter, Jay R., 1991, The Long-run performance of initial public offerings, *Journal of Finance* 46, 3–27.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–16.

Rossi, Alberto G., and Stephen P. Utkus, 2021, Who benefits from robo-advising? Evidence from machine learning, Working paper.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289–315.

de Silva, Tim, and David Thesmar, 2021, Noise in expectations: Evidence from analyst forecasts, NBER Working paper 28963.

So, Eric C., 2013, A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts, *Journal of Financial Economics* 108, 615–640.

Soliman, Mark T., 2008, The use of DuPont analysis by market participants, *The Accounting Review* 83, 823–853.

Stickel, Scott, Predicting individual analyst earnings forecasts, *Journal of Accounting Research*, 28, 409–417.

Stickel, Scott, 1991, Common stock returns surrounding earnings forecast revisions: More puzzling evidence, *The Accounting Review* 66, 402–416.

Stickel, Scott, 1992, Reputation and Performance Among Security Analysts, *Journal of Finance* 47, 1811–1836.

Thomas, Jacob K., and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, K. C. John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *The Journal of Financial and Quantitative Analysis* 39, 677–700.

Webb, Michael, 2020, The impact of artificial intelligence on the labor market, Working paper.

Xing, Yuhang, 2008, Interpreting the value effect through the Q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Zheng, Xiang, 2022, How can innovation screening be improved? A machine learning analysis with economic consequences for firm performance, Working paper.

Zhu, Christina, 2019, Big data as a governance mechanism, *The Review of Financial Studies* 32, 2021–2061.

Zou, Hui, and Trevor Hastie, 2005, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society Series B.* 67, 301–320.

Figure 1: Man vs. Machine: The Performance of Analysts vs. AI

This figure plots the beat ratio, or the proportion of analysts' price forecasts that are more accurate than the corresponding AI price forecasts in each year. The blue line in the middle plots the annual beat ratios, and the surrounding blue-dotted lines indicate the 95% confidence interval of the beat ratio. The red line gives the best linear approximation of the time-series trend in beat ratios. The shaded grey bars represent the NBER recessions.

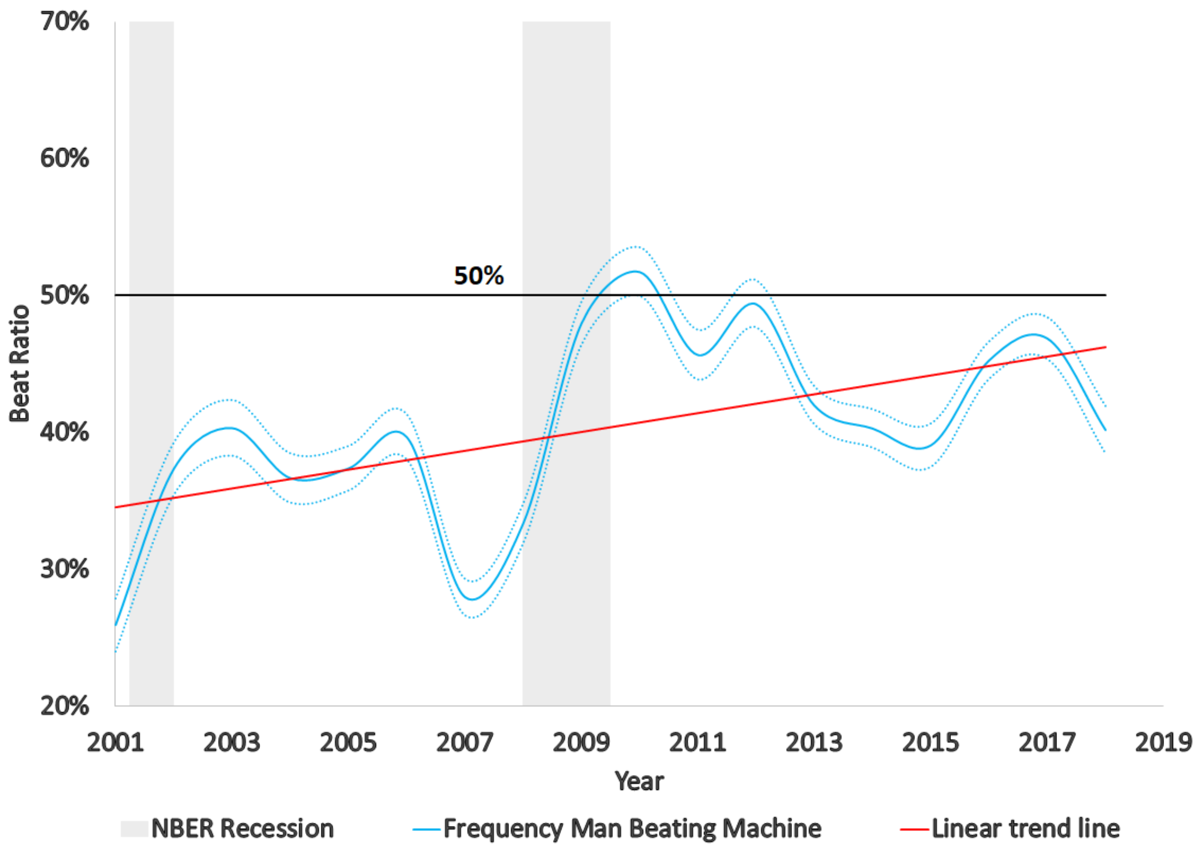

Figure 2: Contribution of Groups of Variables to the AI Prediction

This figure plots the contribution of each group of features to the AI model's price prediction. The features are divided into six groups: Firm Returns (past returns of stocks), Firm Characteristics, Earnings (past firm and industry earnings), Industry Variables, Macro Variables, and Textual Variables. The contribution from each group is the difference in forecast performance between the full information model and one that omits the given group. Specifically, we compute the percentage of times that the AI model with complete information beats the same model but without the given set of information; a probability of 50% indicates no difference. We then scale the percentages (in excess of 50%) representing the incremental effect of each group by the total sum so that they sum up to unit.

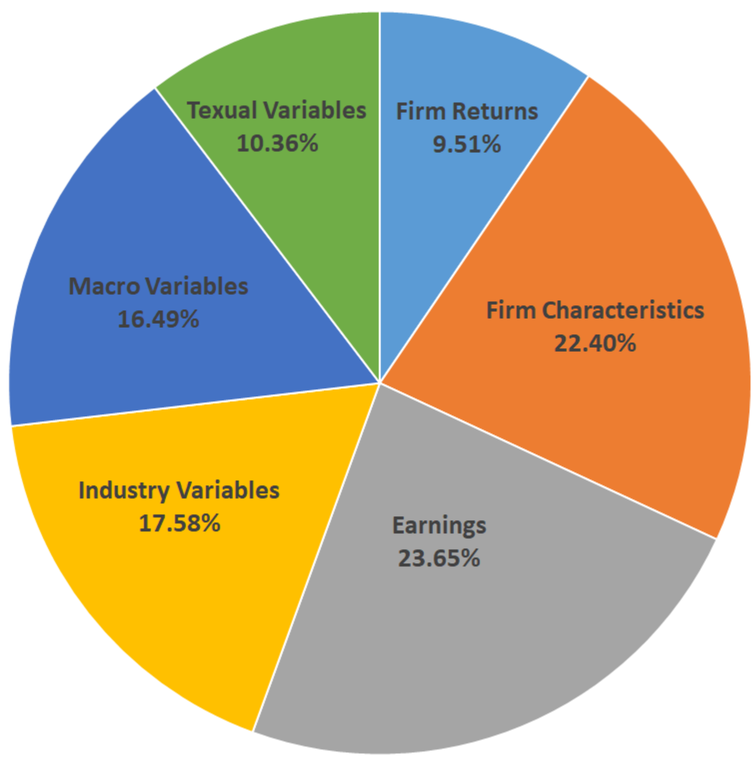

Figure 3: Man + Machine: The Performance of Machine-Debiased Analyst vs. AI

This figure plots the proportion of machine-debiased analyst (MDM) price forecasts that are more accurate than the machine recommendations alone on an annual basis. The blue line in the middle gives the annual machine-debiased analyst beat ratios, the blue-dotted lines above and below are the 95% confidence interval of the beat ratio, the green line represents the analyst beat ratios, and the red line gives the best linear approximation of the trend in beat ratios. The shaded grey bars represent the NBER recessions.

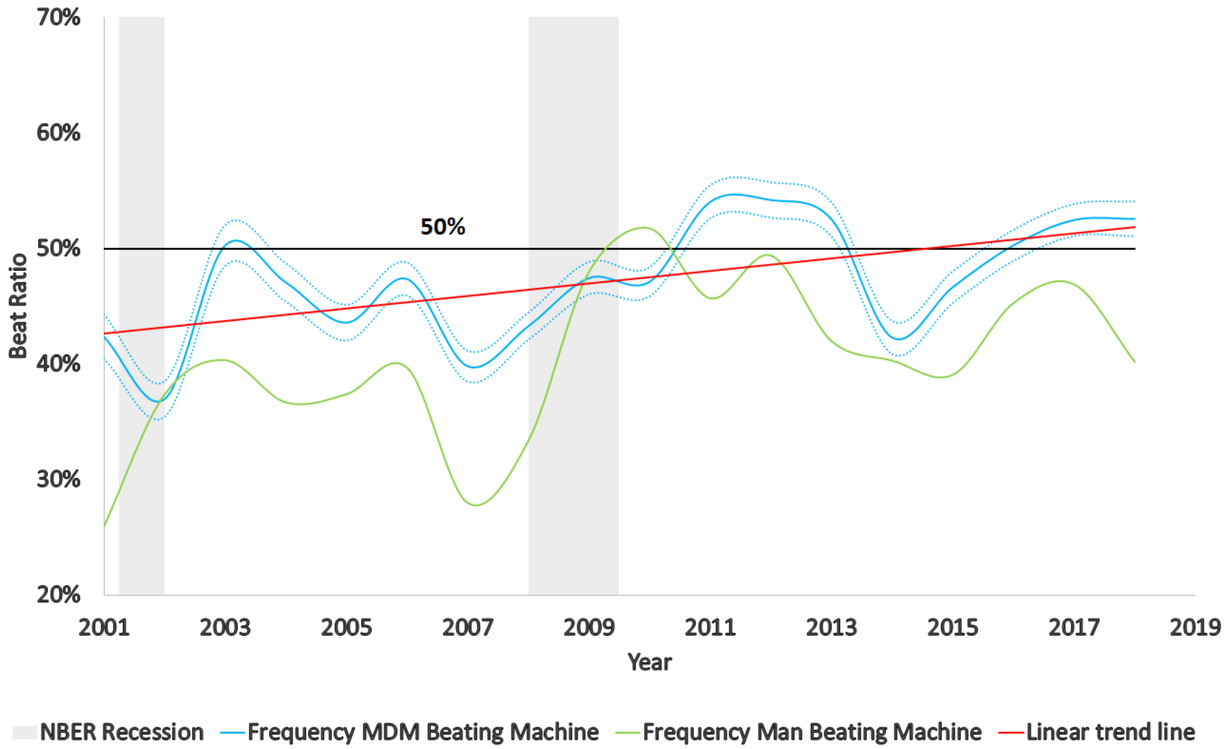

Figure 4: Man + Machine: The Performance of AI-assisted Analysts vs. AI

This figure plots the proportion of AI-assisted analyst price forecasts that are more accurate than the AI recommendations alone on an annual basis, or the "beat ratio." The blue line in the middle gives the annual AI-assisted analyst beat ratios, the blue-dotted lines above and below are the 95% confidence interval of the beat ratio, and the red line gives the best linear approximation of the trend in beat ratios. The shaded grey bars represent the NBER recessions.

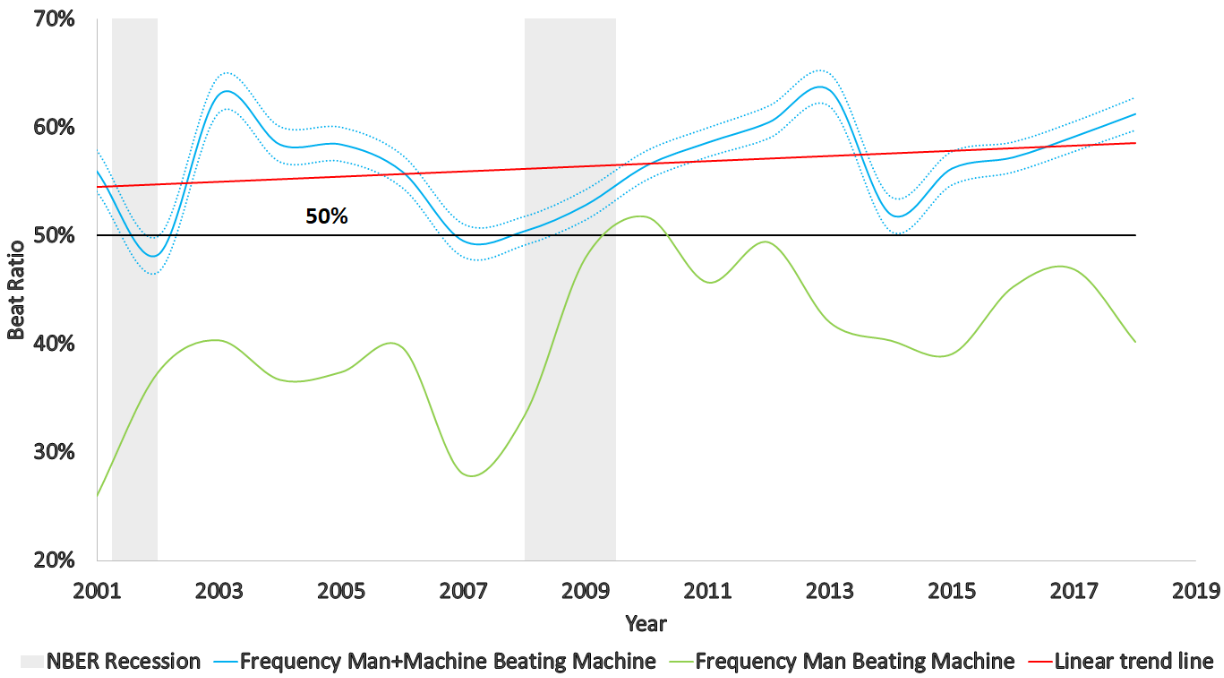

Figure 5: Man vs. Machine: Disagreement

This figure plots the disagreement between man and machine. The disagreement is defined as the squared difference between the log prices predicted by the analysts and the AI. Each year, the average value of the disagreement is calculated. The blue line in the middle gives this average disagreement, the blue-dotted lines above and below are the 95% confidence interval of the disagreement, and the red line gives the best linear approximation of the trend. The shaded grey bars represent the NBER recessions.

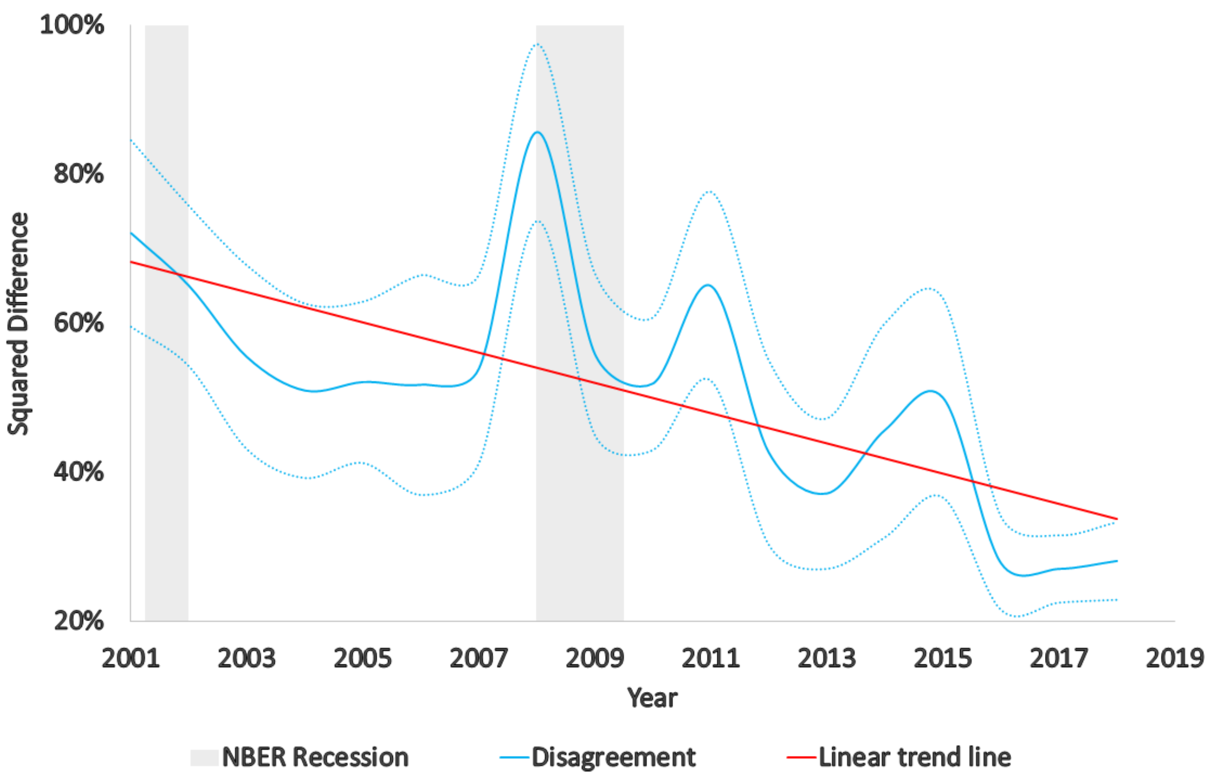

Table 1: Summary Statistics

This table reports the summary statistics of key variables. The firm-level, industry-level, and macroeconomic variables are defined in Section 2.5. The AI and alternative data variables *AI Hiring*, *Alt Data Covered*, and *Post* are defined as follows. *AI Hiring* is the ratio of the number of AI jobs to the total number of job postings. *Alt Data Covered* is an indicator variable equal to one if alternative data are available for the firm by the end of the sample. *Post* is an indicator variable equal to one if a "treated" firm has been covered by alternative data by the given year. For "untreated" firms, *Post* is coded one if the year is after 2014. The mean, median, standard deviation, 25 percentile, 75 percentile, and number of observations are reported in the table.

Variables	Mean	Median	Std	P25	P75	N
Panel A. Firm-level, industry-level, and macroeconomic variables						
Amihud Illiquidity	0.44	0.01	61.55	0.00	0.02	291,331
Market Cap	8.00	7.91	1.67	6.81	9.14	291,331
Standard Deviation of Earnings	0.19	0.11	0.30	0.06	0.22	291,331
% Institutional Holdings	0.66	0.77	2.03	0.53	0.90	291,331
# 8K Reports	2.77	1.00	15.62	0.00	2.00	291,331
Intangible Assets	0.02	-0.14	1.14	-0.42	0.14	291,331
Fluidity	7.16	6.46	3.60	4.52	9.10	291,331
# Analysts in Brokerage Firm	3.45	3.53	0.91	2.89	4.11	291,331
Star Analyst	0.47	0.50	0.21	0.34	0.60	291,331
Book Leverage	1.28	1.08	0.58	0.91	1.39	291,331
Panel B. AI and alternative data variables						
AI Hiring	0.43	0.00	4.73	0.00	0.00	51,469
Alt Data Cover	0.03	0.00	0.17	0.00	0.00	51,469
Post	0.45	0.00	0.50	0.00	1.00	51,469

Table 2: Persistence of Performance of AI Analyst

Each year, analysts are sorted by mean squared prediction errors of log prices based on the past one, past two, and up to five years. If the mean squared error over the last year is below (above) the median during the specified past period, the analyst is in the top (bottom) in the current year. In Panel A, the sorting is based on the full period of the past one, two, . . ., five years. In Panel B, the sorting requires that an analyst be in the top half in each of the past one, two, . . ., five years to be placed in the "top" group. Both panels report the analyst beat ratio, i.e., the number of times analysts beat AI, as a proportion of total number of predictions.

Panel A: Analyst beat ratio sorted by analysts who are above/below median

	1 year	2 years	3 years	4 years	5 years
Analyst top	47.93%	47.77%	47.68%	47.63%	47.60%
Analyst bottom	41.60%	41.69%	41.74%	41.79%	41.82%

Panel B: Analyst beat ratio sorted by analysts who are above median each of the past years

	1 years	2 years	3 years	4 years	5 years
Analyst Persistent top	47.93%	48.66%	49.33%	49.42%	49.88%
Analyst Persistent bottom	41.60%	41.42%	41.09%	40.94%	41.05%

Table 3: Portfolio Performance following Machine vs. Man Recommendations: Monthly Rebalancing

In each month, we gather all predictions made by all analysts and the corresponding AI forecasts in the past 30, 60, 90, and 360 days. For each pair of predictions, if the AI predicts a higher (lower) price than the analyst, it is considered as a buy (sell) signal. During the given time horizon, the portfolio will long the stock if there are more buy than sell signals, and short the stock otherwise. The portfolios are equal weighted and rebalanced monthly, i.e., a position is held for one month or till the signals reverse. The monthly percentage returns of the long-short, long-leg (stocks only with a buy sign) and short-leg portfolios (stocks only with a short sign) as well as the alphas generated from the FF3, FFC4, FF5, and FF6 models are presented. The OLS standard error is used to construct *t*-stats. The *t*-stats are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Portfolio returns – AI vs. Analyst

		AI vs. Analyst			
Long-Short		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	1.39***	1.51***	1.49***	1.05***
		(3.10)	(3.54)	(3.59)	(3.77)
	FF3	1.47***	1.62***	1.58***	1.16***
		(3.28)	(3.74)	(3.77)	(4.22)
	FFC4	1.53***	1.66***	1.61***	1.15***
		(3.43)	(3.84)	(3.86)	(4.16)
	FF5	0.96**	1.14***	1.14***	0.75***
		(2.34)	(2.89)	(2.98)	(3.20)
	FF6	1.03**	1.19***	1.18***	0.76***
		(2.52)	(3.03)	(3.11)	(3.25)
Long-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	2.02***	2.13***	2.12***	1.81***
		(4.30)	(4.56)	(4.61)	(5.07)
	FF3	1.38***	1.44***	1.40***	1.16***
		(4.71)	(4.92)	(5.00)	(9.90)
	FFC4	1.47***	1.51***	1.47***	1.21***
		(5.10)	(5.21)	(5.28)	(10.51)
	FF5	1.21***	1.28***	1.27***	1.00***
		(4.39)	(4.68)	(4.83)	(9.85)
	FF6	1.27***	1.33***	1.31***	1.03***
		(4.73)	(4.92)	(5.07)	(10.62)
Short-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.63	0.62	0.62	0.76
		(1.13)	(1.14)	(1.16)	(1.56)
	FF3	-0.09	-0.18	-0.18	0.00
		(-0.30)	(-0.60)	(-0.61)	(-0.01)
	FFC4	-0.06	-0.15	-0.14	0.06
		(-0.20)	(-0.50)	(-0.51)	(0.25)
	FF5	0.24	0.14	0.12	0.24
		(0.84)	(0.51)	(0.47)	(1.16)
	FF6	0.24	0.14	0.13	0.27
		(0.84)	(0.51)	(0.48)	(1.29)

Table 4: Man vs. Machine: The Relative Advantage of Analyst vs AI

This table presents the coefficients and t-stats of regressing the *Analyst beats AI* indicator (Panel A) and the *Forecast Error Difference: Analyst vs. AI* (Panel B) on the firm-level, industry-level, and macroeconomic variables presented in Table 1. *Analyst beats AI* is an indicator variable equal to one if the analyst beats the AI. *Forecast Error Difference: Analyst vs. AI* is defined as the difference between squared prediction errors between the AI and the analysts, divided by the maximum value of these two prediction errors. The number is positive if the analyst has smaller squared error, i.e., analyst beats AI. The t-statistics are based on standard errors clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Panel A: Analyst beats AI

Variables				
Amihud Illiquidity	0.228**	0.189***	0.033	-0.024
	(2.52)	(2.66)	(0.36)	(-0.33)
Market Cap	-0.064***	-0.037***	-0.086***	-0.053***
	(-9.90)	(-5.87)	(-12.65)	(-8.00)
Standard Deviation of Earnings	-0.050	-0.064	0.045	0.049
	(-0.44)	(-0.53)	(0.43)	(0.42)
% Institutional Holdings	0.050	0.072***	0.059*	0.075***
	(1.04)	(2.68)	(1.66)	(3.62)
# 8K Reports	-0.485***	-0.554***	-0.400***	-0.525***
	(-4.40)	(-4.88)	(-3.42)	(-4.29)
Intangible Assets	0.031***	0.029***	0.039***	0.037***
	(2.99)	(2.97)	(2.84)	(2.79)
Fluidity	0.539***	0.116	0.800***	0.374**
	(3.73)	(0.74)	(4.78)	(2.02)
# Analysts in Brokerage Firm	0.142	0.421***	-0.253	0.157
	(1.06)	(3.24)	(-0.60)	(0.37)
Star Analysts	0.455	0.486	-0.494	-0.478
	(0.99)	(1.07)	(-0.89)	(-0.87)
Book Leverage	0.064**	0.059**	0.099***	0.098***
	(2.17)	(2.54)	(5.19)	(4.71)
Time Trend	0.013***		0.009***	
	(13.64)		(7.84)	
Year Fixed Effect	No	Yes	No	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Analyst Fixed Effect	No	No	Yes	Yes
Observations	291,331	291,331	291,331	291,331
Adjusted R-squared	0.109	0.121	0.165	0.174

Panel B: Forecast Error Difference: Analyst vs. AI

Variables				
Amihud Illiquidity	0.532***	0.462***	0.333***	0.231***
	(4.41)	(5.11)	(2.83)	(2.63)
Market Cap	-0.104***	-0.061***	-0.138***	-0.087***
	(-10.16)	(-6.09)	(-13.11)	(-8.36)
Standard Deviation of Earnings	0.043	0.007	0.148	0.138
	(0.28)	(0.04)	(0.98)	(0.76)
% Institutional Holdings	0.051	0.080*	0.073	0.090***
	(0.72)	(1.96)	(1.40)	(2.89)
# 8K Reports	-0.745***	-0.853***	-0.593***	-0.802***
	(-4.32)	(-4.85)	(-3.23)	(-4.22)
Intangible Assets	0.048***	0.045***	0.064***	0.061***
	(2.95)	(2.92)	(2.98)	(2.92)
Fluidity	0.766***	0.169	1.112***	0.505*
	(3.26)	(0.65)	(4.14)	(1.69)
# Analysts in Brokerage Firm	0.096	0.538***	-0.325	0.311
	(0.47)	(2.72)	(-0.48)	(0.46)
Star Analysts	1.159*	1.197*	0.129	0.156
	(1.66)	(1.73)	(0.16)	(0.19)
Book Leverage	0.114***	0.109***	0.124***	0.125***
	(2.97)	(3.40)	(4.15)	(4.00)
Time Trend	0.021***		0.015***	
	(14.23)		(8.13)	
Year Fixed Effect	No	Yes	No	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Analyst Fixed Effect	No	No	Yes	Yes
Observations	291,331	291,331	291,331	291,331
Adjusted R-squared	0.136	0.148	0.199	0.207

Table 5: Disagreement between Man and Machine

This table presents the coefficients and t-stats of regressing the *Disagreement* indicator on the firm-level, industry-level, and macroeconomic variables presented in Table 1. For each pair of forecasts, we define the indicator variable *Disagreement* to be one if the magnitude of the squared difference between the log predicted prices of the analyst and our AI model is above the 90th percentile among all forecasts on the same firm over the past three years. Conditional on *Disagreement* being positive, we further define two sub-indicators, *Machine wins* and *Man wins*, depending on which side has lower absolute prediction error. We report regression results with these outcome variables. In the *Machine wins* regressions, the observations of *Man wins* in disagreement are excluded because they are not part of the control sample, and vice versa for the *Man wins* regressions. To calculate t-statistics, standard errors are clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Variable	*Disagreement*		*Disagreement &*		*Disagreement &*	
			Machine wins	*Man wins*	*Machine wins*	*Man wins*
Amihud Illiquidity	-0.013	-0.199***	-0.277***	0.112***	-0.537***	0.030
	(-0.48)	(-6.23)	(-7.09)	(3.57)	(-11.79)	(0.47)
Market Cap	0.001	0.002	0.004**	-0.004***	0.007***	-0.005***
	(0.34)	(0.89)	(2.51)	(-3.53)	(2.91)	(-2.99)
Standard Deviation of Earnings	-0.118	-0.130	-0.092	-0.035***	-0.112	-0.022*
	(-0.78)	(-0.48)	(-0.58)	(-4.39)	(-0.40)	(-1.80)
% Institutional Holdings	0.024***	0.060***	-0.056	0.088***	0.684	0.093***
	(3.83)	(16.73)	(-0.13)	(9.26)	(1.17)	(20.45)
# 8K Reports	0.205***	0.214***	0.223***	0.005	0.212***	0.024
	(2.90)	(2.68)	(3.49)	(0.11)	(2.96)	(0.51)
Intangible Assets	0.004*	0.001	-0.418***	0.085*	-0.532***	0.129**
	(1.84)	(0.34)	(-5.71)	(1.68)	(-6.25)	(2.19)
Fluidity	0.177***	0.214***	0.087*	0.115***	0.076	0.174***
	(3.51)	(3.03)	(1.86)	(2.69)	(1.22)	(3.14)
# Analysts in Brokerage Firm	-0.956***	-0.458*	-0.671***	-0.386***	-0.249	-0.296*
	(-10.79)	(-1.90)	(-9.21)	(-7.13)	(-1.21)	(-1.83)
Star Analysts	-0.498*	-0.269	-0.450*	-0.078	-0.275	-0.043
	(-1.69)	(-0.74)	(-1.76)	(-0.41)	(-0.88)	(-0.18)
Book Leverage	-0.026	-0.006	-0.005	-0.023	0.001	-0.007
	(-1.55)	(-0.45)	(-0.79)	(-1.57)	(0.14)	(-0.61)
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Analyst Fixed Effect	No	Yes	No	No	Yes	Yes
Observations	291,331	291,331	280,716	272,108	280,716	272,108
Adjusted R-squared	0.044	0.095	0.051	0.036	0.112	0.067

Table 6: Man + Machine: The Incremental Value of Analyst

This table presents the coefficients and t-stats of regressing the *Analyst + AI beats AI* indicator (Panel A) and *Forecast Error Difference: Analyst + AI vs. AI* (Panel B) on the firm-level, industry-level, and macroeconomic variables presented in Table 1. *Analyst + AI beats AI* is an indicator variable equal to one if Analyst + AI beats AI. *Forecast Error Difference: Analyst + AI vs. AI* is defined as the difference between squared prediction errors between AI and Analyst + AI, divided by the maximum value of these two prediction errors. The number is positive if the analyst has smaller squared error, i.e., Analyst + AI beats AI.

Panel A. Analyst + AI beats AI

Variables				
Amihud Illiquidity	0.139***	0.113***	0.072*	0.041
	(4.04)	(3.28)	(1.84)	(1.09)
Market Cap	-0.027***	-0.015***	-0.036***	-0.021***
	(-6.15)	(-3.44)	(-6.52)	(-3.72)
Standard Deviation of Earnings	0.328***	0.336***	0.322***	0.342***
	(4.17)	(4.83)	(4.15)	(5.30)
% Institutional Holdings	0.076***	0.089***	0.078***	0.088***
	(3.70)	(7.39)	(5.11)	(9.40)
# 8K Reports	-0.176*	-0.192*	-0.168	-0.187
	(-1.66)	(-1.73)	(-1.39)	(-1.48)
Intangible Assets	0.027***	0.026***	0.027***	0.026***
	(3.35)	(3.32)	(3.37)	(3.10)
Fluidity	0.319**	0.107	0.459**	0.239
	(2.04)	(0.66)	(2.43)	(1.22)
# Analysts in Brokerage Firm	-0.234**	-0.153	-0.555	-0.470
	(-1.97)	(-1.30)	(-1.37)	(-1.18)
Star Analysts	0.798*	0.915**	0.986*	1.016*
	(1.70)	(1.97)	(1.66)	(1.72)
Book Leverage	0.006	0.005	-0.023	-0.025
	(0.14)	(0.10)	(-0.48)	(-0.50)
Time Trend	0.002**		0.002**	
	(2.08)		(2.01)	
Year Fixed Effect	No	Yes	No	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Analyst Fixed Effect	No	No	Yes	Yes
Observations	291,331	291,331	291,331	291,331
Adjusted R-squared	0.042	0.047	0.065	0.069

Panel B: Forecast Error Difference: Analyst + AI vs. AI

Variables				
Amihud Illiquidity	0.193***	0.166***	0.208***	0.191***
	(5.70)	(4.96)	(6.76)	(6.24)
Market Cap	-0.020***	-0.007	-0.027***	-0.011**
	(-4.35)	(-1.58)	(-5.06)	(-2.05)
Standard Deviation of Earnings	0.201	0.204	0.260	0.276
	(1.26)	(1.33)	(1.29)	(1.43)
% Institutional Holdings	0.047*	0.060***	0.022	0.030**
	(1.80)	(3.72)	(0.94)	(1.96)
# 8K Reports	-0.272**	-0.219*	-0.256*	-0.217
	(-2.26)	(-1.72)	(-1.89)	(-1.52)
Intangible Assets	0.021**	0.020**	0.018*	0.016*
	(2.32)	(2.24)	(1.86)	(1.65)
Fluidity	0.262	0.112	0.392**	0.257
	(1.61)	(0.67)	(2.01)	(1.25)
# Analysts in Brokerage Firm	-0.223*	-0.135	-0.394	-0.349
	(-1.84)	(-1.13)	(-0.93)	(-0.84)
Star Analysts	0.992**	1.117**	1.100*	1.125*
	(2.09)	(2.37)	(1.86)	(1.92)
Book Leverage	0.059***	0.060***	0.045**	0.045**
	(2.67)	(2.61)	(2.08)	(2.11)
Time Trend	0.002***		0.003***	
	(3.02)		(2.82)	
Year Fixed Effect	No	Yes	No	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Analyst Fixed Effect	No	No	Yes	Yes
Observations	291,331	291,331	291,331	291,331
Adjusted R-squared	0.047	0.053	0.072	0.077

Table 7: Probabilities of Extreme Errors: Man vs. Machine and Man + Machine

This table presents the probabilities of extreme errors by analysts and the AI and how the Man + Machine model helps to correct such errors. We benchmark the forecast error of each forecast to the 90th (or 75th, as a sensitivity check) percentile of squared prediction errors from all analysts on the same firm over past three years. Such a setup leads to four outcomes with regard to who commit(s) an extreme error: (1) both the analyst and the AI model ("Both"); (2) Analyst; (3) AI; and (4) neither commits an extreme error ("Neither"). We examine these four scenarios and compute their empirical frequencies. We then compute the unconditional and conditional probabilities that the Man + Machine model can avoid the extreme error committed in the first three scenarios, and equally importantly, the probability that Man + Machine creates an extreme error in the fourth scenario. Panel A and B show results for extreme errors defined by the 90th percentile and 75th percentile of forecast errors, respectively.

Panel A: Probabilities of Extreme Errors (90th percentile)

	Both	Analyst	AI	Neither
Uncond. Prob.	2.49%	10.75%	12.44%	79.30%

	M+M Avoids Both	M+M Avoids Analyst	M+M Avoids AI	M+M Creates EE
Uncond. Prob.	0.37%	8.21%	4.00%	1.01%

	M+M Avoids Both/ Both	M+M Avoids Analyst/ Analyst	M+M Avoids AI/ AI	M+M Creates EE/ Neither
Conditional Prob.	15.06%	76.33%	32.18%	1.27%

Panel B: Probabilities of Extreme Errors (75th percentile)

	Both	Analyst	AI	Neither
Uncond. Prob.	8.17%	24.86%	22.89%	60.42%

	M+M Avoids Both	M+M Avoids Analyst	M+M Avoids AI	M+M Creates EE
Uncond. Prob.	1.08%	16.54%	5.62%	1.44%

	M+M Avoids Both/ Both	M+M Avoids Analyst/ Analyst	M+M Avoids AI/ AI	M+M Creates EE/ Neither
Conditional Prob.	13.15%	66.51%	24.54%	2.39%

Table 8: Man + Machine Event Study: Alternative Data Coverage

This table presents the coefficients and t-stats of regressing the *Analyst Beats AI* indicator on brokerage *AI Hiring*, *Alt Data Covered*, *Post*, and the interactions among these variables. *Analyst beats AI* is an indicator variable equal to one if the analyst beats the AI. The AI and alternative data variables *AI Hiring*, *Alt Data Covered*, and *Post* are defined as follows. *AI Hiring* is the ratio of the number of AI jobs to the total number of job postings. *Alt Data Covered* is an indicator variable equal to one if alternative data are available for the firm by the end of the sample. *Post* is an indicator variable equal to one if a "treated" firm has been covered by alternative data by the given year. For "untreated" firms, *Post* is coded one if the year is after 2014. The control variables are the firm-level, industry-level, and macroeconomic variables presented in Table 1. Standard errors are clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Variables	Analyst beats AI			
AI Hiring	0.142***	0.050	0.136***	0.044
	(2.69)	(0.80)	(2.59)	(0.69)
Alt Data Cover				-0.065
				(-1.52)
Treat: Alt Data Cover × *Post*			-0.012	0.041
			(-0.22)	(0.78)
Treat × *AI Hiring*			1.065***	1.375***
			(2.69)	(2.65)
Controls	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
Firm fixed effect	Yes	No	Yes	No
Analyst fixed effect	No	Yes	No	Yes
Observations	51,468	51,468	51,468	51,468
Adjusted R-squared	0.123	0.040	0.123	0.040

Appendix A. List of Variables

Table A1: List of All Variables Used in AI Algorithms

All Variables (and the definition/source) used in the machine learning algorithms are provided.

Firm Characteristics	Definition and/or Source
Momentum	Past 12-month return, Jegadeesh and Titman (1993)
Composite Equity Issuance	Daniel and Titman (2006)
Gross Profits-to-Assets	Novy-Marx (2013)
Asset Growth	Cooper, Gulen, and Schill (2008)
Investment-to-Assets	Titman, Wei, and Xie (2004) and Xing (2008)
Net Operating Assets	Hirshleifer, Hou, Teoh, and Zhang (2004)
Accruals	Sloan (1996)
Net Stock Issues	Ritter (1991) and Loughran and Ritter (1995)
Failure Probability	Campbell, Hilscher, and Szilagyi (2008)
O-Score	Ohlson (1980)
Return on Assets	Fama and French (2006) and Chen, Novy-Marx, and Zhang (2014)
Book-to-Market Equity	Rosenberg, Reid, and Lanstein (1985)
Debt-to-Market	Bhandari (1988)
Earnings-to-Price	Basu (1983)
Cash Flow-to-Price	Lakonishok, Shleifer, and Vishny (1994)
Payout Yield	Boudoukh, Michaely, Richardson, and Roberts (2007)
Five-year Sales Growth Rank	Lakonishok, Shleifer, and Vishny (1994)
Enterprise Multiple	Loughran and Wellman (2011)
Sales-to-Price	Barbee, Mukherji, and Raines (1996)
Abnormal Corporate Investment	Titman, Wei, and Xie (2004)
Investment-to-Assets	Cooper, Gulen, and Schill (2008)
Changes in PPE and Inventory/Assets	Lyandres, Sun, and Zhang (2008)
Investment Growth	Xing (2008)
Inventory Changes	Thomas and Zhang (2002)
Operating Accruals	Sloan (1996)
Total Accruals	Richardson, Sloan, Soliman, and Tuna (2005)
Net External Finance	Bradshaw, Richardson, and Sloan (2006)
Return on Net Operating Assets	Soliman (2008)
Profit Margin	Soliman (2008)
Asset Turnover	Soliman (2008)
Operating Profits-to-Equity	Fama and French (2015)
Book Leverage	Fama and French (1992)
Advertising Expense-to-Market	Chan, Lakonishok, and Sougiannis (2001)
R&D-to-Market	Chan, Lakonishok, and Sougiannis (2001)
Operating Leverage	Novy-Marx (2011)
Financial Constraints	Kaplan-Zingales index, Lamont, Polk, and Saá-Requejo (2001)
Asset Liquidity	Scaled by book assets, Ortiz-Molina and Phillips (2014)
Asset Liquidity	Scaled by market assets, Ortiz-Molina and Phillips (2014)
IBES Actual Earning	IBES actual earning 4 quarter before scaled by adjusted price
Number of Institutional Owners	Number of 13F institutional investors that own the stock
Ownership Concentration	Herfindahl-Hirschman Index
Total Institutional Ownership	Percent of shares outstanding owned by 13F investors

Industry Variables	Definition and/or Source
Competition Measure from 10-K	Li, Lundholm, and Minnis (2013)
Fluidity	Product market Fluidity, Hoberg, Phillips, and Prabhala (2014)
48 Industry Dummy	Dummy variables that indicate Fama-French 48 industries
Industry Size	Industry Size within past 3, 6, 9 ,12, 24 and 36 months
Industry Earning	Industry earning within past 3, 6, 9 12, 24 and 36 months
Macro Variables	Definition and/or Source
IP	Industrial Production Index
CPI	Consumer Price Index
Oil price	Crude Oil Price
Tbill3	3-month Treasury Bill
TBond10	10-Year Treasury Constant Maturity Rate
Credit Spread	Baa-AAA yield spread
Textual Variables	Definition and/or Source
Neg 10KQ	Percentage of negative words from 10K/10Q
NegPos 10KQ	Percentage of negative minus positive words from 10K/10Q
Neg 8k	Percentage of negative words from 8K
NegPos 8K	Percentage of negative minus positive from 8K
Neg Other	Percentage of negative words from other reports
NegPos Other	Percentage of negative minus positive from other reports
ML-based Sentiment	ML-based negative tones minus ML-based positive tones scaled by the length of SEC filings, Cao, Kim, Wang, and Xiao (2020)
ML-based Neg Sentiment	ML-based negative tones scaled by the length of SEC filings

Appendix B. Details of the Machine Learning Models

In this section, we briefly describe the basic structure and strengths of machine learning models considered in our paper. Interested readers are referred to representative references for more details, such as Hastie, Tibshirani, and Friedman (2009) and Goodfellow, Bengio, and Courville (2016).

B.1. Linear Models

Linear machine learning models generalize linear regressions and classification models, and are more flexible and can accommodate a larger number of variables than the traditional linear regressions, by their built-in dimension-reduction capabilities. Linear models are typically efficient in model training because they are typically associated with fast algorithms, such as linear and quadratic programming techniques.

B.1.1. Lasso and Elastic-Net

The Lasso and Elastic-Net models are generalizations of the OLS linear regression model. When there is a large number of predictors, the OLS tends to have good in-sample performance (*small bias* in the terms of machine learning) and bad out-of-sample performance (*large variation* in the terms of machine learning). Furthermore, the OLS can generate significant loadings on a large number of independent variables, making the interpretation of the model difficult. One class of models, the shrinkage models, generalize the OLS by imposing a penalty on the number and size of non-zero coefficients in the estimation, effectively limiting the model to focus on a subset of the independent variables and achieving dimension reduction.

The Elastic-Net model (Zou and Hastie (2005)), of which the Lasso model is a special case, is a shrinkage model in which the penalty function is a linear combination of L^1 and L^2 norms of the coefficients. In particular, the Elastic-Net model minimizes the following objective function,

$$\min_{\beta} \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=1}^{p} |\beta_j|^2. \tag{A1}$$

The Elastic-Net model is a generalization of the well-known Lasso and Ridge regression models. When the hyperparameter $\lambda_2 = 0$ in (A1), we recover the Lasso model. When $\lambda_1 = 0$, we recover the Ridge model. In general, the Lasso model tends to select a few strong predictors while setting the coefficients of other predictors to essentially zero, but can make random choices among several strong and correlated variables. The Ridge model usually includes more predictors and shrink the coefficients of correlated variables together. The Elastic-Net model strikes a balance between these characteristics, allowing both a selection of strong features and the averaging of correlated features.

B.1.2. Support Vector Regression

The support vector regression (SVR) is motivated by support vector machines (SVM) for classification problems. Consider a classification problem with 2 classes and n predictors, i.e., each observation belongs to one of two classes and is a point in the n-dimensional space. Given a training sample, i.e., a set of labeled n-dimensional points, a linear classifier is equivalent to a $(n-1)$-dimensional hyperplane that separates the two classes of points in the n-dimensional space. The linear SVM searches for the *maximum separating* $(n-1)$-dimensional hyperplane such that it maximizes the distance from the hyperplane to the closest data points. To deal with the case that separating hyperplanes may not exist, the SVM also tolerates misclassified points within some bounds. The SVM thus focuses on points close to the separating hyperplane, i.e., observations that are on the boundary of the two classes. In fact, the results the SVM do not depend on observations far away from the boundary. One key advantage of the SVM is that it performs well when there is a large number of features relative to the sample size, e.g., in the case of textual and image analysis.

The SVM is regarded as one of the best out-of-the-box machine learning algorithms and has been widely applied in the classification of text, image, hand-writing, and proteins.

The support vector regression optimizes the following objective function.

$$\min_{\beta} \sum_{i=1}^{N} V(y_i - \beta_0 - x_i^T \beta) + \lambda \sum_{j=1}^{p} |\beta_j|^2. \tag{A2}$$

where the cost function $V(\cdot)$ is given by

$$V(z) = \begin{cases} 0, & \text{if } |z| < \epsilon, \\ |z| - \epsilon, & \text{otherwise.} \end{cases} \tag{A3}$$

The cost function is insensitive to the signs/sizes of errors if the error size is less than ϵ, i.e., it is more sensitive to points where the estimation error is larger. These marginal points, in turn, are instrumental in determining the estimated coefficients. The benefit of the support vector regression is that it allows efficient dimension reduction even with a very large number of features. However, it may not perform well if the underlying pattern is far away from being linear.

B.2. Decision-Tree Based Models

The linear models considered above may not work well if there are nonlinear relationships among the predictive variables. In this section, we discuss a class of versatile nonlinear models – decision trees and derived models.

B.2.1. Decision Trees

Decision trees are modeled after human decisions. A decision tree is a series of binary decisions based on cutoffs of independent variables at each branching point. The tree thus will divide the rectangular feature space into smaller rectangular blocks. The decision tree regression then use the sample mean of the dependent variable in each block as the prediction for any point in the block.

Decision trees have the benefit of being easily interpretable because it is modeled after human decisions (similar to a step-by-step instructions) and can also be displayed graphically (as binary trees). Trees are also a flexible non-linear model that can model a variety of nonlinear patterns given the large degree of freedom in specifying the sequences of branching rules.

However, trees do not have a high level of accuracy by themselves because of the restrictive form of the binary branching process, which forces the sample to be split into rectangular regions and may not approximate the real underlying patterns (whether linear or nonlinear) well. Trees are also non-robust. In addition, a small change in the data can lead to large changes in the structure of the estimated tree because the tree structure is discrete, not continuous. Several methods, including random forest and gradient boosting, use trees as basic building blocks to form ensemble predictors and achieve superior performance.

B.2.2. Random Forest

A random forest (introduced by Breiman, 2001) proceeds in the following way. First, it involves drawing a bootstrapped sample (drawing with repetition) from the original sample. Second, on the bootstrapped sample, one builds a decision tree, selecting a splitting predictor among only a random m features of the total p predictors. Third, one repeats the above two steps to build a number of decision trees, and form the ensemble predictor by taking the mean predictor of all the trees.

Random forests perform better than simple trees for several reasons. First, through aggregating predictions over bootstrapped samples, it reduces the variance and non-robustness of single trees. Second, the random feature selection in the second step above ensures that the estimated trees are not too correlated, avoiding relying only on a few prominent features and further reducing the variance of the model.

B.2.3. Gradient Boost

Boosting also combines a number of weak models to generate a stronger model. In boosting of trees, a number of trees are constructed sequentially, i.e., each tree is constructed using information based on the previously constructed trees. In gradient boosting, each decision tree is fit to the residuals of the model, not to the outcome. Once a new tree is obtained, it is added to the predictive function to update it, usually with a learning weight multiplied to the tree predictor to adjust the rate of learning new information. Then new residuals are obtained from the updated predictive function and the process is repeated for a number of times to obtain the final ensemble predictor. Because boosting models aggregate results of decision trees sequentially, each component tree does not need to be very precise and can be simple, i.e., having a low depth.

In a sense, gradient boosting is similar to the Newton's gradient algorithm in optimization. It approximates the true underlying function sequentially by improving on the predicted residuals/errors gradually. This allows the final predictive function to have a much richer and more flexible structure and thus much better performance than single decision trees. It also reduces the non-robustness of single trees through using an ensemble of trees. For these reasons, gradient boosting is one of the best off-the-shelf machine learning methods.

B.3. Deep Learning Model: Long Short-Term Memory Neural Networks

The neural networks models, initial motivated by the neuron structures in the brains of humans and animals, blossomed after breakthroughs in algorithms and computing power (LeCun, Bengio, and Hinton, 2015). Neural networks models, also called deep learning models, have become some of the most powerful models and achieved near- or super-human capabilities in a wide variety of applications, such as natural language processing, speech recognition, computer vision, game playing, and autonomous driving.

There are many different architectures of neural networks, such as the simplest Feedforward Neural Networks for straightforward classification tasks, the Convolutional Neural Networks for image and pattern recognition, and Recurrent Neural Networks (RNN) that can process sequential data such as speech and text. Long Short-Term Memory (LSTM) Neural Networks are a special type of RNN that is the key to the many successes of RNN, including speech recognition, language modeling, and translation.

In a neural network, there are nodes (neurons) that are connected to each other. There are three types of nodes: input nodes that are used to receive data; output nodes that produce desired outcomes or predictions; and intermediate nodes that process the data from input nodes and convert them to outputs. The connections of the nodes determine the structure of the neural network and its features. RNNs are neural networks with loops, or nodes that are connected to themselves.

LSTM networks are introduced by Hochreiter and Schmidhuber (1997) to solve the problem that standard RNNs have trouble retaining "memory" of the much earlier parts of sequential input data, when processing the later parts of the data. Since sequential data may have long-term dependencies, i.e., parts far away in the sequence may be related, it is important to have "long-term memory" to handle them. LSTM networks have a sequence of nodes that are specifically designed to retain long-term information and update it continuously with new information in a flexible way. As a result, LSTM can capture both short-term and long-term relations in sequential or time-series data very well, suggesting its potential applications in financial economics given the abundance of time-series financial data.

Internet Appendix of "From Man vs. Machine to Man + Machine: The Art and AI of Stock Analyses"

Table IA1: Performance Evaluation of Machine Learning Models

We evaluate the performance of a variety of machine learning models in predicting stock prices. We use rolling windows to train and evaluate the models as in Section 3.1. Detailed discussions of the machine learning models can be found in Appendix B. This table presents the average squared prediction errors for each model and each year of our sample. Specifically, in the rolling window prior to a given year, similar to Gu, Kelly, and Xiu (2020), we allocate the data in the final year of the rolling window as the validation sample and utilize data before the final year to train the model. We then apply the trained model to the validation sample to predict stock prices (for example, for year 2001, we predict prices for the validation sample in 2000). The prediction error is the difference between actual log price and the predicted log price. We then take the average value of the squared forecast errors for each year.

Squared Prediction Errors

Year	Lasso	Elastic Net	SVM	Random Forest	Gradient Boost	LSTM
2001	1.58	1.48	0.69	0.61	0.57	0.48
2002	1.94	1.94	0.90	0.46	0.60	0.62
2003	1.49	1.37	0.44	0.35	0.32	0.33
2004	1.44	1.30	0.46	0.29	0.28	0.28
2005	1.50	1.29	0.38	0.25	0.27	0.25
2006	1.34	1.11	0.40	0.21	0.23	0.25
2007	1.57	1.39	0.46	0.24	0.25	0.25
2008	2.28	2.06	0.88	0.61	0.64	0.63
2009	2.21	1.81	0.64	0.52	0.41	0.44
2010	1.68	1.24	0.68	0.31	0.35	0.65
2011	1.69	1.28	0.75	0.58	0.46	0.52
2012	1.80	1.36	0.48	0.26	0.29	0.25
2013	1.46	1.10	0.47	0.24	0.23	0.20
2014	1.36	1.04	0.38	0.21	0.21	0.22
2015	2.24	1.72	0.63	0.38	0.38	0.39
2016	1.84	1.41	0.65	0.37	0.38	0.41
2017	1.17	0.93	0.45	0.30	0.27	0.23
2018	1.38	1.05	0.42	0.22	0.25	0.22

Table IA2: Portfolio Performance following Machine vs. MDM Recommendations: Monthly Rebalancing

In each month, we gather all predictions made by all machine-debiased-man (MDM) analysts and the corresponding AI forecasts in the past 30, 60, 90, and 360 days. For each pair of predictions, if the AI predicts a higher (lower) price than the MDM analyst, it is considered as a buy (sell) signal. During the given time horizon, the portfolio will long the stock if there are more buy than sell signals, and short the stock otherwise. The portfolios are equal weighted and rebalanced monthly, i.e., a position is held for one month or till the signals reverse. The monthly percentage returns of the long-short, long-leg (stocks only with a buy sign) and short-leg portfolios (stocks only with a short sign) as well as the alphas generated from the FF3, FFC4, FF5, and FF6 models are presented. The OLS standard error is used to construct t-stats. The t-stats are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

		AI vs. MDM			
Long-Short		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.65**	0.73***	0.67***	0.51***
		(2.43)	(3.81)	(3.77)	(3.00)
	FF3	0.61**	0.70***	0.62***	0.48***
		(2.18)	(3.51)	(3.35)	(2.65)
	FFC4	0.67**	0.75***	0.67***	0.51***
		(2.42)	(3.82)	(3.65)	(2.81)
	FF5	0.52**	0.58***	0.54***	0.39**
		(1.97)	(3.13)	(3.13)	(2.39)
	FF6	0.54**	0.60***	0.56***	0.41**
		(2.08)	(3.28)	(3.27)	(2.49)
Long-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	1.33***	1.41***	1.41***	1.33***
		(2.78)	(3.15)	(3.18)	(3.16)
	FF3	0.61***	0.64***	0.62***	0.59***
		(3.22)	(5.78)	(5.84)	(5.83)
	FFC4	0.71***	0.73***	0.72***	0.67***
		(3.95)	(7.78)	(7.84)	(7.56)
	FF5	0.68***	0.67***	0.67***	0.63***
		(3.77)	(6.40)	(6.67)	(6.78)
	FF6	0.71***	0.69***	0.69***	0.66***
		(4.12)	(7.79)	(8.09)	(8.00)
Short-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.68	0.68	0.74	0.82*
		(1.40)	(1.45)	(1.61)	(1.85)
	FF3	0.00	-0.06	0.00	0.11
		(-0.00)	(-0.35)	(0.02)	(0.62)
	FFC4	0.04	-0.02	0.05	0.16
		(0.19)	(-0.11)	(0.29)	(0.96)
	FF5	0.16	0.08	0.13	0.24
		(0.81)	(0.51)	(0.84)	(1.56)
	FF6	0.16	0.09	0.14	0.25*
		(0.85)	(0.55)	(0.89)	(1.68)

Table IA3: Portfolio Performance following Machine vs. Man and Machine vs. MDM Recommendations: Semi-annual Rebalancing

For every six months, we gather all predictions made by all analysts and the corresponding AI forecasts in past 30, 60, 90 and 360 days. For each pair of predictions, if AI predicts a higher (lower) price than the Analyst, it is considered as a buy (sell) signal. During the given time horizon, the portfolio will long the stock if there are more buy than sell signals; and short the stock otherwise. The portfolio is equal weighted. The portfolios are rebalanced monthly, i.e., a position is held for one month or till the signals reverse. The monthly percentage returns of the long-short, long-leg (stocks only with a buy sign) and short-leg portfolios (stocks only with a short sign) as well as alphas generated from FF3, FFC4, FF5 and FF6 models are presented. We also compare the portfolio performance between the Machine-debiased analysts (MDM) and AI. The OLS standard error is used to construct *t*-stats. The *t*-stats are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Panel A. Portfolio returns with semi-annual rebalancing – AI vs. Analyst

AI vs. Analyst					
Long-Short		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	1.59***	1.53***	1.36***	1.07***
		(3.81)	(3.69)	(3.50)	(3.04)
	FF3	1.66***	1.60***	1.40***	1.18***
		(3.88)	(3.76)	(3.59)	(3.15)
	FFC4	1.60***	1.53***	1.35***	1.10***
		(3.76)	(3.61)	(3.47)	(2.96)
	FF5	1.46***	1.41***	1.02***	0.83***
		(3.61)	(3.50)	(2.83)	(2.51)
	FF6	1.44***	1.38***	1.01***	0.80**
		(3.55)	(3.42)	(2.81)	(2.41)
Long-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	1.89***	1.88***	1.87***	1.68***
		(4.87)	(4.90)	(4.87)	(4.66)
	FF3	1.18***	1.17***	1.16***	1.05***
		(8.44)	(8.79)	(8.70)	(7.93)
	FFC4	1.26***	1.24***	1.22***	1.09***
		(9.37)	(9.61)	(9.45)	(8.36)
	FF5	1.03***	1.02***	1.01***	0.87***
		(8.08)	(8.46)	(8.37)	(7.62)
	FF6	1.08***	1.07***	1.06***	0.90***
		(9.14)	(9.41)	(9.24)	(8.14)
Short-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.30	0.35	0.51	0.61
		(0.51)	(0.60)	(0.90)	(1.15)
	FF3	-0.48	-0.43	-0.24	-0.14
		(-1.14)	(-1.02)	(-0.65)	(-0.38)
	FFC4	-0.35	-0.29	-0.13	-0.01
		(-0.84)	(-0.71)	(-0.36)	(-0.04)
	FF5	-0.43	-0.39	-0.01	0.04
		(-1.09)	(-0.98)	(-0.04)	(0.11)
	FF6	-0.35	-0.30	0.04	0.11
		(-0.90)	(-0.78)	(0.12)	(0.33)

Panel B. Portfolio returns with semi-annual rebalancing – AI vs. MDM

AI vs. MDM					
Long-Short		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.60***	0.53***	0.54***	0.42***
		(7.20)	(6.73)	(6.88)	(5.61)
	FF3	0.56***	0.48***	0.48***	0.41***
		(6.73)	(6.15)	(6.22)	(5.37)
	FFC4	0.58***	0.49***	0.49***	0.41***
		(7.04)	(6.29)	(6.34)	(5.34)
	FF5	0.56***	0.50***	0.51***	0.41***
		(7.09)	(6.70)	(6.88)	(5.76)
	FF6	0.57***	0.50***	0.51***	0.41***
		(7.24)	(6.76)	(6.93)	(5.76)
Long-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	1.36***	1.36***	1.36***	1.33***
		(2.94)	(2.97)	(2.98)	(3.11)
	FF3	0.51***	0.50***	0.50***	0.54***
		(4.63)	(4.90)	(4.94)	(5.39)
	FFC4	0.59***	0.58***	0.58***	0.61***
		(6.11)	(6.42)	(6.48)	(6.81)
	FF5	0.62***	0.61***	0.62***	0.64***
		(5.97)	(6.48)	(6.64)	(7.05)
	FF6	0.63***	0.63***	0.64***	0.65***
		(6.82)	(7.40)	(7.58)	(7.88)
Short-Leg		30 day inform	60 day inform	90 day inform	360 day inform
Monthly returns	Ret	0.76*	0.83*	0.82*	0.91**
		(1.69)	(1.85)	(1.84)	(2.14)
	FF3	-0.04	0.02	0.02	0.13
		(-0.44)	(0.25)	(0.19)	(1.37)
	FFC4	0.02	0.09	0.08	0.20**
		(0.19)	(1.05)	(0.99)	(2.41)
	FF5	0.05	0.12	0.11	0.22**
		(0.57)	(1.35)	(1.30)	(2.69)
	FF6	0.06	0.13	0.12	0.2**4
		(0.72)	(1.60)	(1.55)	(3.20)