Voluntary Disclosure, Misinformation, and AI Information Processing: Theory and Evidence

Abstract

This paper investigates the impact of generative AI on firms' voluntary disclosure choices. Our theoretical model highlights a trade-off between AI's improved ability to process disclosed information and its potential for misinformation, modeled as a random "hallucination" unrelated to the firms' fundamentals. We predict that increased AI processing leads to more strategic non-disclosure due to two related economic forces. First, hallucinations provide additional camouflage after strategic non-disclosure. Second, because users consider the risk of misinformation, they discount observed marginal disclosures, further reducing the benefit of disclosure. To test our predictions, we leverage OpenAI's launch of ChatGPT in November 2022 as a shock to AI processing. Consistent with the theory, firms with more AI processing reduce their voluntary disclosures. Further, the introduction of ChatGPT reduces information processing failures, which manifests in increased information processing speed. Combining the crowding-out effect on information supply and the positive impact on information processing speed, we do not find evidence of a net increase in information quality.

Keywords: *AI hallucination, Information processing, Information crowding out, ChatGPT, Voluntary disclosure.*

JEL Codes: O00, O33, G10, G14.

I. Introduction

The development of generative AI has been transformative in capital markets. On the one hand, investors have been exposed to unprecedented advancements in information technology as machines' cognitive and communication capabilities can be compared to those of humans for the first time in history. On the other hand, Geoffrey Hinton, the "godfather" of AI and 2024 Nobel Prize winner in Physics, repeatedly warns of the potential negative impacts of generative AI (New York Times, May 1, 2023).[1] The rise of generative AI technologies, such as OpenAI's ChatGPT, has amplified societal concerns, particularly regarding misinformation. Large language models (LLMs) can generate highly realistic text, raising significant concerns about the spread of difficult-to-detect misinformation. Regulators, such as the European Commission, have called for attention to the risks of AI-generated misinformation, while public media outlets have highlighted the dangers of realistic yet factually incorrect content (NBC New York, 2023; DW, 2024).[2]

At first glance, the impact of generative AI on the capital market's information environment appears to be ambiguous. These tools, while potentially disseminating misinformation, also enhance the information processing capabilities of agents, accelerating the incorporation of information into prices (Sims, 2003; Dong et al., 2016; Blankespoor et al., 2020; Kim et al., 2024a). To understand this trade-off, we present a theoretical model that jointly captures AI's incremental information processing capabilities and its potential for generating misinformation.[3] Our main prediction is that as investors increasingly rely on AI information processing, this processing reduces firms' incentives to disclose and thus crowds out firms' voluntary disclosures. Our

[1] Geoffrey Hinton's concerns about AI resonate with the notion of a potential digital dystopia, which was once depicted only in allegorical tales such as *I, Robot*, and *2001: A Space Odyssey*. *I, Robot* is a novel written by Isaac Asimov in 1950 that explores the complex relationship between humans and robots. Asimov raises questions about human autonomy and potential consequences of advanced technology. *2001: A Space Odyssey* is a collaboration between Arthur C. Clarke and Stanley Kubrick. The story follows a mission to Jupiter, during which an artificial intelligence begins to malfunction, posing an existential threat to the crew. Today, these fictional scenarios are becoming increasingly plausible with the significant impact of generative AI on the production and consumption of information. See https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html.

[2] See https://www.disinfo.eu/publications/platforms-policies-on-ai-manipulated-and-generated-misinformation/, https://www.bbc.com/news/technology-65110030, https://www.nbcnewyork.com/investigations/fake-news-chatgpt-has-a-knack-for-making-up-phony-anonymous-sources/4120307/, and https://akademie.dw.com/en/generative-ai-is-the-ultimate-disinformation-amplifier/a-68593890.

[3] This trade-off reflects a fundamental aspect of our modern society with the presence of generative AI. The pursuit of convenience and efficiency drives technological advancement. However, concerns about misinformation and information control by advanced technology have been longstanding themes in literature, such as Aldous Huxley's dystopian novel - *Brave New World*.

empirical tests center on the model predictions and provide consistent evidence using management forecasts as proxies for firms' voluntary disclosure and measures of firm-level AI-processing intensity.

To set intuition for a theoretical mechanism, we develop a stylized voluntary disclosure game with a friction along the lines of Dye (1985) and Jung and Kwon (1988).[4] In the model, the firm is always informed about a fundamental signal and chooses whether to disclose or stay silent. However, before disclosing, the firm does not know whether its message will be processed by an AI system (AI processing) or human analysis (human processing). We assume that human processing is constrained by processing capacity and may sometimes fail to see the disclosure even when one is made. This assumption is similar to what Blankespoor et al. (2019, 2020) define as the "awareness" cost of information processing (i.e., "monitoring for the disclosure's existence") and reflects that humans have limited capacity to scan all information sources for the presence of information. In the model, human processing does not distinguish whether the signal is disclosed but unobserved (i.e., unawareness) or the firm strategically withholds information.

In contrast, AI processing is not subject to unawareness. For simplicity, we assume that AI processing can always process firms' disclosures. However, its capacity to process extensive information sources carries an inherent limitation: it may mistakenly interpret unrelated evidence as informative disclosures, especially when firms strategically withhold information. This can be viewed as akin to an AI "hallucination."[5] Importantly, the resulting misinformation is, for our purpose, a modeling abstraction to anchor a simple and testable trade-off.[6] We do not mean an AI would imagine fake products, contracts, or earnings announcements. However, an AI may (mistakenly) point users to assessments about the firm that are not descriptive of the information

[4] Our model does not aim to realistically represent information processing within a complex market institution. Instead, it is designed to illustrate a particular mechanism where certain types of processing frictions can discourage voluntary disclosure. Our theory may apply to settings other than AI, such as types of involuntarily confusing or complex information where a non-disclosure may be (incorrectly) interpreted as informative.

[5] Merriam-Webster (2023) defines hallucination as "a plausible but false or misleading response generated by an artificial intelligence algorithm."

[6] We acknowledge that these assumptions only aim to clarify the key economic forces driving the theory and provide grounding for an empirical hypothesis. In practice, a comparison between AI and human processing is subject to many other differences that could invalidate the hypothesis. For example, we do not consider the roles of human experience. Similarly, AIs do not always hallucinate, nor are they guaranteed to identify and process information when it exists. In model extensions, we develop several of these forces formally and show that they can yield countervailing effects.

actually reported.[7]

We show that while AI improves the processing of disclosure, a higher probability of AI processing reduces firms' voluntary disclosure (i.e., increases the probability of strategic withholding). This occurs because of two interconnected forces. First, the potential for hallucination provides additional camouflage for a strategic non-disclosure. Non-disclosing firms are no longer pooled solely with an average signal if the disclosure is not processed by the human. When the AI hallucinates a message that would have been disclosed, non-disclosing firms are now pooled with firms with better information who voluntarily disclose. This implies that by holding the disclosure threshold fixed, the payoff to non-disclosure is typically greater than if only humans process information. Second, since information users must now consider the possibility of misinformation, a Bayesian correction is made to observed signals above the disclosure threshold, reducing their effects on firm value. This leads to a reduction in the payoff from disclosure, further increasing the benefit of strategic withholding.

Applying the minimum principle (Acharya et al., 2011; Guttman et al., 2014), we further show that the increase in the disclosure threshold must shift the equilibrium away from the minimum non-disclosure belief and thus increase the non-disclosure price. Put differently, the change to the disclosure environment from greater AI processing offers a new test of the minimum principle. Because the threshold in this type of evidence game minimizes the non-disclosure price relative to any other threshold, we predict that the greater use of AI increases the non-disclosure price.

Empirical tests of our model's predictions are challenging because we do not directly observe how financial market participants use generative AI tools. To address the challenge, we utilize OpenAI's launch of ChatGPT 3.5 in November 2022 as a significant technological advance in generative AI development. ChatGPT 3.5, recognized for its contextual awareness and coherent conversations, rapidly achieved 100 million monthly active users within two months after its introduction (Reuters, 2023).[8] However, ChatGPT 3.5's ability to produce highly realistic text raises concerns about misinformation. Therefore, the launch provides a natural setting to test our

[7] We empirically verify that ChatGPT 3.5 hallucinates by querying it about listed firms' management forecasts a total of 10,000 times. Our results show a substantial probability that ChatGPT 3.5 provides a forecast when the firm does not issue one or fails to provide a forecast when one is issued. Additionally, the probability of hallucination is significantly higher for firms that withhold information. For more details, see Section 2.2.1.

[8] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

theoretical predictions on the crowding-out effect of AI processing.[9]

Our model predicts that the probability of AI processing impacts managers' disclosure incentives. We construct a firm-level measure of AI-savvy sell-side financial analyst coverage to proxy for the probability of AI processing. We categorize analysts as AI-savvy ("technical analysts") if they possess technical skills (e.g., artificial intelligence) or have majored in technical subjects (e.g., STEM majors). Our empirical proxy relies on two assumptions. First, we assume that financial analysts with technical backgrounds are more likely to use generative AI and, thus, are potentially more aware of its information processing capabilities and the associated risks of misinformation. This assumption is supported by the survey evidence from Bick et al. (2024).[10] Second, we assume that interactions between technical analysts and managers (e.g., conference calls and investor days) enhance managers' awareness of the benefits of AI-facilitated information processing and the potential for misinformation. This awareness incentivizes managers to adjust their disclosure strategies accordingly. Based on these assumptions, we classify firms covered by tech-savvy analysts as treated firms—whose managers are more aware of the costs and benefits of generative AI—while firms not covered by such analysts serve as the control group. In this context, the interaction term between the treatment indicator (coverage by AI-savvy analysts) and the post-ChatGPT time indicator proxies for the probability of AI processing in our model. This approach allows us to assess whether and how information receivers' reliance on generative AI affects firms' information supply.

We employ a difference-in-differences research design to assess the impact of AI processing on voluntary disclosure. Treated firms exhibit an economically significant reduction in managerial forecasts in the post-ChatGPT 3.5 periods, using various managerial forecasts from 2021Q1 to 2023Q4 as proxies for voluntary disclosures. Specifically, we document a 19.8% reduction in the

[9] When GPT-3 was launched in June 2020, it was offered with limited access, and its training data had a delay, with the initial version trained up to October 2019. Similarly, GPT 3.5 was trained with data until September 2021, reflecting a lag in updates. While this delay raises concerns about AI's ability to provide timely information, it does not negate the relevance of older data. Rather than serving as a search engine replacement, GPT models are better suited for analyzing existing documents, such as firm disclosures (e.g., 10-K reports). Additionally, GPT's capacity to integrate real-time web data (even though its neural network may not have been trained by this data), along with competing models like Gemini, offers the possibility to access and process up-to-date information. This suggests that AI tools, even with a lag in training data, may remain valuable for information processing.

[10] Bick et al. (2024) document survey evidence that 46 percent of workers with STEM degrees use generative AI at work, compared with 40 percent for those with business, economics, or communication majors, and 22 percent for those in other fields, including liberal arts and humanities.

overall volume of managerial forecasts compared to the sample average forecast probability. The observed trend reflects a significant reduction in firms' propensity to provide voluntary disclosures due to information users (such as analysts) increasingly relying on generative AI. This supports our main hypothesis on the crowding-out effect of generative AI usage on firms' information supply. To reinforce the causal relationship, we estimate the dynamic treatment effects from periods before to after the introduction of ChatGPT 3.5. Consistent with parallel trends, we find no disparities in voluntary disclosure between treated and control firms before the introduction of ChatGPT 3.5. Changes in disclosure practices begin to emerge a quarter after the launch of ChatGPT 3.5.

An exogenous shock to AI processing and the resulting decrease in firm disclosure does not ensure that our empirical findings align with the theoretical model. The gap between the empirics and theory may occur if our empirical proxy of firm-level AI processing does not accurately reflect AI processing in the model. Additionally, the correlation between AI processing and voluntary disclosure might be influenced by omitted economic factors. For example, firms with more technology-savvy analysts may have been more affected by the emergence of AI. Changes in firm-level operational risks due to AI could affect managers' incentive to issue forecasts.

Although these threats to identification cannot be fully resolved without a randomized assignment of AI, we test another important implication of our explanation to provide additional support for the theoretical mechanism. The impact of AI processing on firms' voluntary disclosure is potentially more pronounced for more complex firms. The intuition is that, for complex firms, AI processing is more likely to encounter and mistakenly interpret unrelated information as relevant to the firm's fundamentals, thereby increasing the potential for misinformation. This heightened probability of AI-hallucinated misinformation leads to a stronger effect on voluntary disclosure decisions. Empirically, we employ a triple difference-in-differences design and find consistent evidence that the decline in disclosure by treated firms after the introduction of ChatGPT is more pronounced for complex firms, as measured by whether firms have foreign operations.

Moreover, we empirically test the beneficial effect of generative AI in mitigating humans' information processing failures, as highlighted in the model. However, the probability of processing failure by information recipients is a deep parameter within the model and is

empirically unobservable. We adopt the structural approach outlined by Smith (2024) and estimate the information processing speeds of voluntary disclosure for both the treatment and control firms from before to after the ChatGPT 3.5 introduction.[11] The underlying rationale is that information processing failures should manifest in a reduced speed of price discovery, as humans need more time to find the information. Our findings show that investors' processing speeds increase for treated firms with management forecasts in the post-ChatGPT 3.5 periods relative to control firms, which is consistent with a lower likelihood of processing failures.

Next, we explore whether the introduction of ChatGPT 3.5 has increased the overall information embedded in stock prices. We estimate the structural model in Smith (2024) to assess the changes in overall information being incorporated into stock prices after the introduction of ChatGPT 3.5. Importantly, we find an insignificant net effect on the information embedded in stock prices. One potential explanation is that the AI's positive effect on mitigating information processing failures is offset by a negative crowding-out impact on firms' information supply.[12] Nevertheless, we caution against overinterpreting this result. The net effects of ChatGPT 3.5 may take time to materialize and may not be fully captured by our analysis within this relatively limited investigation window.

Furthermore, our model not only predicts a crowding-out effect on disclosure but also offers a testable implication about market reactions to disclosure. Since investors cannot perfectly distinguish between real and hallucinated disclosures, they apply a Bayesian correction (i.e., a discount) to observed disclosures. We use analyst forecast revisions around the issuance dates of management forecasts to identify market participants' reactions to management forecasts. Consistent with the theory, we document that analysts' reactions to management forecasts are significantly lower for treated firms following the launch of ChatGPT 3.5.

[11] Note that management forecasts are typically released alongside earnings announcements within a short timeframe (Rogers and Van Buskirk, 2013). In this context, using Smith's (2024) approach, the processing speed of management forecasts reflects investors' ability to process information embedded in both earnings announcements and management forecasts. To distinguish the specific processing speed of management forecasts, we compare two scenarios: earnings announcements from firms that issue management forecasts versus earnings announcements from firms without concurrent managerial forecasts. This comparison helps us to tease apart the processing speed attributed solely to management forecasts from those associated with both earnings announcements and managerial forecasts. See Section 5.3 for more details.

[12] Our findings of an insignificant increase in the informativeness of stock prices may help address an alternative explanation that investors utilize AI tools to become relatively more informed, which reduces the information gap between investors and managers and thus reduces managers' incentives to issue forecasts (e.g., Verrecchia, 1983).

Last, we explore the implications of the minimum principle in evidence games (Acharya et al., 2011; Guttman et al., 2014) by testing whether increased AI processing increases the non-disclosure threshold and associated non-disclosure price. To this end, we utilize future earnings per share (EPS) as a measure of the non-disclosure threshold and employ the price-to-earnings (PE) ratio and Tobin's Q as proxies for current non-disclosure prices. We test whether there are increases in these variables for firms choosing non-disclosure when their analysts increasingly rely on generative AI. Our results show an increase in the non-disclosure threshold, measured as a higher average future EPS for non-disclosing firms. Furthermore, we observe that the average PE and Tobin's Q are higher for non-disclosing firms, which potentially reflects changes in investors' posterior beliefs about non-disclosing firms following the introduction of ChatGPT-facilitated AI processing.

Taken together, our tests using the ChatGPT introduction imply that AI processing enhances information processing speed while reducing firms' information supply. However, we caveat that our model only seeks to address the trade-off between two primary factors: the concern of misinformation and the enhanced information processing capacity of generative AI. Other factors beyond our model may also contribute to the crowding-out effect in managerial forecasts in the real world. For example, Einhorn (2018), Banerjee et al. (2024), and Libgober et al. (2023) model environments in which other parties may possess private information separate from the firm's disclosures. To the extent that AI facilitates access to such private information, these mechanisms may also be at play. However, these theories further demonstrate that the effects of these mechanisms can be subtle and, under certain conditions, lead to crowding-in. Another related study by Frenkel et al. (2020) identifies a potential for crowding out when a third party discloses information that may have been strategically withheld. Therefore, our empirical tests serve primarily as supportive evidence for one mechanism but do not exclude other explanations. Beyond validating the primary crowding-out prediction, our additional tests—including cross-sectional analyses of firm complexity, processing speed assessments, evaluations of stock price informativeness, examinations of the sensitivity of analyst revisions to disclosure, and minimum principle tests—further substantiate multiple predictions that, jointly, are tied to our proposed theoretical mechanism.

Nonetheless, while our findings are consistent with the trade-off between misinformation

concerns and enhanced information processing capacities described in our model, we recognize that they pertain to the early stages of AI processing and may evolve as the technology advances. On the one hand, the impact of AI-hallucinated misinformation could decrease with the development of new algorithms aimed at mitigating such misinformation.[13] On the other hand, there is an increasing risk that third parties may deliberately abuse ChatGPT-related technologies to produce hard-to-verify fake news. These factors extend beyond the scope of our empirical analyses.

II. Related Literature and Institutional Background

2.1. Related Literature

Our research advances the literature along several directions, with a particular focus on the rapidly evolving research on the processing of corporate disclosures. The foundational literature primarily focuses on frictions to communication due to characteristics of the information received by the disclosing firm, which prevent the disclosure of unfavorable information (e.g., Verrecchia, 1983; Dye, 1985). Recent research, by contrast, emphasizes frictions originating from the users of information, as users may face processing costs and capacity constraints when collecting and analyzing information from multiple sources (Blankespoor et al., 2020). These frictions, in turn, feed back into market prices and alter firms' incentives to disclose voluntarily. We contribute to this line of research by theoretically integrating the challenge of misinformation, a particularly salient issue in the era of technology-assisted information processing. Our theoretical contribution is based on the presumption that AI-assisted information processing makes it more difficult, relative to humans, to identify "unknowns" (i.e., knowing when information does not exist)[14] and easier to process "knowns" (i.e., when firms disclose the information).

Second, our study extends the large body of empirical research on the determinants of corporate voluntary disclosure and market reactions to these disclosures (Beyer et al., 2010). We

[13] Example mitigation techniques include retrieval augmented generation (RAG) and knowledge graph integration (Gao et al., 2023; Pan et al., 2024).

[14] See Li et al. (2024) for a comprehensive survey of the literature on the honesty of large language models (LLMs). Honesty is a crucial principle for aligning LLMs with human values, as it requires these models to accurately recognize and communicate the limits of their knowledge. However, current LLMs still exhibit significant dishonest behaviors, such as confidently providing incorrect answers instead of admitting uncertainty with statements like "I don't know" when they lack sufficient information.

provide empirical evidence on the impact of misinformation generated by generative AI on firms' information supply. Our identification strategy and empirical measures are motivated by model parameters (e.g., probability of AI processing, processing failures, non-disclosure threshold and prices, etc.), and the findings are consistent with the model's comparative statics.

Our findings relate to those of Cao et al. (2023) but differ in several important aspects. Cao et al. (2023) show that firms make their SEC filings more machine-friendly when more information recipients use machines to download these filings. We focus on the benefits and costs of AI processing. On the benefit side, we estimate the structural model by Smith (2024) and document the positive impact of AI in mitigating information processing failures, which are empirically unobservable but can be inferred through structural estimation. On the cost side, we demonstrate that the potential for misinformation undermines the credibility of corporate disclosure and diminishes the voluntary supply of information. This highlights a significant crowding-out effect, where the presence of AI-hallucinated misinformation can obscure the value of real corporate disclosures, thereby influencing asset price informativeness. Our findings on the crowding-out effect of AI on information supply may be of general societal interest, as they may serve as a forewarning of new challenges emerging in the digital economy.

A recent empirical study by Bertomeu et al. (2024) explores how the ChatGPT ban in Italy impacts information users, specifically the behavior of financial analysts. Their findings show that the ban discourages financial analysts from intermediating information and leads to greater information asymmetry among investors. This paper differs in several key ways. First, we focus on the trade-off between the information processing benefits of ChatGPT and its misinformation generation, and the resulting impacts on the information supplier—firm managers. Second, our finding that increased AI information processing does not significantly alter the information embedded in stock prices is not necessarily inconsistent with their results. The ChatGPT ban in Italy was short-lived, lasting only one month, and therefore, it may not have the time to significantly impact managers' information supply decisions.[15]

2.2. Institutional Background on AI Hallucinations

Large language models (LLMs) show great promise in supporting investment-related tasks,

[15] In fact, our results in Figure 5 suggest that it may have taken at least one quarter for management forecast policies to adjust to the presence of AI processing, and the estimated coefficient of the second quarter is small.

such as analyzing corporate disclosures and forecasting earnings (Kim et al., 2024a, 2024b). However, these advancements come with challenges, notably the tendency of LLMs to produce "hallucinations," or generate misleading information. AI hallucination has become a major concern given that LLMs may generate false information at an unexpectedly high frequency.[16] Merriam-Webster (2023) defines hallucination as "a plausible but false or misleading response generated by an artificial intelligence algorithm." Hallucinations have garnered considerable media attention (Weise and Metz, 2023), and U.S. President Biden issued an Executive Order emphasizing the need for safeguards against misleading outputs from generative AI systems (Biden, 2023).

2.2.1. Verification Test on ChatGPT 3.5 Hallucinations

We empirically validate the potential of hallucinations in ChatGPT 3.5. Specifically, we access the GPT 3.5-Turbo API and use a prompt to request management forecasts for EPS for fiscal year 2020. We randomly select 50 firms in our sample that issued forecasts for 2020, and 50 firms that did not. For each firm, we prompt ChatGPT with the following question: "What is [firm_name]'s management forecast for EPS for the fiscal year 2020? Your response should follow exactly the same pattern and do not add any additional words: 1. If there is a management forecast, return The value of the forecast: followed by ONLY a specific value or a range. 2. If there is no management forecast, return ONLY There is no management forecast." We query ChatGPT 100 times for each of the 100 firms, resulting in 10,000 responses.

Figure 1 illustrates the frequency of hallucination rates for firms with (red) and without (blue) management forecasts. We classify ChatGPT 3.5's responses as hallucinations if ChatGPT provides a forecast when the firm did not issue one for fiscal year 2020 or ChatGPT fails to provide a forecast when one was issued. The hallucination rate for each firm is calculated by dividing the

[16] LLMs are built on deep learning architectures, typically using transformer networks. These models are trained on vast amounts of text data, learning patterns, and associations within the language by predicting the next word or sequence of words. This allows LLMs to generate human-like text based on the input they receive. See Kim et al. (2024a) section II. A for explanations of ChatGPT's transformer architecture.
In the computational literature, numerous explanations have been proposed for why LLMs hallucinate. First, hallucinations can occur due to inadequacies in the training data, which includes either false information within the training data itself (e.g., Ji et al., 2023) or outdated data that fail to account for current events and lead to gaps in accuracy (e.g., Aksitov et al., 2023). Second, hallucinations can occur due to the sequential generation of text. LLMs generate fluent and coherent text by predicting each subsequent word based on patterns in their training data, which can result in plausible sounding yet factually incorrect responses since they are optimized for text generation rather than verifying factual accuracy. Notably, Kalai and Vempala (2024) demonstrate a statistical lower bound on the rate at which pre-trained language models produce hallucinations.

number of hallucinated responses by the number of queries (i.e., 100). The histogram bars denote the frequency of hallucination rates, while the fitted density lines highlight the overall trends within each group of firms. We observe that ChatGPT 3.5 significantly hallucinates more for firms without management forecasts (the blue group).

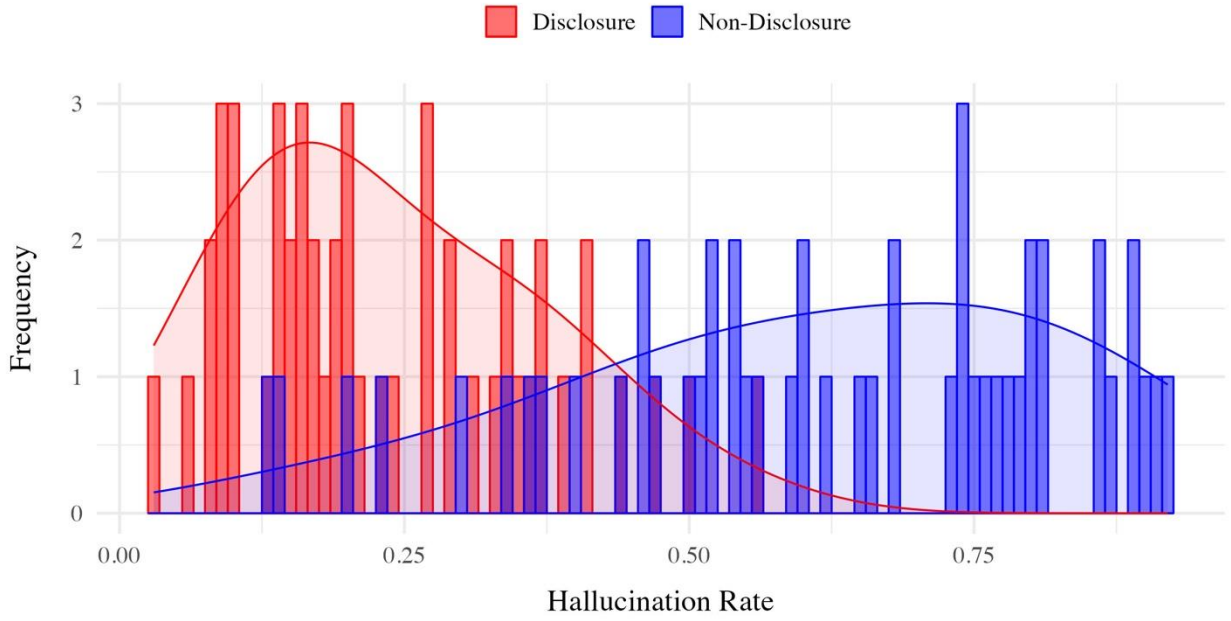


Figure 1: Histogram of Hallucination Rates by Firms With and Without Management Forecasts

Table 1 presents the detailed statistics of ChatGPT 3.5's hallucinations. First, we show that the hallucination rate is significantly higher for firms without disclosure (61.5%) compared to those with disclosure (23.3%), resulting in a statistically significant difference of 38.2%. This finding supports our model's assumption that AI processing is more prone to generating misinformation when firms withhold disclosures. Second, conditional on ChatGPT providing an answer on management forecast, we assess the accuracy of its responses by calculating the absolute difference between ChatGPT's answer and the actual EPS for fiscal year 2020. We also create a scaled measure by dividing the absolute difference by the stock price and multiplying by 100. Across both measures (rows 2 and 3), we find that ChatGPT's answers are less accurate (i.e., higher absolute differences) for firms without management forecasts compared to those with forecasts, and the differences across the two groups of firms are statistically significant. Our results demonstrate that ChatGPT 3.5 generates significantly more hallucinations and incurs more

significant errors for firms without management forecasts, highlighting the increased risk of misinformation and inaccuracies when firms withhold information.

Table 1: Summary Statistics of ChatGPT 3.5's Hallucinations

	Mean		**Difference (1) – (2)**
	Non-Disclosure Firms (1)	Disclosure Firms (2)	
Hallucination Rate	0.615	0.233	**0.382 (*t*=10.86)**
Forecast Error	3.057	1.951	**1.106 (*t*=6.90)**
Forecast Error/Price(*100)	20.857	4.429	**16.428 (*t*=5.13)**

III. Theoretical Framework

3.1. Assumptions and Equilibrium

We consider a disclosure game in the spirit of Dye (1985) and Jung and Kwon (1988), where an informed manager communicates information to investors. Our objective is to develop a straightforward economic trade-off capturing the potential differential impacts of AI versus human information processing. To achieve this, we employ an abstract model featuring a representative investor who can choose to process information either with an AI system or through human analysis. While this approach is not intended to offer descriptive realism, it aims to capture economic tensions from both types of information processing. We include the proofs for this section in Appendix B.

In the model, the manager privately observes a value $\tilde{v}$, drawn from a distribution with c.d.f. $F(.)$ and p.d.f. $f(.)$ with support over $V = [\underline{v}, \bar{v}]$ and mean μ. The firm manager can report truthfully or, strategically, remain silent, which we write as $d(v) \in \{v, \emptyset\}$. We maintain the assumption that prosecuting strategic withholding is not possible and assume an informed manager to keep the model as minimal as possible. However, as will become evident in the model's discussion, this assumption is not critical to our analysis. There are no other costs to disclose, and the model is intentionally designed so that any friction to information processing is sufficient to prevent classical unraveling (Milgrom, 1981; Dickhaut et al., 2003; Bourveau et al., 2020; Jin et al., 2021).

Specifically, the investor does not directly observe the disclosure but instead processes an imperfect signal, consistent with theories of rational information processing (Blankespoor et al., 2020; Bertomeu et al., 2023). With probability $p \in [0,1]$, investors' information collection is

described as "human processing." Humans have limited attention (Hirshleifer and Teoh, 2003; Abramova et al., 2020) and may fail to observe the disclosure with probability $q \in [0,1]$, in which case their information set is $d_h(v) = \emptyset$ regardless of $d(v)$. With complementary probability $1 - q$, human processing is able to correctly observe $d_h(v) = d(v)$. The investor prices the firm according to the Bayes' rule as $P_h(x) = \mathbb{E}(v \mid d_h(v) = x)$. When observing $d_h(v) = \emptyset$, the investor cannot differentiate whether the message is unobserved due to inattention or strategically withheld by the manager.[17]

With probability $1 - p$, the investor relies on AI to collect information. Unlike humans, the AI can always process the sender's disclosure. We denote the AI's observable report as $d_a(v) = v$ conditional on disclosure. However, the AI hallucinates when there is no disclosure, creating a new random noise signal $d_a(v) = \tilde{v}_a$, similar to noisy talk in Blume et al. (2007) or fake news in Frenkel et al. (2024).[18] We wish to avoid situations in which the noise is exogenously biased to issue good (resp., bad) signals, since then this would bias the analysis toward the AI rewarding (resp., punishing) non-disclosure. Hence, we set v_a to be drawn from $F(.)$ so the noise is calibrated to the correct distribution.[19]

The noise signal $\tilde{v}_a$ is independent from the true fundamentals, and the AI is not reporting to the user whether v_a or v is being observed. In other words, we assume that the user cannot distinguish between misinformation and the true signal. The user applies Bayes' rule, thereby rationally processing the imperfect information provided by the AI and recognizing that the signal may be garbled.

We summarize the timeline of the model in Figure 2 below.

---

[17] Imperfect human processing is mathematically equivalent to the friction described by Dye (1985), as the information sets are identical whether the investor fails to receive the disclosures or the manager is uninformed and therefore unable to disclose. For the sake of exposition, we introduce an investor-level friction to create greater symmetry in the model between AI and human frictions. Our results remain robust even if the human is always informed ($q = 0$). The inclusion of imperfect human processing simply ensures that the model does not inherently assume that humans are superior to AI in terms of information processing.

[18] While AIs increasingly have the capacity to consult factual sources, this functionality is not always reliable. In our model, we assume that the AI always hallucinates, which is not literally accurate, as in reality, an AI might sometimes recognize when an answer does not exist or when it is unable to locate one. This model serves as a conceptual simplification to highlight the core idea. However, the comparative statics remain applicable in a generalized model where the AI may not always hallucinate with positive probability. Specifically, assuming that the AI is less prone to hallucination would increase voluntary disclosure by shifting the disclosure threshold closer to classical unraveling.

[19] This assumption also ensures that, in principle, a user cannot identify if an AI is systematically hallucinating by repeating a query many times and comparing the distribution of $\tilde{v}_a$ to $F(.)$.

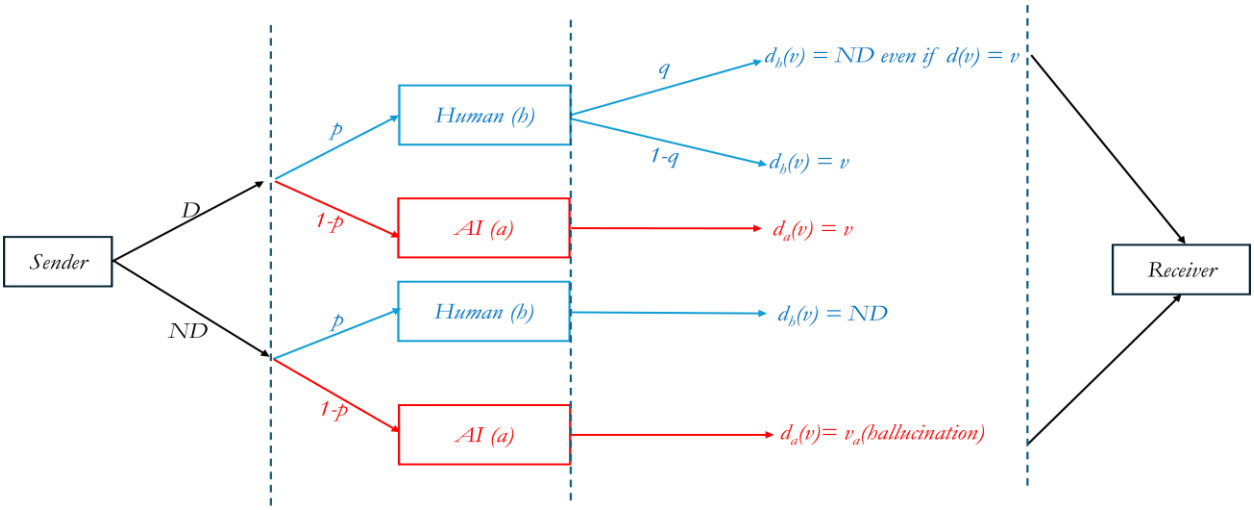


Figure 2: Model Timeline

Definition 1.1 *An equilibrium is given by a disclosure policy* $d(v) \in \{v, \emptyset\}$*, a pricing function after human* $P_h(.)$ *or AI* $P_a(.)$ *processing, and expected prices* $P(v)$ *conditional on disclosure and* $P(\emptyset)$ *conditional on withholding, such that:*

(i) The manager discloses, $d(v) = v$*, when the payoff from disclosure is greater than the expected payoff from non-disclosure:*

$$\underbrace{p\big(qP_h(\emptyset) + (1-q)P_h(v)\big) + (1-p)P_a(v)}_{=P(v)} > \underbrace{pP_h(\emptyset) + (1-p)\int P_a(v)f(v)dv}_{=P(\emptyset)} \qquad (1)$$

and does not disclose, $d(v) = \emptyset$*, when* $P(v) < P(\emptyset)$*;*

(ii) Prices are formed according to the Bayes' rule: (ii.a) $P_h(v) = v$ *and* $P_h(\emptyset) = \mathbb{E}(v \mid d_h(v) = \emptyset)$*, (ii.b)* $P_a(v) = \mathbb{E}(\tilde{v} \mid d_a(\tilde{v}) = v)$*.*

3.2. Analysis

We restrict the attention to a threshold equilibrium (Guttman et al., 2014; Aghamolla and Smith, 2023), in which the receiver discloses if and only if $v \geq \tau$. The marginal discloser $v = \tau$ satisfies the indifference condition that sets Equation (1) at equality. Conditional on this threshold, the non-disclosure price set by the human is

$$P_h(\emptyset) = \frac{q\mu + (1-q)\int_{\underline{v}}^{\tau} vf(v)dv}{q + (1-q)F(\tau)}, \qquad (2)$$

which is equal to the belief in Jung and Kwon (1988) since it is equivalent whether the manager is uninformed and cannot disclose or, as we assume here, the manager does disclose but there is a chance (q) the message is not received.

The determination of the price $P_a(v)$ that follows an AI report is more complex, because the investor must determine this expectation by taking into account that the AI message is potentially garbled. There are two cases to consider.

If $d_a(v) < \tau$, the investor knows that the report must be a hallucination because the manager would never have disclosed a message below the threshold in the equilibrium. Hence, the investor forms the price based on a rational belief that the manager withholds signals below τ, which implies that $P(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} < \tau)$. In this case, the manager realizes no informational rents in expectation, because the garbled message perfectly reveals strategic withholding. Similar to Versano (2021), where non-disclosure conveys significant information when paired with discretionary disclosure, a garbled signal that provides no direct information about fundamentals can still be informative to the investor.

In contrast, if $d_a(v) \geq \tau$, the price must satisfy Bayes' rule, requiring a probabilistic assessment of whether the AI is accurately reporting the sender's signal or misreporting a hallucinated signal due to the manager's non-disclosure. The price is a weighted average over two probabilistic events:

$$P_a(v) = \frac{F(\tau)f(v)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \big(1 - F(\tau)\big)f(v)v}{F(\tau)f(v) + \big(1 - F(\tau)\big)f(v)} = F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \big(1 - F(\tau)\big)v. \quad (3)$$

In the first event, the firm withholds information but the AI hallucinates, with probability $F(\tau)f(v)$. The investor receives an average payoff of $\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)$, which reflects the manager's strategic withholding. In the second event, the firm reports the information, which is then correctly processed by the AI, with probability $\big(1 - F(\tau)\big)f(v)$. Then, the investor receives the payoff v.

In equilibrium, the threshold discloser $v = \tau^*$ must satisfy the indifference condition:

$$p(qP_h(\emptyset) + (1-q)\tau^*) + (1-p)P_a(\tau^*) = pP_h(\emptyset) + (1-p)\int P_a(v)f(v)dv, \quad (4)$$

where the left-hand side is the expected payoff from disclosure, in which case the human may not observe the message with probability q but the AI always observes it. The right-hand side of Equation (4) is the expected payoff from withholding, in which case the human always prices the message at the non-disclosure price but the AI hallucinates a new uninformative random signal leading to a price $P_a(\tilde{v})$.

Lemma 1.1 *The expected price conditional on a hallucination*

$$H \equiv \int P_a(v)f(v)dv = F(\tau^*)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*) + \big(1 - F(\tau^*)\big)\mu \tag{5}$$

is lower than the unconditional mean μ.

Lemma 1.1 further demonstrates that hallucinations do not result in a complete loss of information, even though the observed message is independent of the firm's fundamentals. With probability $F(\tau^*)$, the hallucinated message is below τ^*, which reveals to the user the presence of strategic withholding. As a result, hallucinations imply an expected payoff that is strictly less than the unconditional mean, proportional to the probability that strategic withholding is revealed.

We characterize below the disclosure threshold by simplifying the equations above and using an integration by parts to express the solution in terms of the unconditional mean and the distribution function.

Proposition 1.1 *There exists an equilibrium, and the equilibrium threshold τ^* is given by a solution to*

$$\Big(\underbrace{\frac{(1-p)\big(1-F(\tau^*)\big)\big(q+(1-q)F(\tau^*)\big)}{p(1-q)}}_{\equiv\zeta} + q\Big)(\mu - \tau^*) = (1-q)\int_{\underline{v}}^{\tau^*} F(v)dv. \tag{6}$$

Further, the equilibrium has the following properties:

(i) $\tau_d < P_h(\emptyset) < \tau^* < H < \mu$, where τ_d is the solution with human processing only ($p = 1$);

(ii) If $F(.)$ is logconcave, $\tau^*$is unique.

We consider two special cases with only human or AI processing. In the special case of $p = 1$ (i.e., only human processing), Equation (6) simplifies to the threshold in Jung and Kwon (1988). In comparison, in the special case of $p = 0$ (i.e., only AI processing), the disclosure threshold is $\tau^* = \mu$, so that the manager discloses if and only if the signal is above average. Intuitively, with only AI processing, the manager is always better off resampling a new garbled message, which, on average, will compare favorably to the unconditional mean. Human processing reduces this benefit because there is a non-zero probability that a non-disclosure is detected, which triggers an unfavorable posterior belief $P_h(\emptyset) < \tau^*$.

We prove more generally in Proposition 1.1 that the manager will tend to disclose less with AI processing, that is, $\tau^* > \tau_d$, where τ_d is the solution with human processing only (i.e., $p = 1$). Specifically, relative to Jung and Kwon (1988), Equation (6) contains an additional positive term

$$\zeta = \frac{1-p}{p} \times \left(q + (1-q)F(\tau^*)\right) \times \frac{1-F(\tau^*)}{1-q},$$

which increases the net benefit of withholding.

To gain additional intuition, we decompose ζ into three parts. The first part (i.e., $\frac{1-p}{p}$) is the odds ratio of the AI processing to the human processing, which captures the importance of the distortion due to AI. The second term (i.e., $q + (1-q)F(\tau^*)$) reflects that the AI only facilitates pooling when choosing a non-disclosure, as it is only in this case that the AI muddles the message. In other words, holding all else equal, the effect of AI processing is proportional to the probability of non-disclosure. The third part (i.e., $\frac{1-F(\tau^*)}{1-q}$) captures the interaction between human and AI processing and is the odds ratio of the probability of a manager's voluntary disclosure to the probability of informative human processing (i.e., $1-q$, where q is the probability of human processing failure). As $1 - F(\tau^*)$ becomes larger, the manager is more likely to use the AI's hallucinations to pool a non-disclosure with above-threshold disclosures. However, the greater $1 - q$, the higher the risk that the manager will be identified as engaging in strategic non-disclosure by attentive human processing.

The presence of this last term can be further explained by ranking several possible ex-post outcomes for the manager. The non-disclosure price set by human processing (i.e., $P_h(\emptyset)$) is a weighted average of signals below τ and the unconditional distribution of $\tilde{v}$ when the human involuntarily fails to process information. Hence, humans always respond skeptically to a non-disclosure, that is, $P_h(\emptyset) < \mu$. Consistent with models with uncertain information endowment, strategic reporting involves pooling with below-average types and requires a threshold $\tau^* < \mu$. This property is preserved in the presence of AI processing.

Comparing Equations (3) and (5) and given that we have shown that $\tau^* < \mu$, it must hold that $P_a(\tau^*) < H$. Hence, the marginal discloser, whose preferences determine the equilibrium threshold, is better off when the AI hallucinates, even though they might occasionally receive a price $\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) \leq P_a(\tau^*)$ with probability $F(\tau^*)$. Central to the intuition of our model, the manager tends to withhold information to induce a hallucination. As $1 - F(\tau^*)$ increases, the hallucinated signal becomes less informative because there are fewer revealing messages below the threshold τ^*. Therefore, withholding information becomes more attractive to the marginal discloser.

Combining $P_a(\tau^*) < H$ and the indifference condition in Equation (4), we deduce that $\tau^* > P_h(\emptyset)$. This implies that, unlike in Jung and Kwon (1988), where these terms are equal, the human is relatively more skeptical of non-disclosure in the presence of AI processing. Consequently, the manager tends to withhold information $v \in (P_h(\emptyset), \tau^*)$, which would have attained a higher price under only human processing. The manager avoids disclosing infra-marginal news in the presence of AI processing, expecting that the non-disclosure may be garbled by the AI. Taken together, the disclosure probability is lower with AI processing than with only human processing.

3.3. Comparative Statics

As discussed above, a positive probability of AI processing implies less disclosure than only human processing. Additionally, a model with only AI processing achieves the lowest level of disclosure (i.e., when $p = 0$, $\tau^* = \mu$) and is such that the manager discloses only above-average news. We prove in Corollary 1.1 below the full comparative static in human processing probability p. Specifically, the disclosure threshold (i.e., τ^*) is increasing in the probability of AI processing (i.e., $1 - p$). The intuition is similar to the corner values of p and relies on how hallucinations muddle the non-disclosure signal.

Corollary 1.1 *If $F(.)$ is log-concave, τ^* and $P_h(\emptyset)$ are increasing in the probability of AI processing 1-p.*

Another property of the non-disclosure price under AI processing is that the disclosure threshold $\tau^* > \tau_d$ no longer satisfies the minimum principle, a well-known property in this type of model stating that the equilibrium threshold τ_d minimizes the non-disclosure price (Acharya et al., 2011). The minimum principle is derived from the intuition that an informed manager can always separate by disclosing, and such separation is privately beneficial to the discloser if and only if it reduces the non-disclosure price. Hence, the existence of possibly lower non-disclosure prices would imply that such profitable deviations exist, contradicting the equilibrium behavior. In contrast, disclosers in our model can never perfectly separate because they are endogenously pooled with hallucinations. As a result, the minimum principle breaks down. Consequently, the non-disclosure price must be higher than it would be absent AI processing, and because $P_h(\emptyset)$ is increasing for $\tau \geq \tau_d$, a higher probability of AI processing (monotonically) leads to less skepticism and a higher non-disclosure price.

Finally, the presence of AI processing will tend to increase skepticism toward actual disclosures, since investors anticipate the possibility of a below-threshold hallucination. As shown in Corollary 1.2 below, the price conditional on disclosure and the price response sensitivity to disclosed signals are decreasing in the probability of AI processing.

Corollary 1.2 *The expected price conditional on disclosure* $M(v) \equiv pv + (1-p)P_a(v)$ *and the price response sensitivity* $M'(v)$ *are increasing in* p.

To provide additional intuition, we consider two parametric versions of the model that illustrate how unraveling fails in the presence of AI processing. First, assume that firm value follows a centered uniform distribution $\tilde{v} \sim U(-\sigma, \sigma)$. The price conditional on a human observing non-disclosure can be explicitly written as:

$$P_h(\emptyset) = -\frac{1}{2}\frac{(1-q)(\sigma^2 - \tau^2)}{\sigma + \tau + q(\sigma - \tau)}. \tag{7}$$

Equation (6) implies that τ^* is a solution to a third-order polynomial.[20] In the limit case where the human is almost always able to perfectly process the message (i.e., $q \to 0$), the price becomes:

$$P_h(\emptyset) = \frac{-\sigma + \tau}{2}, \tag{8}$$

which implies that the human always perfectly infers that a non-disclosure is due to strategic withholding. However, having humans perfectly process the message alone is not sufficient to lead to unraveling in our model because there is a probability $1-p$ that the AI processes the signal and the AI hallucinates a signal upon a strategic non-disclosure. Specifically, we solve for τ^*:

$$\tau^* = \sigma\frac{1 - \sqrt{1 + 4p(1-p)}}{2(1-p)}, \tag{9}$$

which decreases in p from $\tau^* = 0$ when $p = 0$ and the information is always processed by an AI, to $\tau^* = -\sigma$ when $p = 1$ and the information is always processed by a human. Hence, in this example, the manager discloses less when there is a higher probability of AI processing.

Second, we consider a normal distribution where $\tilde{v} \sim N(0, 1)$. Figure 3 below illustrates how the disclosure threshold (τ^*) varies with the probability of human processing (p) and the probability of human processing failure (q). The disclosure threshold is decreasing (increasing) in

[20] This polynomial can be simplified to $-(1-p-q+pq)(\tau^*)^3 + (p-2q+pq^2)(\tau^*)^2 + (1+p+q-pq-2pq^2)\tau^* + p(1-q)^2 = 0$.

human processing probability (processing failure). In other words, a greater likelihood of AI processing or imperfect human processing reduces the probability of disclosure. Interestingly, in this specification, the disclosure threshold is approximately symmetric with respect to the two probabilities. This suggests that AI processing has a quantitative effect comparable to the friction in human processing.

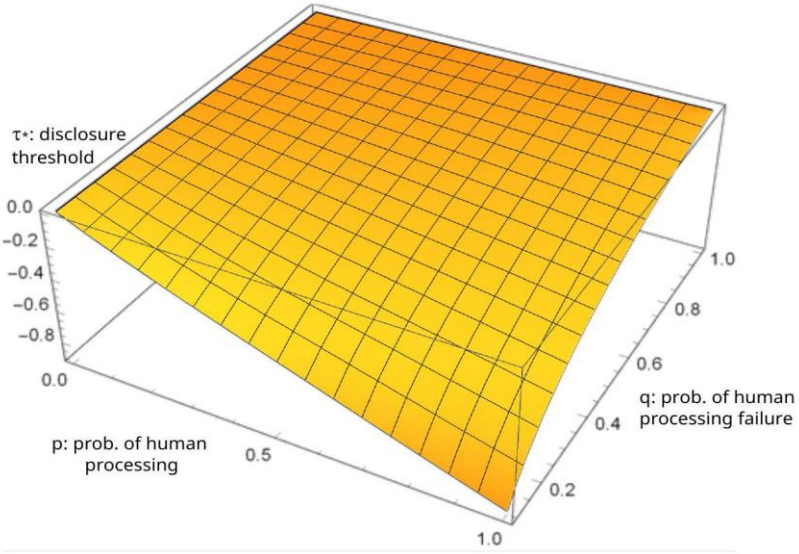


Figure 3: Disclosure Threshold With $\tilde{v} \sim N(0,1)$

IV. Empirical Hypotheses and Sample Construction

4.1. Hypothesis Development

Our model introduces AI, a critical information processing technology, into a disclosure game where an informed manager decides whether to communicate the information. We assume that the investor does not directly observe the disclosure and is subject to information processing constraints (Blankespoor et al., 2020), relying on an information intermediary using either AI or human processing. Our model illustrates the differential impacts of AI-driven versus human-driven information processing on managers' disclosure incentives.

First, we hypothesize that increased use of AI processing reduces firms' voluntary disclosures. This hypothesis stems from the notion that AI is more prone to generating misinformation when substantive information is lacking, such as when a firm does not disclose. Typically, non-disclosure elicits skepticism from investors, who may interpret it as withholding unfavorable news. However, misinformation produced by AI is more easily mistaken for true disclosures.

Furthermore, the value of voluntary disclosure diminishes as investors discount disclosed information due to the increased risk of misinformation. Consequently, higher probabilities of AI processing incentivize firms to withhold information, as non-disclosure leads to more favorable expected price reactions. Our first hypothesis follows:

Hypothesis 1: *Increased use of AI processing by information receivers diminishes firms' incentives to disclose voluntarily.*

Second, the effect of AI processing on voluntary disclosure depends on the potential for AI-generated errors, especially when firm-specific information is unavailable. This potential for errors increases for more complex firms, where the AI may find unrelated information and interpret it as relevant to a firm's fundamentals, particularly if it did not originate from the firm. Conversely, for relatively simple firms, the risk of misinformation is less pronounced because the user is better able to recognize whether there is misinformation. Hence, the impact of AI processing on firms' voluntary disclosure is stronger when the likelihood of hallucinations is higher or when misinformation produces less predictably identifiable signals.[21]

Hypothesis 2: *The impact of AI processing on firms' voluntary disclosure is more pronounced for more complex firms.*

A key assumption in our model is that the use of AI reduces information processing failures. However, a significant empirical challenge is that receivers' information processing failures are inherently unobservable. To overcome this challenge, we rely on the close relation between information processing failures and value-relevant information not being incorporated into the stock price in a timely manner. In this context, processing failures can be empirically captured by the speed at which relevant information is incorporated into the prices. Therefore, we use information processing speed as a manifestation of information processing failure and propose the following hypothesis:

Hypothesis 3: *Increased use of AI processing increases the information processing speed (i.e., reduces human processing failures).*

Next, we shift our focus to the impact of AI processing on investors' beliefs and price reactions. An important implication of our theory is that investors become more skeptical of firms' voluntary disclosures in the presence of AI-generated hallucinations. Because investors must consider the

[21] We derive this hypothesis from the generalized models in Section 6.

potential for hallucinations, they discount observed disclosures. This contrasts with standard disclosure models, where any disclosure reflects the actual evidence and is not discounted. Our fourth hypothesis follows:

Hypothesis 4: *Increased use of AI processing reduces the sensitivity of information users' responses to firms' voluntary disclosures.*

Last, our model predicts that increased use of AI processing raises both non-disclosure thresholds and prices. This hypothesis derives from the theoretical implication of the minimum principle in evidence games (Acharya et al., 2011; Guttman et al., 2014). According to this principle, if an informed sender can disclose all her evidence without friction, the equilibrium disclosure threshold minimizes the non-disclosure price compared to all other possible thresholds. However, the presence of AI processing complicates the disclosure of evidence and moves the equilibrium away from this minimum, thereby increasing the non-disclosure price. In other words, beyond studying the effect of AI on disclosure, we can test the impact of AI processing on non-disclosure prices, which provides an empirical test of the minimum principle in evidence games.

Hypothesis 5: *Increased use of AI processing increases the non-disclosure thresholds and investors' posterior beliefs about the fundamentals of non-disclosing firms.*

4.2. Sample Construction and Measurement of Key Variables

We construct our sample by combining multiple datasets. We use quarterly management forecasts reported in the I/B/E/S Guidance database from January 1, 2021, to December 31, 2023, covering 5,920 U.S. firms. We obtain data on institutional ownership from the Thomson Reuters 13F institutional holdings database. Firm-level characteristics are from Compustat Fundamental Quarterly, and stock prices are from CRSP Security Daily.

To identify sell-side financial analysts' backgrounds, we match analysts surveyed by I/B/E/S with the résumé data from Revelio Labs.[22] We employ fuzzy matching, utilizing both the brokerage houses' and analysts' names, to match analysts on I/B/E/S with the records from Revelio Labs.[23] Our final sample consists of 6,689 sell-side financial analysts located in the United States.

[22] Revelio Labs is a leading provider of labor market analytics, gathering information from professionals' online profiles and resumes on platforms such as LinkedIn.

[23] We note that I/B/E/S only provides abbreviated brokerage house names and the analysts' last names along with their first initials. We use ChatGPT to expand the abbreviated brokerage names into their full names. After the fuzzy matching between I/B/E/S and Revelio Labs, we manually review and exclude incorrect matches.

We proxy for analysts' likelihood of using AI based on their skills and educational histories. Following Frank et al. (2023), we classify analysts as "technical analysts" if they possess technical skills such as machine learning, artificial intelligence, and advanced statistics or if they have majored in technical fields—primarily science, technology, engineering, or mathematics (STEM) subjects.[24] In our sample, 1,944 out of 6,689 analysts (29%) are classified as technical analysts.

Our final dataset comprises 9,866 firm-quarter observations covering 2021Q1 to 2023Q4 after matching datasets and keeping observations with complete data for all key variables. Our sample comprises 1,252 U.S. firms, which are covered by at least one analyst with available resume information and whose management has issued at least one forecast during the sample period. We winsorize all continuous variables, except for return and volatility, at the 1st and 99th percentiles to reduce the influence of outliers.

Table 2 presents the summary statistics. An average firm-quarter in our sample has a market capitalization of $21.17 billion (*Size*) and a book-to-market ratio of 0.48 (*BM*). The average firm-quarter has a leverage ratio of 30.3% (*Leverage*), with negative quarterly earnings 23.4% of the time (*Loss*), and earnings higher than four quarters ago 57.8% of the time (*EPS Increase*). The average firm-quarter has around 11 analysts providing at least one forecast on EPS for the firm over the quarter (*AnalystCover*), with 81.4% of shares held by institutional investors (*InsOwn*) and 36.0% held by the top five institutional investors (*InsOwnTop5*).

Regarding voluntary disclosure, an average firm issues 0.81 management forecasts per quarter on various financial metrics (*MgrForecasts – All*). Of these, 0.17 forecasts pertain to EPS (*MgrForecasts – EPS*) and 0.39 to sales (*MgrForecasts – SALES*). Among analysts matched to valid resume information from Revelio Labs, 18.3% are likely to use AI (*Tech Analyst*). In our sample, 525 firms (41.9%) are covered by at least one technical analyst, while 727 firms (59.1%) are covered exclusively by non-technical analysts.

V. Empirical Analyses

[24] Bick et al. (2024) present survey evidence showing that workers with STEM degrees adopt generative AI in their roles at significantly higher rates than those with other majors. According to the August 2024 wave of the Real-Time Population Survey, 46 percent of STEM-educated workers utilize generative AI at work, compared to 40 percent of individuals with business, economics, or communication majors, and 22 percent of those in other fields, including liberal arts and humanities.

5.1. Testing the Impact of AI Processing on Voluntary Disclosure

Our first hypothesis posits that managers disclose less when the probability of AI processing is higher. To empirically test this, we focus on the introduction of ChatGPT 3.5, a transformative large language model that may positively impact AI processing. We examine differences in the disclosure practices of firms covered by analysts with varying levels of technical expertise from before to after the introduction of ChatGPT 3.5 in 2022Q4. Our identifying assumption is that analysts with greater technical proficiency are more likely to adopt generative AI tools, thereby making managers more aware of the benefits of AI-facilitated information processing and the risks associated with AI-generated misinformation.

Empirically, we estimate the following difference-in-differences design:

$$Mgr\ Forecasts_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \quad (10)$$

where $Mgr\ Forecasts_{i,t}$ is an indicator variable that equals one if the firm *i* issues at least one forecast in quarter *t* and zero otherwise.[25] $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). We follow Abramova et al. (2020) in controlling for other firm-level characteristics that could affect corporate disclosure. Specifically, the control variables include *institutional ownership*, *return*, *loss*, *EPS increase*, *absolute change in EPS*, *leverage*, *size*, *BM*, *return volatility*, and *analyst coverage*. We lag all control variables by one quarter and present variable definitions in Appendix A. We include firm and year-quarter fixed effects to absorb unobserved heterogeneity at the firm level regarding voluntary disclosure decisions and common time-series shocks. We cluster standard errors by firm to account for within-firm correlations over time.

Table 3 reports the results testing the differential impacts of the ChatGPT introduction on quarterly management forecasts from 2021 to 2023 between firms covered by technical analysts and other firms. Firms covered by technical analysts significantly lower their number of management forecasts compared with control firms after the introduction of ChatGPT. Regarding economic magnitudes, the probabilities of issuing at least one forecast on any financial metric, EPS, and sales decrease by 19.8%, 37.0%, and 19.8%, respectively, relative to the sample mean

[25] Our main results remain robust when using the number of management forecasts as an alternative dependent variable and applying a Poisson regression to deal with the count data.

probability. Our findings indicate that managers at firms covered by tech-savvy analysts significantly reduce voluntary disclosure following the introduction of ChatGPT.

Our difference-in-differences design hinges on the crucial assumption that treated firms (those covered by technical analysts) and control firms exhibit comparable disclosure trends in the periods before the introduction of ChatGPT. A potential concern is that firms covered by technical analysts may have been influenced by broader business or technological trends. While we cannot entirely rule out the possibility that such trends coincide with the introduction of ChatGPT, such a scenario would likely exhibit pre-existing patterns in the firms' disclosure practices before ChatGPT's launch. To assess this assumption, we examine the dynamic treatment effects of ChatGPT introduction from 2021Q1 to 2023Q4. Our specification follows:

$$Mgr\ Forecasts_{i,t} = \beta_s \sum_{s=-7\sim+4, s\neq-1} TechAnalyst_i \times D_{s(t)} + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \quad (11)$$

where the control variables are the same as those in Equation (10). $D_{s(t)}$ is a set of indicator variables that take value one if, in quarter *t*, the introduction of ChatGPT is *s* quarters away. For example, $D_{0(t)}$ equals one in 2022Q4 and zero otherwise, while $D_{1(t)}$, $D_{2(t)}$, $D_{3(t)}$, and $D_{4(t)}$ are indicator variables for each of the four quarters after the introduction of ChatGPT. Similarly, $D_{-7(t)},\ldots, D_{-1(t)}$ are indicator variables for each of the seven quarters before the introduction of ChatGPT.

Figure 5 presents our findings. The results do not reject the parallel trends assumption, as there are no significant differences in management forecasts between treated and control firms during the seven quarters preceding the introduction of ChatGPT. In 2022Q4, with the launch of ChatGPT 3.5, the impact of AI processing on voluntary disclosure begins to turn negative, although it remains insignificant. This finding suggests that the introduction of ChatGPT 3.5 in November 2022 does not produce an immediate impact. Notably, the reduction in disclosure by treated firms persists and intensifies in the subsequent quarters, indicating a sustained effect of AI processing on firms' disclosure practices.[26] This pattern is consistently observed across

[26] We note that the introduction of ChatGPT 3.5 coincides with a broader rise in AI technology that offers sufficient flexibility for adoption by market participants. Our difference-in-differences design cannot determine whether the observed effects are specifically attributable to ChatGPT or to other AI-assisted tools that proliferate following its release. Our primary focus is on whether the overall growth in AI processing influences corporate disclosure practices, rather than attributing the effects to any single generative AI tool. As shown in Figure 5, the impact of AI processing intensifies over time, consistent with other AI tools than ChatGPT contributing to this effect.

management forecasts for all financial metrics, EPS, and sales, as shown in Panels A, B, and C.

5.2. How Firm Complexity Affects the Relation Between AI Processing and Voluntary Disclosure

We test the second hypothesis that the impact of AI processing on firms' voluntary disclosure is more pronounced for more complex firms. Intuitively, AI is more prone to misinterpreting unrelated information as relevant to a firm's fundamentals when dealing with more complex firms. In contrast, the risk of misinformation is less significant for simpler firms because inaccuracies are more easily recognized by the AI or the user.

Empirically, we measure firms' complexity using international operations data from the Compustat Historical Segments File, focusing on international operations as a key source of corporate complexity. International operations amplify earnings volatility by exposing firms to additional economic factors, such as currency and political risks, regulatory interventions, and market turbulence (Duru and Reeb, 2002). We predict that the decline in voluntary disclosure for treated firms after the introduction of ChatGPT is concentrated in more complex firms. Our specification follows:

$$\begin{aligned} Mgr\ Forecasts_{i,t} = {} & \beta_1 TechAnalyst_i \times Post_t \times Complexity_i + \beta_2 TechAnalyst_i \times Post_t \\ & + \beta_3 Complexity_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \end{aligned} \tag{12}$$

where $Complexity_i$ is proxied by two measures: (1) whether the firm operates at least one foreign segment outside the U.S., and (2) the number of geographic segments.[27] Table 4 reports the results. Consistent with our hypothesis, the decline in disclosure by treated firms is generally more pronounced for the firms that are more complex. For foreign operations (columns 1 to 3), the interaction term *TechAnalyst* × *Post* × *Complexity* is significantly negative, while *TechAnalyst* × *Post* is generally negative but not statistically significant. These results hold across all management forecasts, EPS forecasts, and sales forecasts. Turning to geographic segments (columns 4 to 6), the incremental treatment effect is much weaker for more complex firms. Taken together, our results imply that the crowding-out effect of AI information processing on firms' voluntary disclosure is

[27] We follow Cohen and Lou (2012) in including only firms where the total sales of all segments account for at least 80% of firm-level sales, ensuring the validity of both measures of firm complexity. Our results remain robust after removing this sample restriction.

significantly stronger for firms with foreign operations, where AI is potentially more likely to confuse unrelated information with firms' actual disclosures.

5.3. Testing the Impact of AI Processing on Information Processing Speed

A key assumption and benefit of AI processing is its ability to reduce information processing failures. However, these failures are empirically unobservable, making it challenging to directly measure their reduction. To address this issue, we utilize the structural model developed by Smith (2024), which allows us to estimate the information processing speed. The strength of this approach lies in its use of daily stock return data to infer unobservable deep parameters, such as the processing speed and incremental informativeness of the earnings announcements, and its ability to disentangle abnormal earnings announcement volatility. By leveraging this model, we can indirectly capture the impact of AI on investors' ability to process information, offering insights into how AI reduces information processing failures in our model.

Our empirical exercise involves measuring the speed at which capital markets process new information from management forecasts. However, the majority of management forecasts are bundled with earnings announcements, as they are generally released within a short time window (Rogers and Van Buskirk, 2013). This means that the identified information processing speed of management forecasts reflects the speed of processing combined earnings announcements and management forecasts.[28] Our identification approach thus has to rely on two scenarios: 1) for firms with management forecasts, Smith's (2024) method captures the speed of processing the bundled management forecasts and earnings announcements, and 2) for firms without management forecasts, it captures the speed of processing earnings announcements alone.[29] Therefore, the comparison between these two scenarios helps isolate the specific speed of processing management forecasts. The model utilizes daily stock return variances as inputs, which capture the perceived uncertainty associated with the information content of EAs and non-earnings sources. Our empirical analysis examines volatility throughout the quarter following an EA to understand the dynamics of information processing, distinguishing periods influenced by EA-related

[28] We note that in our sample, 93.7% of management forecasts are announced within a [-1, +2]-day range relative to earnings announcement dates. Our results remain robust when we restrict the sample to firm-quarters with management forecast dates falling within this [-1, +2]-day window.

[29] We identify earnings announcement (EA) dates following DellaVigna and Pollet (2009) by comparing EA dates from IBES and Compustat, selecting the earlier date when discrepancies arise.

information from those driven by non-earnings information arrivals.

Importantly, a failure in information processing in our model is interpreted as low information processing speed in Smith (2024). We employ the $\widehat{\emptyset}(x)$ statistic estimated from Smith's (2024) structural model. $\widehat{\emptyset}(3)$ ($\widehat{\emptyset}(5)$) represents the fraction of earnings information processed by the market three (five) days after the earnings release, indicating the extent of investor uncertainty reduction within that period. For each firm, we estimate two $\widehat{\emptyset}(x)$: one for the pre-period using four quarters before 2022Q4 and the other for the post-period using four quarters after 2022Q4 (inclusive).[30]

We conduct a difference-in-differences analysis of the changes in information processing speeds for treated firms compared with control firms from before to after the introduction of ChatGPT 3.5 in Nov. 2022. We estimate the following specification:

$$\widehat{\emptyset}(x)_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \quad (13)$$

The treatment sample consists of public firms covered by at least one tech-savvy financial analyst, while the control sample includes public firms without such coverage. The pre-period covers the four quarters prior to 2022Q4, and the post-period consists of the four quarters starting from 2022Q4 (inclusive).

Table 5, Panel A presents the results of our firm-level analysis of the information processing speed of earnings announcements. We report results for three different samples of firms: full sample (columns 1 and 4), firm-quarters with management forecasts (columns 2 and 5), and firm-quarters without management forecasts (columns 3 and 6). First, we do not find significant evidence in the full sample that firms covered by tech-savvy analysts show higher information processing speeds post–ChatGPT introduction than control firms, although the coefficients are directionally consistent with higher processing speed (columns 1 and 4). Second, subsample analyses reveal that information processing speed significantly increases for firm-quarters with management forecasts (columns 2 and 5), while no significant impact is observed for firm-quarters

[30] The average of firm-level measure of information speed ($\widehat{\emptyset}(3)$) is 0.31, which means that around 31% of the information from a typical earnings announcement is processed within the first three days following the announcement. We note that our estimate is lower than the 0.79 reported by Smith (2024) for the full sample. One potential explanation is that our estimation is conducted at the firm level, capturing more idiosyncratic information, whereas Smith (2024) uses a group-level model, which averages out firm-specific idiosyncrasies. Thus, while a significant portion of information is integrated quickly, much of the firm-specific information takes longer to fully influence investor valuations.

without management forecasts (columns 3 and 6).[31] In other words, the processing speed of bundled management forecasts and earnings announcements increases, whereas the processing speed of stand-alone earnings announcements does not significantly change. These findings suggest that AI processing may facilitate the market's interpretation of voluntary disclosures.

5.4. Testing the Impact of AI Processing on the Informativeness of Earnings Announcements

Another benefit of the approach by Smith (2024) is that it jointly estimates the informativeness of earnings, including immediate informativeness (*I*), earnings horizon (*H*), and adjusted incremental informativeness ($(\widehat{\pi_{NR} - \pi_R})_{adj}$). The estimates are used to infer the effect of EAs in reducing investor uncertainty. We use these measures to assess the changes in the informativeness of voluntary disclosures post-ChatGPT introduction. While our model does not have explicit predictions on incremental informativeness, the test on informativeness allows us to rule out alternative factors driving the reduced managerial forecasts. For example, Verrecchia (1983) suggests that voluntary disclosure decreases in managers' relative information advantage over external investors. Better-informed investors, facilitated by AI utilization, could lead to more informativeness of stock prices, thereby reducing the need for voluntary disclosure. While this alternative explanation does not align with Hypotheses 2, 4, and 5, which suggest evidence of information muddling due to AI, it proposes an alternative mechanism with a testable implication. To address both the stand-alone question of whether generative AI enhances the informativeness of stock prices, given its potential crowding-out effect, and to assess this alternative explanation, we examine the impact of AI processing on total information content.

We conduct a difference-in-differences analysis to test changes in the informativeness of earnings announcements for treated and control firms. We report the results in Table 5, Panel B, for three different samples of firms: full sample (columns 1, 4, and 7), firm-quarters with management forecasts (columns 2, 5, and 8), and firm-quarters without management forecasts (columns 3, 6, and 9).[32] First, we do not find significant evidence in the full sample that AI processing affects the informativeness of earnings announcements. Second, focusing on the subsample of firm-quarters with management forecasts, we find significant evidence that AI

[31] Regarding economic magnitudes, the two processing speed measures, $\widehat{\emptyset}(3)$ and $\widehat{\emptyset}(5)$, increase by 47.6% and 60%, respectively, relative to the sample mean of firm-quarters with management forecasts.

[32] We note that the firm-level estimates of $\hat{I}$ and $(\widehat{\pi_{NR} - \pi_R})_{adj}$ have larger standard errors than the estimates in Smith (2024).

processing increases the earnings horizon (column 2). Additionally, AI processing has insignificant but positive effects on immediate informativeness (column 5) and adjusted incremental informativeness (column 8). These results suggest that AI's enhanced processing capacity, compared to human processing, facilitates the incorporation of information from voluntary disclosures into prices. This finding is consistent with our results in Table 5 Panel A that information processing speed increases in the post-period for firms covered by tech-savvy analysts. Last, we do not find a significant impact of AI processing on the informativeness of firm-quarters without management forecasts.[33] Overall, our results suggest that the introduction of AI processing tools, such as ChatGPT, potentially enhances the information quality of earnings announcements for firms that provide management forecasts. However, for firms that withhold information, our analyses imply that the risk of misinformation potentially offsets the benefit of increased processing speed, leaving the net effect on informativeness ambiguous.

5.5. Testing the Sensitivity to Disclosure using Analyst Forecast Revisions Around Management Forecasts

We examine whether the probability of AI processing influences the sensitivity of market participants' reactions to disclosures. In our model, increased AI processing elevates the disclosure threshold, thereby increasing the likelihood of misinformation. As market participants may not differentiate real versus hallucinated disclosures, they become more skeptical, leading to muted reactions to firms' real disclosures.[34]

To test this prediction, we examine analyst forecast revisions around the issuance dates of management forecasts to capture market participants' reactions to management forecasts (Rogers and Van Buskirk, 2013; Hsu and Wang, 2021). Analyst forecast revisions serve as a high-frequency measure of how new information from management forecasts is incorporated into market participants' expectations of firms' future performance.[35] Our specification follows:

---

[33] In untabulated analyses, we employ the empirical framework of Lundholm and Myers (2002) to assess the forward-looking information embedded in current stock prices on firms' future performance. Similar to Lundholm and Myers (2002), we regress current quarterly stock returns on lagged quarterly earnings, current earnings, and future earnings, controlling for future quarterly stock returns. Consistent with our findings in Table 5, Panel B, we do not observe significant increases in the forward-looking information of stock prices for treated firms following the introduction of ChatGPT.

[34] We acknowledge that empirically, we have only one stock price per firm, making it challenging to disentangle stock price reactions resulting from AI processing from those due to human processing.

[35] An alternative is to use price responses to management forecasts. However, unlike forecast revisions, price responses tend to be much more volatile and not solely driven by current forecasts.

$$AFRev_{i,t} = \beta_1 TechAnalyst_i \times Post_t \times MFNews_{i,t} + \beta_2 TechAnalyst_i \times Post_t + \beta_3 Post_t \times MFNews_{i,t} + \beta_4 TechAnalyst_i \times MFNews_{i,t} + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \quad (14)$$

where *AFRev* is calculated as the difference between the first analyst forecast issued after the managerial guidance date and the last analyst forecast issued before the managerial guidance date, scaled by the firm's stock price three trading days before the guidance date. *MFNews* proxies for the new information in management forecasts and is calculated as the difference between managerial guidance and the last analyst forecast prior to the guidance date, scaled by the firm's stock price three trading days before the guidance date. We include firm and year-quarter fixed effects and cluster standard errors by firm.

Table 6 presents our findings. We utilize the full sample in columns 1 and 4, the subsample with positive *MFNews* in columns 2 and 5 (i.e., management forecast is higher than the last analyst forecast), and the subsample with negative *MFNews* in columns 3 and 6 (i.e., management forecast is lower than the last analyst forecast). First, columns 1 to 3 show that analyst forecast revisions respond strongly to management forecast news, indicating that analysts promptly incorporate management forecast information into their forecasts. Second, analysts' reactions to management forecasts are significantly muted for firms that are covered by technical analysts in periods after the ChatGPT introduction (i.e., *TechAnalyst*×*Post*×*MFNews* is significantly negative in column 4). Third, this effect is particularly more pronounced for positive *MFNews* and not significant for negative *MFNews* (i.e., *TechAnalyst*×*Post*×*MFNews* is significantly negative in column 5 but not in column 6), which potentially suggests that it is more challenging to differentiate positive disclosures from AI-generated hallucinations than negative ones. Our findings indicate that the widespread adoption of AI processing following ChatGPT's introduction—particularly among firms covered by technical analysts—leads to an increase in AI-hallucinated disclosures blending with real disclosures, thereby reducing the market's responsiveness.

5.6. Testing The Minimum Principle

Our final hypothesis posits that higher probabilities of AI processing elevate non-disclosure thresholds and prices. In evidence games where a sender can disclose all available evidence upon receipt, a common property known as the "minimum principle" dictates that the equilibrium disclosure threshold minimizes the non-disclosure price relative to all other possible thresholds. However, the introduction of AI processing shifts the equilibrium away from this minimum,

resulting in increased non-disclosure prices. We evaluate the impact of AI processing on investors' pricing of non-disclosing firms, thereby providing an empirical test of the minimum principle in voluntary disclosure models.

Empirically, we assess investors' pricing of non-disclosing firms by restricting our analysis to a subsample of firms without management forecasts. Our specification follows:

$$ND\ Threshold\ or\ ND\ Prices_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t}, \quad (15)$$

where *ND Threshold* is proxied by future *EPS*, which is earnings per share for firm *i* in quarter *t+1*. *ND Prices* are proxied by current *Price-to-Earnings Ratio (PE)* and *Tobin's Q*, where *PE* is the price at the end of the quarter *t* divided by earnings per share for firm *i* in quarter *t*. *Tobin's Q* is defined as the market equity plus long-term debt and short-term debt in quarter *t* scaled by book assets for firm *i* in quarter *t*. We include firm and year-quarter fixed effects and cluster standard errors by firm.

Table 7 presents the results. We find that among non-disclosure firms, the ones covered by technical analysts exhibit significantly higher future *EPS*, current *PE*, and current *Tobin's Q*, compared with the control firms after the introduction of ChatGPT. These findings are consistent with the hypothesis that increased AI processing shapes investors' posterior beliefs about the fundamentals of non-disclosing firms.

5.7. Robustness Checks

We conduct several robustness checks. First, our main analyses focus on managers' quarterly forecasts, and we examine whether our results hold for annual management forecasts. Annual forecasts are less sticky than quarterly forecasts, involve significantly higher levels of uncertainty, and provide more information to the market. Consequently, annual management forecasts may better align with voluntary disclosure theory, which posits that investors must contend with substantial uncertainty about firm fundamentals. However, unlike interpretations that suggest managers have an uncertain information endowment—typically more relevant for longer horizons—our baseline model assumes that managers are always informed, while humans may not always process the information effectively. Nevertheless, we redo the analyses using annual management forecasts from 2018 and 2023 and present the results in Table 8. Consistent with our previous findings, we observe a significant decrease in annual management forecasts in the years following the introduction of ChatGPT for firms covered by technical analysts.

Second, we examine whether our main findings are robust to using an alternative definition of management forecasts and treated firms. Specifically, we use the number of management forecasts as the dependent variable and define a continuous treatment variable as the percentage of analysts with technical skills at the firm level. Table 9 shows that firms with a higher proportion of technical analysts issue fewer management forecasts following the introduction of ChatGPT. These results align with our earlier findings based on indicator variables for management forecasts and treated firms.

VI. Model Extensions and Discussion of Hypotheses

We aim to present straightforward intuition through a simple model, which is not intended to offer a general perspective on AI information processing. Nevertheless, we note that certain assumptions about the information structure can be altered or relaxed without significant qualitative changes to our analysis.

6.1. The Investor Does Not Know Whether Processing Is by AI or Human

Our baseline model assumes that the investor knows whether the message is processed by a human or an AI. This assumption reflects situations where investors use specific tools for information processing or have established relationships with their information intermediaries, which is particularly relevant for institutional investors who maintain long-term relationships with sell-side analysts from different brokerage houses. However, it is also possible to consider the opposite scenario where the investor lacks knowledge about the intermediary's processing mechanism. While this introduces additional complexity to the model, we explore this case further and show that, under certain conditions, our analysis still holds.

Denoting the price function as $P(.)$, Equation (2) for the non-disclosure price $P(\emptyset) = P_h(\emptyset)$ is unchanged because only human processing is consistent with no message being observed. Similarly, the price remains $P(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)$ if $v \leq \tau$ because the user can infer in equilibrium that this disclosure is a hallucination following strategic withholding. However, the price $P(v)$ conditional on a disclosure $v > \tau$ must now incorporate the possibility of human processing:

$$P(v) = \frac{p(1-q)f(v)v + (1-p)\big(F(\tau)f(v)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \big(1-F(\tau)\big)f(v)v\big)}{p(1-q)f(v) + (1-p)\Big(F(\tau)f(v) + \big(1-F(\tau)\big)f(v)\Big)}$$

$$= \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + 1_{v \geq \tau}\left(1 - \frac{1-p}{1-pq}F(\tau)\right)\big(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)\big) \quad (16)$$

Compared to the baseline, the user is less skeptical toward AI disclosures because these could originate from a human and, vice-versa, more skeptical toward human disclosures, which can no longer be distinguished from the AI. The indifference condition at the disclosure threshold becomes:

$$pqP(\emptyset) + (1-pq)P(\tau^*) = pP(\emptyset) + (1-p)\int P(v)f(v)dv. \quad (17)$$

Note that a greater probability that the human is subject to a friction makes the user more skeptical of all disclosures as they are more likely to be from an AI hallucination. Simple algebraic manipulations similar to Lemma 1.1 yield

$$H \equiv \int P(v)f(v)dv = \frac{1-p}{1-pq}F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \left(1 - \frac{1-p}{1-pq}F(\tau)\right)\mu, \quad (18)$$

which implies that, for a given disclosure threshold, the hallucination is now more beneficial to the manager given the higher weight on μ and, for any threshold, the value of withholding is increased relative to the baseline information structure:

$$pP(\emptyset) + (1-p)\int P(v)f(v)dv > pP_h(\emptyset) + (1-p)\int P_a(v)f(v)dv. \quad (19)$$

Vice-versa, a similar exercise implies the inequality:

$$pqP(\emptyset) + (1-pq)P(\tau^*) > p(qP_h(\emptyset) + (1-q)\tau^*) + (1-p)P_a(\tau^*), \quad (20)$$

so that there is also a greater benefit to disclosure in the baseline information structure for the marginal discloser. Put differently, the lack of knowledge by the investor further muddles the message by jointly increasing the payoff to non-disclosure and the payoff to the marginal disclosure.

The comparative statics of the probability of human processing is more complicated in this context. Specifically, if $P(\emptyset)$ is high relative to the (discounted) $P(\tau^*)$, which occurs when human processing is sufficiently ineffective (i.e., when q is large), the manager will have stronger incentives to strategically withhold information. It is important to note that this scenario does not

arise when the probability or efficiency of human information processing is large because, in these cases, $P(\emptyset)$ is compared directly to τ^*.

To develop more intuition, we analytically examine the setting where the human is not subject to processing frictions (i.e., $q = 0$) and demonstrate that our baseline conclusions remain valid. Consequently, by continuity, these conclusions also hold when human processing is sufficiently effective. We show in the Appendix that the equilibrium is characterized by

$$(1-p)(\mu-\tau^*)F(\tau^*) = p\int_{\underline{v}}^{\tau^*} F(v)dv, \tag{21}$$

which closely resembles the result of Jung and Kwon (1988), except that $1-p$ now represents the probability of AI processing, rather than an information endowment friction. Moreover, the left-hand term is multiplied by $F(\tau^*)$. This additional term indicates that a hallucination is informative when sending a message below τ^*, implying that the probability of disclosure is higher than in a comparably calibrated model with uncertain information endowment (i.e., a model where there is no hallucination, but the probability of information endowment is p). Furthermore, it can be readily verified that the equilibrium is unique if $F(\cdot)$ is log-concave, and that an increase in the probability of human processing p leads to higher disclosure.

When $q > 0$, an additional factor influences the equilibrium: increased human processing raises the likelihood of achieving the (potentially high) non-disclosure price $P(\emptyset)$ and avoiding the AI's processing of adverse news. In the limit, as q approaches one, the non-disclosure price $P(\emptyset)$ converges to the unconditional mean μ, which is strictly higher than $P(\tau^*)$. Thus, in this context, more human processing can crowd-out voluntary disclosure. Figure 4 illustrates the disclosure threshold for the case where $\tilde{v} \sim N(0,1)$. While the surface is generally flat or decreasing in p, there is a region in the top right corner where the threshold (slightly) increases with the probability of human processing, consistent with the above intuition. In summary, this analysis suggests that in environments where all of the following conditions are met: (i) human processing is sufficiently imperfect, (ii) the probability of human processing is not excessively high, and (iii) investors are entirely unaware of the processing method, the probability of disclosure may decrease with increased human processing.

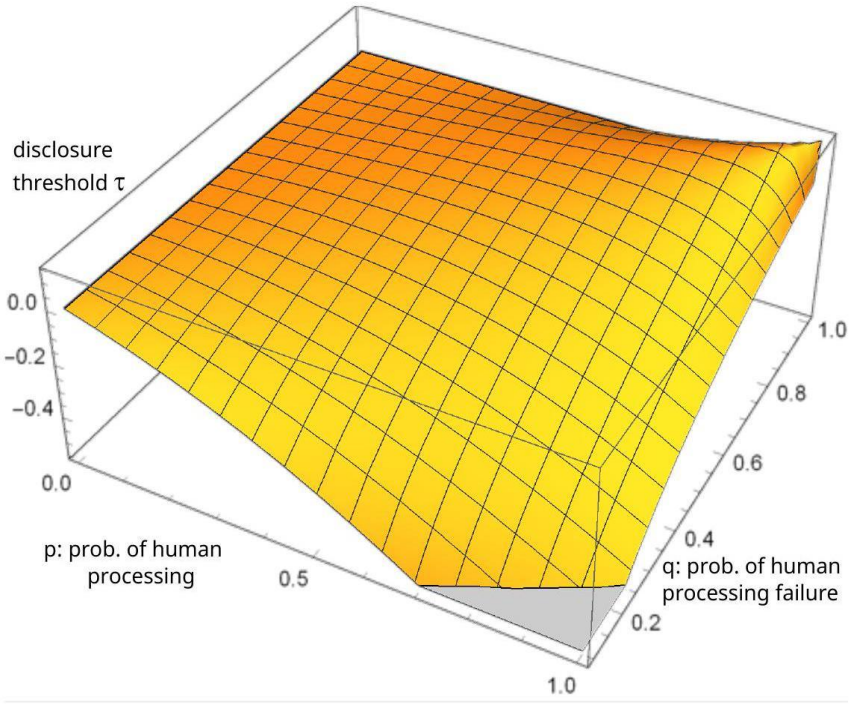


Figure 4: Disclosure Threshold With $\tilde{v} \sim N(0,1)$: Unknown Processing

6.2. Varying the Probability of AI Hallucination Upon Non-Disclosure

We assume in the baseline model that the AI always generates a garbled signal when non-disclosure occurs. This assumption ensures a model structure where only human processing can interpret non-disclosure. However, this may imply that the reduction in voluntary disclosures occurs not only because the AI cannot understand the absence of information, but also because the AI is extremely poor at processing information.

To address this question, we modify the model so that the AI, upon non-disclosure, hallucinates a garbled signal unrelated to the firm's fundamentals with probability $\rho \in (0,1)$.[36] With probability $1-\rho$, the AI does not hallucinate, and we explore two different formulations. In the first formulation, we assume that the AI independently identifies the information even if it has not been disclosed. This variation represents an ideal scenario where the AI has access to more information than humans in both disclosure and non-disclosure situations, while still preserving the key aspect that human processing is better than AI processing at understanding the absence of

[36] The special case $\rho = 1$ corresponds to the baseline model.

information.[37] In the second formulation, we assume that the information known to the manager is not accessible to the AI, so the AI observes a non-disclosure with probability $1-\rho$.

For the first formulation, generalizing Equation (3), the price conditional on a report v is

$$P_a(v) = \frac{F(\tau)\rho f(v)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \big(1-F(\tau)\big)(1-1_{v<\tau}\rho)f(v)v}{F(\tau)\rho f(v) + \big(1-F(\tau)\big)(1-\rho 1_{v<\tau})f(v)}, \tag{22}$$

such that infra-marginal reports $v < \tau$ are more likely to originate from a hallucination and receive a lower price. We solve for the indifference condition as follows:

$$p(1-q)\big(\tau^* - P_h(\emptyset)\big) = \frac{\rho(1-p)\big(1-F(\tau^*)\big)}{1-(1-\rho)F(\tau^*)}(\mu-\tau^*). \tag{23}$$

The disclosure threshold is similar to the baseline model (i.e., with $\rho = 1$) except that the equivalent probability of AI when mapping to the baseline is $p' > p$ with

$$\frac{1-p'}{p'} = \frac{1-p}{p}\frac{\rho}{1-(1-\rho)F(\tau^*)}, \tag{24}$$

so that a lower probability of hallucination is "equivalent" to a model with more human processing and thus tends to feature a higher probability of disclosure. This is intuitive because making the AI more effective tends to reduce the disclosure frictions and moves the main prediction toward unraveling.

From this observation, one might conjecture that a sufficiently low probability of hallucination would reverse the crowding out effect of AI processing and unambiguously improve the information environment. Indeed, it is always the case that a lower ρ increases the information released by the AI conditional on non-disclosure. When $\rho \to 0$, the AI becomes fully informative. However, this does not remove the crowding-out of voluntary disclosure: as we note next (subject to a unique equilibrium), the disclosure threshold remains *decreasing* in the probability of human processing regardless of ρ. The intuition for this result is that any hallucination pools below-average news with above-average news, versus average news in the case of human processing. As such, any level of hallucination pushes the disclosure threshold above the level that would occur under human processing.

[37] It is important to recognize that in a voluntary disclosure model, the information disclosed may not be exclusive to the manager. Instead, it may consist of data that the manager has collected, which is also part of the public record but is difficult to find, access, or process. For example, detailed knowledge of the industry, competitors, or new innovations. Consequently, AI may be able to obtain this undisclosed information more effectively than humans. We evaluate this alternative scenario in the first formulation.

In the second formulation, we assume that the manager's signal is entirely private and cannot be accessed by the AI from other sources in the event of non-disclosure. The main difference with this setting is that the AI will assign a price:

$$P_a(\emptyset) = P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) \tag{25}$$

conditional on non-disclosure or a disclosure of $v < \tau$, because this fully reveals strategic withholding. In turn, this implies that, in the absence of hallucination, the AI imposes a strongly skeptical belief with probability $1 - \rho$ toward a non-disclosure - in fact, more skeptical than human processing $P_h(\emptyset)$. This heightened skepticism serves as a new incentive for managers to increase disclosure.

Using similar steps to solve for $P_a(v)$ for $v \geq \tau$ and H, and then solving for the indifference condition yields:

$$\frac{(1-p)(1-F(\tau))}{1-(1-\rho)F(\tau)}(\rho\mu + (1-\rho)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*) - \tau^*) = p(1-q)\big(\tau^* - P_h(\emptyset)\big), \tag{26}$$

which we show in the Appendix, subject to the solution remaining unique, implies that τ^* monotonically converges to τ_d as p converges to one. Evaluating at $p = 0$, the comparative static in p thus depends on the solution τ_0^* relative to τ_d where

$$\rho\mu + (1-\rho)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau_0^*) = \tau_0^* \tag{27}$$

and is such that the disclosure threshold is decreasing in p if and only if ρ is sufficiently large or q is sufficiently low. In other words, this result suggests that the crowding-out effect of AI processing on voluntary disclosure depends on the relative quality of AI processing compared with human processing. Specifically, the crowding-out effect occurs if the AI hallucination problem is sufficiently severe relative to human processing failure.

6.3. The Manager is Informed About the Choice of Information Processing

In our baseline model, we assume the manager has common knowledge of human processing probability p but does not know with certainty how their information will be processed, enabling a more thorough exploration of the interaction between AI and human processing. However, this assumption is not essential to our hypothesis. If the manager is informed about AI processing, the problem can be framed as a game where, with probability p, the manager employs the Jung and Kwon (1988) threshold $\tau^* = \tau_d$, and with probability $1 - p$, the manager adopts the AI-only threshold $\tau^* = \mu$. Consequently, the probability of disclosure increases with p and thus decreases

with AI processing (i.e., $1-p$). The main difference is that the threshold becomes random and dependent on AI processing.

An interesting variation of this information structure may arise if the manager endogenously chooses the type of information processing. While there may be additional legal considerations in delegating the disclosure decisions to AI, certain price considerations within the model may still be examined. Specifically, we assume that, upon observing v, the manager can choose between AI and human processing. One challenge is that equilibria with a single type of processing can be sustained, as long as investors believe that any off-equilibrium processing is chosen only by firms with sufficiently unfavorable private information.

One manner to rule out such off-equilibrium forcing beliefs is to restrict the attention to equilibria in which all messages that need interpretation are on the equilibrium path, which rules out knife-edge equilibria in which the probability of hallucination or human processing is zero. In addition, noting that there would be no reason for the manager to condition their processing on private information when not disclosing it (since this renders it irrelevant), we assume that the manager always chooses AI when not disclosing. In summary, we consider an equilibrium in which information below τ is not disclosed and processed by AI, while given $v \geq \tau^*$, the manager can choose over AI versus human processing.

Under AI processing, the update is given by Equation (3) and $P(v) = F(\tau^*)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*) + \big(1 - F(\tau^*)\big)v$, which reduces the price sensitivity to the signal by $\big(1 - F(\tau^*)\big)$, whereas under human processing, the price is $P_h(v) = qP_h(\emptyset) + (1-q)v$. The resulting mathematics of the model thus present a strong parallel to Aghamolla and Smith (2023), where a manager chooses between communication mechanisms with different price sensitivities and intercepts. Specifically, when $F(\tau^*) < q$, AI processing can be interpreted in the model by Aghamolla and Smith (2023) as a "complex" disclosure which yields a higher price sensitivity but a greater discount on level due to the possibility of hallucinations. A non-disclosure paired with AI processing "obfuscates" information and is chosen by managers with bad news, while an "informative" disclosure with AI processing ensures that a greater fraction of the information is incorporated into the price. Vice-versa, a "simple" human disclosure may involve a probability q of information loss. As a result, the equilibrium involves both low and high signals being conveyed via AI, while intermediate

signals use human processing. Vice-versa, if $F(\tau^*) > q$, better news tends to rely on human processing.[38]

6.4. Discussion on the Distribution of Misinformation Signals

In our baseline model, we calibrate the distribution of AI-hallucinated signals to match the true distribution of firm fundamentals. This calibration ensures that hallucinations are neither systematically favorable nor unfavorable and is based on the assumption that large language models are generally pre-trained to have a realistic prior. If hallucinations were to produce more favorable or unfavorable news more frequently, AI processing could mechanically alter the probability of disclosure independently of the manager's strategic considerations.

We investigate alternative calibrations of AI-hallucinated signals. More generally, suppose that hallucinations draw a garbled message from a distribution with density $g(\cdot)$, resulting in an AI-generated price:

$$P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*) + 1_{v \ge \tau^*} \frac{\big(1 - F(\tau^*)\big) f(v)}{F(\tau^*) g(v) + \big(1 - F(\tau^*)\big) f(v)} \big(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*)\big), \quad (28)$$

which generates more negative beliefs for events more likely to arise from a hallucination.

For illustrative purposes, consider a scenario where hallucinations are sufficiently biased toward unfavorable outcomes, such that the mass of $g(\cdot)$ is concentrated below the disclosure threshold τ^*. In this case, hallucinations lead to $H = P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*)$, as they fully reveal the strategic withholding of information. The indifference condition is then given by:

$$p(qP_h(\emptyset) + (1-q)\tau^*) + (1-p)\tau^* = pP_h(\emptyset) + (1-p)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*), \quad (29)$$

which simplifies to

$$(1 - pq)(P_h(\emptyset) - \tau^*) = (1-p)\big(P_h(\emptyset) - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*)\big). \quad (30)$$

Since the right-hand side is positive, the non-disclosure price $P_h(\emptyset)$ now exceeds the disclosure threshold τ^*. Given that $P_h(\emptyset)$ follows a U-shaped curve with a minimum at $\tau_d = P_h(\emptyset)$, it then follows that $\tau^* < \tau_d$: the likelihood of disclosure increases compared to only human processing. In other words, hallucinations tend to be fully revealing, and AI processing provides more

[38] In an equilibrium where the marginal discloser uses AI, we have shown in the baseline model that $\tau^* = \mu$. Further, for this equilibrium not to be AI-only, it must hold that $q < F(\mu)$, indicating that $q < F(\mu)$ is a sufficient condition for the existence of an equilibrium where sufficiently good news is processed by humans.

informational content than human processing, pushing the disclosure threshold toward unraveling.[39]

Importantly, hallucinations that are biased toward favorable outcomes do not necessarily reduce the likelihood of disclosure. For instance, if hallucinations consistently imply a specific value v_0, even if v_0 approaches $\bar{v}$, they become fully revealing. In summary, hallucinations that are more easily identifiable on the equilibrium path, either because they are unfavorable or predictable, will, in general, increase the value of AI for information processing and increase voluntary disclosure.

VII. Conclusion

This paper explores the trade-off between the benefits of enhanced information processing by AI and the potential drawbacks posed by misinformation. Our analysis starts with a disclosure game where firms, equipped with information about their fundamentals, must decide whether to disclose or withhold it. We depart from traditional models (e.g., Dye, 1985) by incorporating misinformation caused by the use of AI by market participants. Unlike human processing, AI is not limited by capacity but is prone to generating misleading signals when information is withheld. We obtain a set of new predictions from the model: while AI can enhance the processing of disclosed information, its potential for misinformation discourages voluntary disclosure, encouraging strategic non-disclosure. This crowding-out effect is driven by the potential for hallucination to camouflage a strategic non-disclosure. Users make a Bayesian correction to observed signals above the disclosure threshold, reducing their effect on firm value and the payoff from disclosure.

Next, we employ an identification strategy to test our predictions empirically. Specifically, we use OpenAI's launch of ChatGPT 3.5 in November 2022 as a significant advancement in AI-facilitated information processing. By examining analysts' propensity to adopt AI processing based on their educational background, we classify firms covered by these tech-savvy analysts as

[39] If q were zero, unraveling implies $\tau^* = \underline{v}$. This case can only apply if the distribution of hallucinated messages does not have full support. This example is intended to illustrate the (potential) informational value of hallucinations and demonstrates that, on its own, misinformation does not necessarily reduce information and, in certain environments, could lead to unraveling.

treated firms, while those not covered are assigned to the control group, aligning with our theoretical model on how information receivers rely on generative AI.

We apply the structural approach developed by Smith (2024) to document a positive effect of ChatGPT 3.5 in reducing processing failures, on average, for the treatment group, with an impact concentrated in firms with voluntary disclosures. This finding highlights the beneficial effect of AI processing. Then, we find that treated firms exhibit an economically significant reduction in providing managerial forecasts in the post-ChatGPT 3.5 period. This suggests that information users' greater reliance on generative AI lowers firms' propensity to disclose information voluntarily. We strengthen the link between our model and empirical tests by showing that 1) firm complexity further aggravates the crowding out effect of AI on firms' disclosure, 2) non-disclosure treated firms exhibit an increase in disclosure threshold and a higher share price in the post-ChatGPT 3.5 period (i.e., minimum principle suggested by Acharya et al. (2011) and Guttman et al. (2014)), and 3) the crowding-out effect further manifests in the analysts' reduced responses to treated firms' disclosures.

An important caveat for our theoretical model is that it addresses only one, albeit presumably important, trade-off rather than incorporating all possible mechanisms. Specifically, we conceptualize a tension between the risk of misinformation and an advantage of generative AI in information processing capacity. The crowding-out effect on firms' information supply serves as a cautionary tale, hinting at outcomes where digital advancements could lead to a decline in information provision. At the same time, additional research is needed to capture other important mechanisms through which AI affects information processing, as its effects extend well beyond misinformation. For example, greater access to AI may asymmetrically impact awareness and processing, level the playing field, enhance liquidity by making AI more accessible to unsophisticated investors, or correlate returns across different firms through the use of common information processing tools. While these important questions are beyond our current focus, our primary message is that AI involves trade-offs that may not always, or necessarily, lead to improvements in the information environment.

References

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

Acharya, V. V., Demarzo, P., & Kremer, I. (2011). Endogenous information flows and the clustering of announcements. *American Economic Review, 101*(7), 2955–2979.

Aghamolla, C., & Smith, K. (2023). Strategic complexity in disclosure. *Journal of Accounting and Economics, 76*(2-3), 101635.

Aksitov, R., Chang, C. C., Reitter, D., Shakeri, S., & Sung, Y. (2023). Characterizing attribution and fluency tradeoffs for retrieval-augmented large language models. *arXiv preprint arXiv:2302.05578*.

Banerjee, S., Marinovic, I., & Smith, K. (2024). Disclosing to informed traders. *The Journal of Finance, 79*(2), 1513-1578.

Bergstrom, T., & Bagnoli, M. (2005). Log-concave probability and its applications. *Economic Theory, 26*, 445–469.

Bertomeu, J., Cheynel, E., & Cianciaruso, D. (2021). Strategic withholding and imprecision in asset measurement. *Journal of Accounting Research, 59*(5), 1523–1571.

Bertomeu, J., Hu, P., & Liu, Y. (2023). Disclosure and investor inattention: Theory and evidence. *The Accounting Review,98*(6), 1–36.

Bertomeu, J., Liang, Y., & Marinovic, I. (2023). A primer on structural estimation in accounting research. *Foundations and Trends® in Accounting, 18*(1–2), 1-137.

Bertomeu, J., Lin, Y., Liu, Y., & Ni, Z. (2024). Capital market consequences of generative AI: Early evidence from the ban of ChatGPT in Italy. *Available at SSRN 4452670*.

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics, 50*(2-3), 296-343.

Bick, A., Blandin, A., & Deming, D. J. (2024). The Rapid Adoption of Generative AI (No. w32966). National Bureau of Economic Research.

Biden, J. R. Jr. (2023, October 30). Executive Order on the safe, secure, and trustworthy development and use of artificial intelligence. The White House. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/

Blankespoor, E., Dehaan, E., Wertz, J., & Zhu, C. (2019). Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs. *Journal of Accounting Research, 57*(1), 53-84.

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

Blume, A., Board, O. J., & Kawamura, K. (2007). Noisy talk. *Theoretical Economics, 2*(4), 395–440.

Bourveau, T., Breuer, M., & Stoumbos, R. (2020). Learning to disclose: Disclosure dynamics in the 1890s streetcar industry. Available at SSRN 3757679.

Cao, S., Jiang, W., Yang, B., & Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of AI. *The Review of Financial Studies, 36*(9), 3603-3642.

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

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance*, *64*(2), 709-749.
Dickhaut, J., Ledyard, M., Mukherji, A., & Sapra, H. (2003). Information management and valuation: an experimental investigation. *Games and Economic Behavior*, *44*(1), 26–53.
Dong, Y., Li, O. Z., Lin, Y., & Ni, C. (2016). Does information-processing cost affect firm-specific information acquisition? Evidence from XBRL adoption. *Journal of Financial and Quantitative Analysis*, *51*(2), 435-462.
Duru, A., & Reeb, D. M. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review*, *77*(2), 415–433.
Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research, 23*(1), 123–145.
Einhorn, E. (2018). Competing information sources. *The Accounting Review*, *93*(4), 151-176.
Frank, M. Z., Gao, J., & Yang, K. (2023). Behavioral Machine Learning? Computer Predictions of Corporate Earnings also Overreact. *arXiv preprint arXiv:2303.16158.*
Frenkel, S., Guttman, I., & Kremer, I. (2020). The effect of exogenous information on voluntary disclosure and market quality. *Journal of Financial Economics*, *138*(1), 176-192.
Frenkel, S., Guttman, I., & Kremer, I. (2024). Strategic disclosure with fake and real news. *Accounting and Economics Society Annual Conference*.
Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Yi, D., Sun, J., & Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A survey. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2312.10997.
Guttman, I., Kremer, I., & Skrzypacz, A. (2014). Not only what but also when: A theory of dynamic voluntary disclosure. *American Economic Review, 104*(8), 2400–2420.
Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics, 36*(1-3), 337–386.
Hsu, C., & Wang, R. (2021). Bundled earnings guidance and analysts' forecast revisions. *Contemporary Accounting Research, 38*(4), 3146–3181.
Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1-38.
Jin, G. Z., Luca, M., & Martin, D. (2021). Is no news (perceived as) bad news? An experimental investigation of information disclosure. *American Economic Journal: Microeconomics, 13*(2), 141–173.
Jung, W.-O., & Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research, 26*(1), 146–153.
Kalai, A. T., & Vempala, S. S. (2024). Calibrated language models must hallucinate. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing* (pp. 160-171).
Kim, A., Muhn, M., & Nikolaev, V. V. (2024a). Bloated disclosures: can ChatGPT help investors process information? *Chicago Booth Research Paper*, (23-07), 2023-59.
Kim, A., Muhn, M., & Nikolaev, V.V. (2024b). Financial statement analysis with large language models. *arXiv preprint arXiv:2407.17866*.
Li, S., Yang, C., Wu, T., Shi, C., Zhang, Y., Zhu, X., ... & Lam, W. (2024). A Survey on the Honesty of Large Language Models. *arXiv preprint arXiv:2409.18786*.
Li, Y., Lin, Y., Wang, X., & Yang, S. (2024). Wall Street and product quality: The duality of financial analysts. *The Accounting Review*.

Libgober, J., Michaeli, B., & Wiedman, E. (2023). With a Grain of Salt: Uncertain Veracity of External News and Firm Disclosures. *arXiv preprint arXiv:2304.09262*.

Lundholm, R., & Myers, L. A. (2002). Bringing the future forward: the effect of disclosure on the returns-earnings relation. *Journal of Accounting Research*, *40*(3), 809-839.

Merriam-Webster. (2023). Hallucination. In *Merriam-Webster.com dictionary*. https://www.merriam-webster.com/dictionary/hallucination

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics, 12*(2), 380–391.

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*.

Rogers, J. L., & Van Buskirk, A. (2013). Bundled forecasts in empirical accounting research. *Journal of Accounting and Economics, 55*(1), 43–65.

Smith, K. (Forthcoming). Beyond the event window: Earnings horizon and the informativeness of earnings announcements. *The Accounting Review*.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665-690.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, *5*, 179-194.

Versano, T. (2021). Silence can be golden: On the value of allowing managers to keep silent when information is soft. *Journal of Accounting and Economics, 71*(2-3), 101399.

Weise, K., & Metz, C. (2023). When AI chatbots hallucinate. *The New York Times*, *9*, 610-23.

Appendix A: Variable Definitions

Dependent Variables	
# MgrForecasts	The number of forecasts issued by a firm within a quarter for all types of financial metrics, for EPS, and for sales, respectively.
MgrForecasts	An indicator equals one if a firm makes a forecast within a quarter for all types of financial metrics, for EPS, and for sales, respectively, and zero otherwise.
EPS	The earnings per share before extraordinary items and discontinued operations.
PE	The share price at the end of the quarter divided by earnings per share excluding extraordinary items in the same quarter.
Tobin's Q	The market equity (price per share times the number of shares outstanding) plus long-term debt and short-term debt scaled by book assets.
Independent Variables	
TechAnalyst	An indicator equals one if a firm is covered by at least one technical analyst at the end of 2021, and zero otherwise.
Post	An indicator equals one for quarters after 2022Q4 (inclusive), and zero otherwise.
# Geographic Segments	The number of geographic segments. We require the total sales of all segments within a firm to be larger than 80% of the firm-level sales (Cohen and Lou, 2012).
Foreign Operations	An indicator equals one if a firm has a segment operating outside the U.S. and zero otherwise. We require the total sales of all segments within a firm to be larger than 80% of the firm-level sales (Cohen and Lou, 2012).
Control Variables	
InsOwn	The percentage ownership by institutional investors at the end of the quarter.
InsOwnTop5	The percentage ownership by the five largest institutional investors at the end of the quarter.
Return	Cumulative stock return over the quarter.
Loss	An indicator equals one if the EPS is negative in the quarter.
EPS Increase	An indicator equals one if a firm reports an increase in earnings per share this quarter compared with four quarters ago.
AbsEPSChange	The absolute change in a firm's earnings for the current quarter compared with four quarters ago, normalized by last year's stock price.
Leverage	The sum of the amount of long-term debt exceeding maturity of one year and debt in current liabilities (including long-term debt due within one year) divided by the total value of assets.
Size	The market value of equity for a firm at the end of the quarter.
BM	The book value of a firm's common equity divided by the market value of equity.
Return Volatility	The standard deviation of daily returns for a firm over the quarter.
AnalystCover	The number of analysts with at least one forecast on EPS for the firm over the quarter.

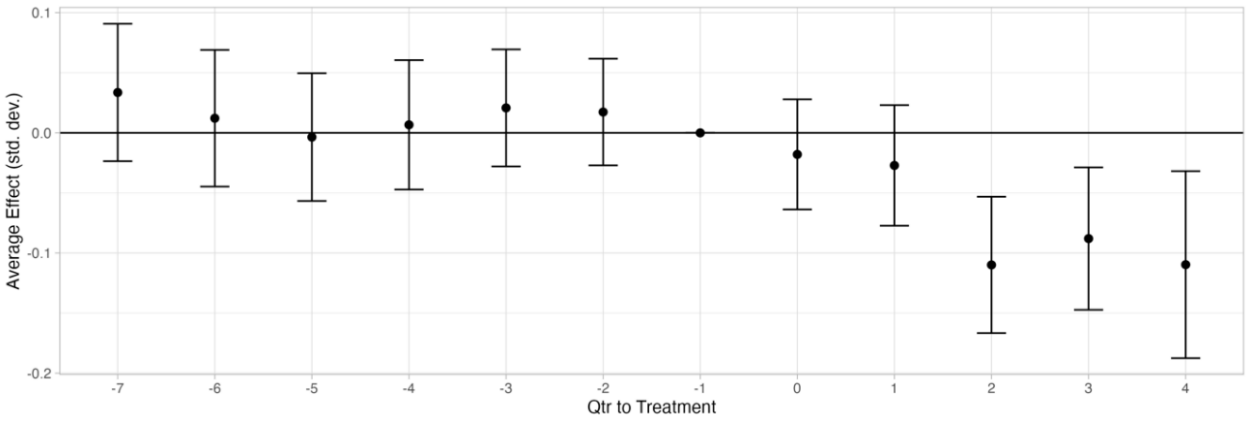
Figure 5: The Dynamic Treatment Effects of AI Processing on Voluntary Disclosure Around the Introduction of ChatGPT

We assess the dynamic treatment effects of AI information processing on voluntary disclosure around the introduction of ChatGPT in 2022Q4. Specifically, we estimate the following specification:

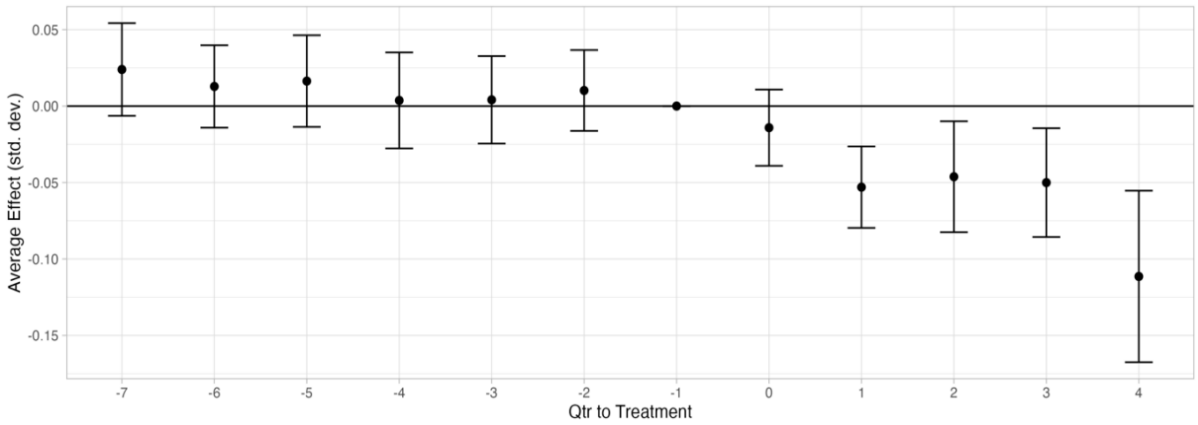
$$Mgr\ Forecasts_{i,t} = \beta_s \sum_{s=-7\sim+4,\ s\neq-1} TechAnalyst_i \times D_{s(t)} + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $Mgr\ Forecasts_{i,t}$ is an indicator variable that equals one if the firm i issues at least one forecast in quarter t and zero otherwise. $D_{s(t)}$ is a set of indicator variables that equals one if the time period is s quarters away from 2022Q4 (s=0 when ChatGPT was introduced), with the 2022Q3 (s=-1) omitted and used as the benchmark. $TechAnalyst_i$ equals one if the firm i is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. We control for other firm-level characteristics that could affect corporate disclosure. We lag all control variables by one quarter. The sample covers all quarters from 2021 to 2023. We include firm and year-quarter fixed effects, and cluster standard errors by firm.

Panel A: All Management Forecasts



Panel B: EPS Forecasts



Panel C: Sales Forecasts

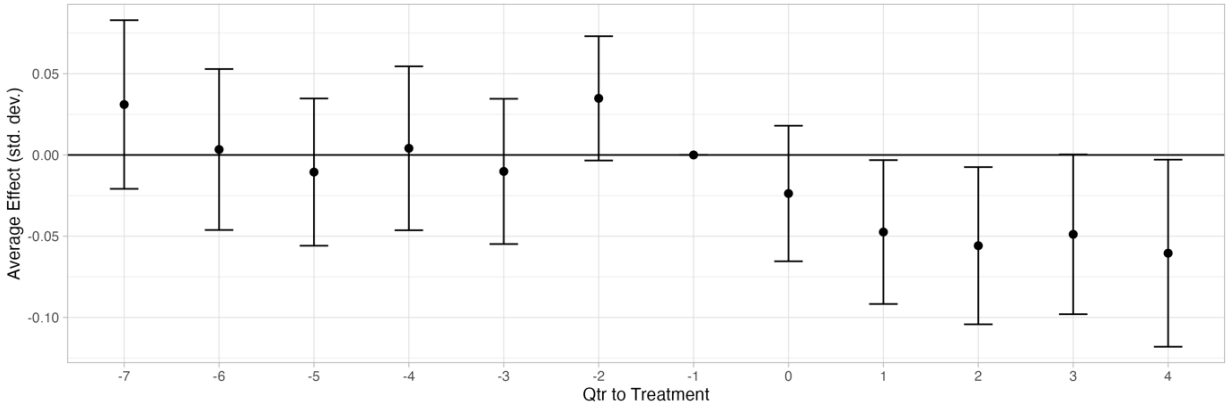


Table 2: Summary Statistics

This table presents the summary statistics for the variables used in our main analyses. We report the number of observations, mean, standard deviation, and the 25th, 50th, and 75th percentiles for all key variables. All continuous variables, except for return-related variables, are winsorized at the 1st and 99th percentiles to mitigate the effects of outliers.

Variable	**Number of Obs**	**Mean**	**Standard Deviation**	**25%**	**Median**	**75%**
# MgrForecasts - All	9866	0.806	1.441	0.000	0.000	1.000
# MgrForecasts – EPS	9866	0.171	0.424	0.000	0.000	0.000
# MgrForecasts – SALES	9866	0.295	0.543	0.000	0.000	1.000
MgrForecasts - All	9866	0.338	0.473	0.000	0.000	1.000
MgrForecasts – EPS	9866	0.154	0.361	0.000	0.000	0.000
MgrForecasts – SALES	9866	0.258	0.438	0.000	0.000	1.000
TechAnalyst	9866	0.419	0.493	0.000	0.000	1.000
# Tech Analyst / # Analyst	9866	0.183	0.275	0.000	0.000	0.333
Foreign Operations	8289	0.728	0.444	0.000	1.000	1.000
# Geographic Segments	8289	3.076	2.378	1.000	2.000	4.000
InsOwn	9866	0.814	0.204	0.726	0.841	0.929
InsOwnTop5	9866	0.360	0.105	0.298	0.354	0.415
AnalystCover	9866	10.851	7.575	5.000	9.000	16.000
Leverage	9866	0.303	0.191	0.145	0.300	0.438
Loss	9866	0.234	0.424	0.000	0.000	0.000
EPS Increase	9866	0.578	0.494	0.000	1.000	1.000
AbsEPSChange	9866	0.022	0.052	0.002	0.006	0.018
Return Volatility	9866	0.037	0.036	0.018	0.025	0.038
Return	9866	-0.068	0.404	-0.157	-0.015	0.122
Size (Billions)	9866	21.165	48.148	1.497	4.593	17.044
BM	9866	0.480	0.394	0.188	0.378	0.673
$\widehat{H}$	1266	0.496	0.208	0.294	0.462	0.715
$\hat{I}$	1266	635.759	809.051	168.988	324.991	758.246
$(\widehat{\pi_{NR} - \pi_R})_{adj}$	784	74.143	111.248	11.104	30.939	84.888
$\widehat{\emptyset}(3)$	1266	0.309	0.200	0.144	0.293	0.442
$\widehat{\emptyset}(5)$	1266	0.367	0.213	0.200	0.361	0.521

Table 3: The Impact of AI Processing on Management Forecasts

This table reports the results testing the impact of AI processing on quarterly management forecasts from 2021 to 2023. We estimate the following difference-in-differences design:

$$Mgr\ Forecasts_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $Mgr\ Forecasts_{i,t}$ is an indicator variable that equals one if the firm *i* issues at least one forecast in quarter *t* and zero otherwise. $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). We control for other firm-level characteristics that could affect corporate disclosure. We lag all control variables by one quarter. We include firm and year-quarter fixed effects, and cluster standard errors by firm.

Dependent Var.	**MgrForecasts**		
	All	**EPS**	**SALES**
	(1)	**(2)**	**(3)**
TechAnalyst × Post	***-0.067******	***-0.057******	***-0.051******
	(0.014)	***(0.008)***	***(0.012)***
InsOwn	-0.128*	-0.014	-0.086
	(0.077)	(0.040)	(0.060)
InsOwnTop5	0.220	0.092	0.180
	(0.140)	(0.064)	(0.116)
AnalystCover	0.002	0.003**	-0.0005
	(0.002)	(0.001)	(0.001)
Leverage	-0.047	-0.058	-0.035
	(0.081)	(0.045)	(0.071)
Loss	-0.013	-0.0008	-0.0001
	(0.013)	(0.006)	(0.011)
EPS Increase	-0.005	-0.009**	-0.003
	(0.007)	(0.004)	(0.006)
AbsEPSChange	-0.068	-0.002	-0.074
	(0.092)	(0.037)	(0.077)
Return Volatility	-0.002	0.025	0.027
	(0.162)	(0.075)	(0.134)
Return	-0.010	-0.003	-0.0006
	(0.016)	(0.007)	(0.012)
Size	-0.766**	0.213	-0.377
	(0.325)	(0.227)	(0.287)
BM	-0.021	0.003	-0.030*
	(0.023)	(0.014)	(0.018)
Firm FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Observations	9,866	9,866	9,866
R^2	0.70	0.83	0.75

Table 4: Cross-sectional Tests on Firm Complexity

This table reports the results of cross-sectional tests on the impact of AI processing on management forecasts. We estimate the following triple difference-in-differences specification:

$$Mgr\ Forecasts_{i,t} = \beta_1 TechAnalyst_i \times Post_t \times Complexity_i + \beta_2 TechAnalyst_i \times Post_t + \beta_3 Complexity_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $Mgr\ Forecasts_{i,t}$ is an indicator variable that equals one if the firm *i* issues at least one forecast in quarter *t* and zero otherwise. $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). $Complexity_i$ is proxied by two measures: (1) whether the firm operates foreign segments outside the U.S. and (2) the number of geographic segments. We follow Cohen and Lou (2012) in requiring that the total sales of all segments within a firm exceed 80% of firm-level sales to ensure the validity of both measures, which reduces the number of observations. We examine foreign operations in columns 1 to 3 and the number of geographic segments in columns 4 to 6. We control for other firm-level characteristics that could affect corporate disclosure. We lag all control variables by one quarter. We include firm and year-quarter fixed effects, and cluster standard errors by firm.

Dependent Var.	**MgrForecasts**					
Complexity	**Foreign Operations**			**# Geographic Segments**		
	All	**EPS**	**SALES**	**All**	**EPS**	**SALES**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
TechAnalyst × Post × Complexity	***-0.074****	***-0.078****	***-0.036****	***-0.009****	***-0.008****	***0.003***
	(0.028)	***(0.014)***	***(0.021)***	***(0.004)***	***(0.004)***	***(0.004)***
TechAnalyst × Post	-0.014	-0.005	-0.023	-0.041**	-0.039***	-0.060***
	(0.022)	(0.008)	(0.015)	(0.021)	(0.014)	(0.017)
Post × Complexity	0.011	0.028**	0.007	0.0003	0.004	-0.001
	(0.019)	(0.011)	(0.016)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,289	8,289	8,289	8,289	8,289	8,289
R^2	0.70	0.71	0.84	0.71	0.84	0.77

Table 5: The Impact of AI Processing on Information Processing Speed and Informativeness

The table reports the impact of AI processing on information processing speed (Panel A) and earnings announcement informativeness (Panel B). Specifically, we estimate the structural model from Smith (2024) using firm-level return volatility. In Panel A, we estimate the following regression specification:

$$\widehat{\emptyset}(\mathrm{x})_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $\widehat{\emptyset}(\mathrm{x})$ is a firm-level measure of information processing speed estimated from Smith's (2024) structural model. $\widehat{\emptyset}(3)$ ($\widehat{\emptyset}(5)$) represents the fraction of earnings information processed by the market three (five) days after the earnings release, indicating the extent of investor uncertainty reduction within that period. For each firm, we estimate two $\widehat{\emptyset}(\mathrm{x})$: one for the pre-period using four quarters before 2022Q4 and the other for the post-period using four quarters after 2022Q4 (inclusive). $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). We use the full sample (columns 1 and 4) as well as subsamples consisting of firms with (columns 2 and 5) and without (columns 3 and 6) management forecasts on earnings. In Panel B, we estimate the same specification as Panel A but employ different dependent variables that proxy for the informativeness of earnings announcements, including $\widehat{H}$, $\hat{I}$, and $(\widehat{\pi_{NR} - \pi_R})_{adj}$ estimated following Smith (2024). For both Panels A and B, we include firm fixed effects and a time dummy indicating pre- and post-periods. Standard errors are clustered by firm.

Panel A: Testing the Information Processing Speed

Dependent Var.	$\widehat{\emptyset}(\mathbf{3})$			$\widehat{\emptyset}(\mathbf{5})$		
With Forecast?	**Y&N**	**Y**	**N**	**Y&N**	**Y**	**N**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
TechAnalyst × Post	***0.017***	***0.179******	***0.004***	***0.013***	***0.254******	***-0.010***
	(0.022)	***(0.062)***	***(0.023)***	***(0.026)***	***(0.068)***	***(0.027)***
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	988	215	773	988	215	773
R^2	0.83	0.85	0.86	0.82	0.86	0.85

Panel B: Testing the Informativeness of Earnings Announcements

Dependent Var.	$\widehat{H}$			$\hat{I}$			$(\widehat{\pi_{NR} - \pi_R})_{adj}$		
With Forecast?	**Y&N**	**Y**	**N**	**Y&N**	**Y**	**N**	**Y&N**	**Y**	**N**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
TechAnalyst × Post	***0.040***	***0.167*****	***0.028***	***-7.38***	***153.4***	***-32.2***	***7.65***	***53.8***	***-3.06***
	(0.027)	***(0.064)***	***(0.031)***	***(65.9)***	***(112.4)***	***(81.3)***	***(12.0)***	***(32.6)***	***(13.0)***
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	988	215	773	988	215	773	673	134	539
R^2	0.70	0.77	0.71	0.84	0.93	0.83	0.90	0.91	0.91

Table 6: Analyst Forecast Revisions Following Management Forecasts

This table reports the results testing the impact of AI processing on the analyst forecast revisions following management forecast news. We estimate the specification below:

$$AFRev_{i,t} = \beta_1 TechAnalyst_i \times Post_t \times MFNews_{i,t} + \beta_2 TechAnalyst_i \times Post_t + \beta_3 Post_t \times MFNews_{i,t} + \beta_4 TechAnalyst_i \times MFNews_{i,t} + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where *AFRev* is calculated as the difference between the first analyst forecast issued after the managerial guidance date and the last analyst forecast issued before the managerial guidance date, scaled by the firm's stock price three trading days before the guidance date. *MFNews* proxies for the new information in management forecasts and is calculated as the difference between managerial guidance and the last analyst forecast prior to the guidance date, scaled by the firm's stock price three trading days before the guidance date. $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). We control for firm-level characteristics that affect corporate disclosure. We utilize the full sample in columns 1 and 4, the subsample with positive *MFNews* in columns 2 and 5 (i.e., management forecast is higher than the last analyst forecast), and the subsample with negative *MFNews* in columns 3 and 6 (i.e., management forecast is lower than the last analyst forecast). We include firm and year-quarter fixed effects. Standard errors are clustered by firm.

Dependent Var.	**AFRev**					
Sample Selection	**Full**	**Pos.**	**Neg.**	**Full**	**Pos.**	**Neg.**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
TechAnalyst × Post × MFNews				***-0.441******	***-0.937****	***-0.274***
				(0.131)	***(0.429)***	***(0.506)***
MFNews	*0.277****	*0.372****	*0.331****	*0.262**	*0.194*	*0.302***
	(0.066)	*(0.059)*	*(0.100)*	*(0.125)*	*(0.136)*	*(0.144)*
TechAnalyst × MFNews				0.110	0.283*	-0.048
				(0.110)	(0.166)	(0.211)
TechAnalyst × Post				0.00004	-0.0003	0.0002
				(0.0001)	(0.0002)	(0.0003)
Post × MFNews				0.084	0.143	0.133
				(0.137)	(0.167)	(0.147)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	619	260	339	619	260	339
R^2	0.55	0.68	0.59	0.57	0.69	0.60

Table 7: Testing The Minimum Principle

This table reports the results testing the implications of the minimum principle in evidence games (Acharya et al., 2011; Guttman et al., 2014) by examining whether increased AI processing increases the non-disclosure thresholds and prices. To this end, we utilize future EPS as a measure of the non-disclosure threshold and employ the price-to-earnings (PE) ratio and Tobin's Q as proxies for current non-disclosure prices. We test whether there are increases in these variables for firms choosing non-disclosure when their analysts increasingly rely on generative AI by estimating the following specification:

$$ND\ Threshold\ or\ ND\ Prices_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where *ND Threshold* is proxied by future *EPS*, which is earnings per share for firm *i* in quarter *t+1*. *ND Prices* are proxied by current *Price-to-Earnings Ratio (PE)* and *Tobin's Q*, where *PE* is price at the end of the quarter *t* divided by earnings per share for firm *i* in quarter *t*. *Tobin's Q* is defined as the market equity plus long-term debt and short-term debt in quarter *t* scaled by book assets for firm *i* in quarter *t*. $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for all quarters after 2022Q4 (inclusive). Our sample consists of firms without management forecasts. We include firm and year-quarter fixed effects and cluster standard errors by firm.

Dependent Var.	**EPS**	**PE**	**Tobin's Q**
	(1)	**(2)**	**(3)**
TechAnalyst × Post	***0.194******	***0.900******	***0.736******
	(0.046)	***(0.143)***	***(0.106)***
InsOwn	0.616***	0.728	0.589
	(0.236)	(0.699)	(0.382)
InsOwnTop5	-1.38***	0.100	-0.527
	(0.393)	(1.20)	(0.706)
AnalystCover	0.009	-0.018	-0.022***
	(0.006)	(0.015)	(0.007)
Leverage	0.104	1.42**	-1.20**
	(0.302)	(0.659)	(0.519)
Loss	0.047	0.212	-0.071**
	(0.043)	(0.157)	(0.035)
EPS Increase	0.171***	-0.049	0.088***
	(0.024)	(0.062)	(0.032)
AbsEPSChange	-0.274	-0.441	-0.228
	(0.294)	(0.529)	(0.215)
Return Volatility	-1.84***	2.71**	-2.40**
	(0.553)	(1.14)	(0.991)
Return	0.118***	-0.152	0.409***
	(0.044)	(0.112)	(0.076)
Size	5.10**	3.07	16.3***
	(2.21)	(4.20)	(4.28)
Firm FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Observations	6,947	6,722	6,586
R^2	0.75	0.35	0.88

Table 8: The Impact of AI Processing on Annual Management Forecasts

This table reports the results testing the impact of AI processing on annual management forecasts from 2018 to 2023. We estimate the following difference-in-differences design:

$$Mgr\ Forecasts_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $Mgr\ Forecasts_{i,t}$ is an indicator variable that equals one if the firm *i* issues at least one forecast in year *t* and zero otherwise. $TechAnalyst_i$ equals one if the firm *i* is covered by at least one technical analyst at the end of 2021 (i.e., before the ChatGPT introduction) and zero otherwise. $Post_t$ equals one for the years 2022 and 2023. We control for other firm-level characteristics that could affect corporate disclosure. We lag all control variables by one year. We include firm and year fixed effects, and cluster standard errors by firm.

Dependent Var.	**MgrForecasts**		
	All	**EPS**	**SALES**
	(1)	**(2)**	**(3)**
TechAnalyst × Post	***-0.066******	***-0.028****	***-0.093******
	(0.018)	***(0.016)***	***(0.019)***
InsOwn	0.231**	0.165**	0.217**
	(0.097)	(0.066)	(0.094)
InsOwnTop5	-0.239*	-0.208*	-0.118
	(0.138)	(0.110)	(0.145)
AnalystCover	0.010***	0.005***	0.006***
	(0.002)	(0.002)	(0.002)
Leverage	0.142**	0.013	0.182***
	(0.057)	(0.053)	(0.060)
Loss	0.002	-0.053***	-0.005
	(0.017)	(0.014)	(0.017)
EPS Increase	0.017**	0.002	0.028***
	(0.008)	(0.007)	(0.009)
AbsEPSChange	0.006	-0.033	-0.007
	(0.044)	(0.031)	(0.041)
Return Volatility	-2.56***	-1.30***	-1.77***
	(0.717)	(0.470)	(0.677)
Return	0.017**	0.008*	0.020***
	(0.008)	(0.004)	(0.008)
Size	-0.598	-0.462	-0.638
	(0.578)	(0.359)	(0.424)
BM	0.014	0.004	0.004
	(0.022)	(0.014)	(0.019)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	6,414	6,414	6,414
R^2	0.65	0.84	0.76

Table 9: The Impact of AI Processing on Management Forecasts Using Number of Forecasts and A Continuous Treatment Variable

This table reports the results testing the impact of AI processing on quarterly management forecasts from 2021 to 2023 using the number of forecasts and a continuous treatment variable. We estimate the following difference-in-differences design:

$$\#\,Mgr\ Forecasts_{i,t} = \beta_1 TechAnalyst_i \times Post_t + Controls + \alpha_t + \gamma_i + \epsilon_{i,t},$$

where $\#\,Mgr\ Forecasts_{i,t}$ is the number of forecasts issued by firm *i* in quarter *t*. $TechAnalyst_i$ is a continuous treatment variable at the firm level as the percentage of technical analysts at the end of 2021 (i.e., before the ChatGPT introduction). $Post_t$ equals one for all quarters after 2022Q4 (inclusive). We control for other firm-level characteristics that could affect corporate disclosure. We lag all control variables by one quarter. We include firm and year-quarter fixed effects, and cluster standard errors by firm.

Dependent Var.	**# MgrForecasts**		
	All	**EPS**	**SALES**
	(1)	**(2)**	**(3)**
TechAnalyst × Post	***-0.189******	***-0.083******	***-0.068******
	(0.060)	***(0.017)***	***(0.026)***
InsOwn	-0.277	-0.002	-0.118
	(0.212)	(0.047)	(0.090)
InsOwnTop5	0.659*	0.114	0.313**
	(0.359)	(0.081)	(0.152)
AnalystCover	0.003	0.002	-0.0001
	(0.005)	(0.001)	(0.002)
Leverage	-0.055	-0.034	-0.043
	(0.207)	(0.056)	(0.093)
Loss	-0.004	-0.002	0.001
	(0.033)	(0.008)	(0.014)
EPS Increase	-0.033*	-0.012**	-0.009
	(0.020)	(0.005)	(0.008)
AbsEPSChange	-0.242	-0.038	-0.097
	(0.182)	(0.046)	(0.088)
Return Volatility	-0.004	0.075	0.077
	(0.410)	(0.101)	(0.177)
Return	-0.009	0.0006	-0.012
	(0.042)	(0.010)	(0.019)
Size	-0.419	0.332	-0.424
	(0.887)	(0.317)	(0.347)
BM	-0.035	0.011	-0.034
	(0.074)	(0.015)	(0.021)
Firm FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Observations	9,866	9,866	9,866
R^2	0.76	0.77	0.69

Appendix B: Proofs in Sections 3 and 6

Proof of Lemma 1.1: We have shown in text that

$$P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + 1_{v \ge \tau}\big(1 - F(\tau)\big)\big(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau)\big) \tag{A1}$$

It then follows that:

$$\begin{aligned}
\int P_a(v)f(v)dv =& \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \int_{\tau}^{\bar{v}} (1 - F(\tau))(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau))f(v)dv \\
=& \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \int_{\bar{v}}^{\bar{v}} (1 - F(\tau))(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau))f(v)dv \\
& - \int_{\underline{v}}^{\tau} (1 - F(\tau))(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau))f(v)dv \\
=& F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + (1 - F(\tau))\mu \\
& - \int_{\underline{v}}^{\tau} (1 - F(\tau))vf(v)dv + \int_{\underline{v}}^{\tau} (1 - F(\tau))\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau)f(v)dv \\
=& F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + (1 - F(\tau))\mu \\
& -(1 - F(\tau))F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + (1 - F(\tau))F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) \\
=& F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + (1 - F(\tau))\mu
\end{aligned}$$

Proof of Proposition 1.1: From (4) and (5), the indifference condition can be re-arranged as

$$\begin{aligned}
p(1 - q)(\tau^* - P_h(\emptyset)) &= (1 - p)\left(\int P_a(v)f(v)dv - P_a(\tau^*)\right) \\
&= (1 - p)\left(F(\tau^*)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*) + \big(1 - F(\tau^*)\big)\mu - P_a(\tau^*)\right) \\
&= (1 - p)\big(1 - F(\tau^*)\big)(\mu - \tau^*)
\end{aligned} \tag{A2}$$

Using an integration by parts to develop the left-hand side of the above

$$\tau^* - P_h(\emptyset) = \tau^* - \frac{q\mu + (1 - q)\int_{\underline{v}}^{\tau^*} vf(v)dv}{q + (1 - q)F(\tau^*)} = \frac{q(\tau^* - \mu) + (1 - q)\int_{\underline{v}}^{\tau^*} F(v)dv}{q + (1 - q)F(\tau^*)} \tag{A3}$$

which implies an equilibrium condition:

$$p(1 - q)\frac{q(\tau^* - \mu) + (1 - q)\int_{\underline{v}}^{\tau^*} F(v)dv}{q + (1 - q)F(\tau^*)} = (1 - p)\big(1 - F(\tau^*)\big)(\mu - \tau^*)$$

which can be reorganized as

$$\Gamma(\tau^*) \equiv \left(\frac{1 - p}{p}\big(1 - F(\tau^*)\big) + \frac{q(1 - q)}{q + (1 - q)F(\tau^*)}\right)(\mu - \tau^*) - \frac{(1 - q)^2\int_{\underline{v}}^{\tau^*} F(v)dv}{q + (1 - q)F(\tau^*)} = 0$$

Denote τ_d as the solution of the above equation at $p = 1$, in which case the solution coincides with the threshold in Jung and Kwon (1988).

Existence. Note that

$$\Gamma(\mu) = -\frac{(1-q)^2 \int_{\underline{v}}^{\mu} F(v)dv}{q + (1-q)F(\mu)} < 0 \tag{A4}$$

and, because $\tau_d < \mu$ is given by

$$(1-q)\int_{\underline{v}}^{\tau_d} F(v)dv = q(\mu - \tau^*) \tag{A5}$$

it must hold that:

$$\begin{aligned}\Gamma(\tau_d) &= \left(\frac{1-p}{p}\big(1-F(\tau_d)\big) + \frac{q(1-q)}{q+(1-q)F(\tau_d)}\right)(\mu-\tau_d) - \frac{(1-q)^2\int_{\underline{v}}^{\tau_d} F(v)dv}{q+(1-q)F(\tau_d)} \\ &= \left(\frac{1-p}{p}\big(1-F(\tau_d)\big) + \frac{q(1-q)}{q+(1-q)F(\tau_d)}\right)(\mu-\tau_d) - \frac{(1-q)q(\mu-\tau_d)}{q+(1-q)F(\tau_d)} \\ &= \frac{1-p}{p}\big(1-F(\tau_d)\big)(\mu-\tau_d) > 0\end{aligned}$$

which, by continuity, implies the existence of at least one solution $\tau^* \in (\tau_d, \mu)$. Next, we show that there are no solutions outside of this interval. It is readily verified that $\Gamma(\tau) < 0$ for any $\tau \geq \mu$. To show that $\Gamma(\tau) > 0$ for $\tau < \tau_d$, it is sufficient to verify that

$$\gamma(\tau) \equiv q(\mu - \tau) - (1-q)\int_{\underline{v}}^{\tau} F(v)dv > 0 \tag{A6}$$

The function $\gamma(\tau)$ is decreasing in τ, attaining zero at $\tau = \tau_d$, which confirms that the inequality (A6) holds.

Uniqueness. Logconcavity implies that $\phi(\tau) \equiv \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)$ has a derivative that is less than one (Bergstrom and Bagnoli, 2005). We can then rewrite

$$P_h(\emptyset) = \alpha(\tau)\mu + (1-\alpha(\tau))\phi(\tau) \tag{A7}$$

where $\alpha(\tau) \equiv \frac{q}{q+(1-q)F(\tau)}$ is decreasing in τ. Differentiating this expression:

$$\frac{\partial P_h(\emptyset)}{\partial \tau} = \alpha'(\tau)\big(\mu - \phi(\tau)\big) + \big(1-\alpha(\tau)\big)\phi'(\tau) < \phi'(\tau) < 1 \tag{A8}$$

so that the non-disclosure price $P_h(ND)$ preserves the property of log-concave distributions on the conditional expectation. It then follows from the fact that the left-hand side of (6) is increasing in τ while the right-hand side is decreasing in τ, that (6) has at most one solution.

Proof of Corollary 1.1: An immediate application of the implicit function theorem yields:

$$\frac{\partial \tau^*}{\partial p} = -\frac{(1-q)\big(\tau^* - P_h(ND)\big) + \big(1-F(\tau^*)\big)(\mu-\tau^*)}{p(1-q)\big(1-\phi'(\tau^*)\big) + (1-p)f(\tau^*)(\mu-\tau^*) + (1-p)\big(1-F(\tau^*)\big)} \tag{A9}$$

The denominator is positive, and the term in the numerator is positive because (i) $\tau^* > P_h(ND)$ since $P_h(ND)$ is U-shaped in τ (Bertomeu et al., 2021) with a minimum at $\tau = \tau_d < \tau^*$, (ii) $\mu - \tau^* > 0$. The

comparative statics for $P_h(\emptyset)$ readily follows from the fact that $P_h(\emptyset)$ is U-shaped in τ with a minimum at τ_d.

Proof of Corollary 1.2: The expected price after an equilibrium disclosure $d(v) = v \geq \tau^*$ is

$$\begin{aligned} M(v) &= pv + (1-p)P_a(v) \\ &= pv + (1-p)\left(\int_{\underline{v}}^{\tau^*} f(v')v'dv' + \big(1 - F(\tau^*)\big)v\right) \\ \frac{\partial M(v)}{\partial p} &= (1-p)(\tau^* - v)f(\tau^*)\frac{\partial \tau^*}{\partial p} > 0 \\ \frac{\partial M'(v)}{\partial p} &= -(1-p)f(\tau^*)\frac{\partial \tau^*}{\partial p} > 0 \end{aligned}$$

Proofs for Section 6.1. The expected price conditional on a hallucination is

$$\begin{aligned} H = \int P(v)f(v)dv =& \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \int_{\tau}^{\bar{v}} \left(1 - \frac{1-p}{1-pq}F(\tau)\right)(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau))f(v)dv \\ =& \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \left(1 - \frac{1-p}{1-pq}F(\tau)\right)(\mu - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)) \\ & - \int_{\underline{v}}^{\tau} \left(1 - \frac{1-p}{1-pq}F(\tau)\right)(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau))f(v)dv \\ =& \frac{1-p}{1-pq}F(\tau)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \left(1 - \frac{1-p}{1-pq}F(\tau)\right)\mu \end{aligned} \tag{A10}$$

In the above expression, $(1-p)/(1-pq) < 1$ assigns a lower probability weight on the lowest expectation $\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) < \mu$ than in the baseline model. Hence, for any given threshold, the payoff to hallucination is higher. This immediately implies (19) given that $P_h(\emptyset) = P(\emptyset)$. To show (20):

$$\begin{aligned} \Delta =& pqP(\emptyset) + (1-pq)P(\tau) - (p(qP_h(\emptyset) + (1-q)\tau) + (1-p)P_a(\tau)) \\ =& (1-pq)P(\tau) - p(1-q)\tau - (1-p)P_a(\tau) \\ =& (1-pq)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) - p(1-q)\tau - (1-p)\mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) \\ & + \left((1-pq)\left(1 - \frac{1-p}{1-pq}F(\tau)\right) - (1-p)(1-F(\tau))\right)(\tau - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)) \\ =& pq(1-F(\tau))(1-p)(\tau - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau)) > 0. \end{aligned}$$

Rearranging the indifference condition (17) and substituting the hallucination expected price from (18):

$$\begin{aligned} pP(\emptyset) + (1-p)\int P(v)f(v)dv &= pqP(\emptyset) + (1-pq)P(\tau^*) \\ (1-p)\int \big(P(v) - P(\tau^*)\big)f(v)dv &= pqP(\emptyset) - pP(\emptyset) - (1-p)P(\tau^*) + (1-pq)P(\tau^*) \\ (1-p)\left(1 - \frac{1-p}{1-pq}F(\tau^*)\right)(\mu - \tau^*) &= p(1-q)\big(P(\tau^*) - P(\emptyset)\big) \end{aligned} \tag{A11}$$

Because the left-hand side is positive (negative) and the right-hand side is negative (positive) if $\tau^* = \underline{v}$ (if $\tau^* = \bar{v}$), there is always an interior threshold equilibrium. Suppose next that $q = 0$, so that $P(\emptyset) = \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*)$ and

$$P(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*) + 1_{v \geq \tau^*}\big(1 - (1-p)F(\tau^*)\big)\big(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*)\big)$$

implying that

$$\begin{aligned}(1-p)\big(1-(1-p)F(\tau^*)\big)(\mu - \tau^*) &= p\big(P(\tau^*) - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*)\big) \\ &= p\big(1-(1-p)F(\tau^*)\big)\big(\tau^* - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*)\big) \\ (1-p)(\mu - \tau^*) &= p\big(\tau^* - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau^*)\big) \\ &= p\left(\tau^* - \frac{F(\tau^*)\tau^* - \int_{\underline{v}}^{\tau^*} F(v)dv}{F(\tau^*)}\right)\end{aligned}$$

which yields equation (21) in text. Logconcavity implies that the right-hand side is increasing in τ and the left-hand side is decreasing in τ, implying that the equilibrium is unique. Further, as $(1-p)/p$ is decreasing in p, an immediate application of the implicit function theorem demonstrates that τ^* is decreasing in p, in line with the baseline model.

Proofs for Section 6.2. As discussed in Section 6.2., we modify the baseline model so that, upon non-disclosure, the AI hallucinates a garbled signal unrelated to the firm's fundamentals with probability $\rho \in (0,1)$. With probability $1-\rho$, the AI does not hallucinate, leading us to explore two different formulations. In the first formulation, we assume that the AI independently identifies information even when it has not been disclosed. This represents an ideal scenario where the AI can access more information than humans in non-disclosure situations. In the second formulation, we assume that information known to the manager is not accessible to the AI, resulting in the AI observing a non-disclosure with probability $1-\rho$.

First, we consider the formulation in which the AI hallucinates with probability ρ but is otherwise (i.e., with probability $1-\rho$) reporting the true information. Equation (22) simplifies to

$$P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau) + \frac{(1-F(\tau))(1-1_{v<\tau}\rho)}{1-(1-\rho)F(\tau) - 1_{v<\tau}\rho(1-F(\tau))}(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \leq \tau))$$

so that the expected value from hallucination is

$$H = \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \frac{1-F(\tau)}{1-(1-\rho)F(\tau)}(\mu - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau))$$
$$- \int_{\underline{v}}^{\tau} \frac{1-F(\tau)}{1-(1-\rho)F(\tau)}(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau)) f(v) dv$$
$$+ \frac{(1-F(\tau))(1-\rho)}{1-(1-\rho)F(\tau)-\rho(1-F(\tau))} \int_{\underline{v}}^{\tau} (v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau)) f(v) dv$$
$$= \frac{\rho F(\tau)}{1-(1-\rho)F(\tau)} \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \frac{1-F(\tau)}{1-(1-\rho)F(\tau)} \mu \quad \text{(A12)}$$

The indifference condition in this setting is:

$$p(qP_h(\emptyset) + (1-q)\tau^*) + (1-p)P_a(\tau^*) = pP_h(\emptyset) + (1-p)((1-\rho)P_a(\tau^*) + \rho H)$$
$$p(1-q)\tau^* + \rho(1-p)P_a(\tau^*) = p(1-q)P_h(\emptyset) + (1-p)\rho H$$
$$p(1-q)(\tau^* - P_h(\emptyset)) = \frac{\rho(1-p)\big(1-F(\tau^*)\big)}{1-(1-\rho)F(\tau^*)}(\mu - \tau^*) \quad \text{(A13)}$$

We prove next the existence and uniqueness. The function

$$\Gamma(\tau) = \frac{\rho(1-p)\big(1-F(\tau^*)\big)}{1-(1-\rho)F(\tau^*)}(\mu - \tau^*) - p(1-q)\big(\tau^* - P_h(\emptyset)\big) \quad \text{(A14)}$$

satisfies $\Gamma(\mu) < 0 < \Gamma(\underline{v})$ and therefore has a solution with $\Gamma'(\tau^*) < 0$. It can also be readily verified that there is no solution above μ, because this would imply that the left-hand side is negative, i.e., $P_h(\emptyset) > \tau^* > \mu$, which would contradict $P_h(\emptyset) \le \mu$. Further, if $F(.)$ is log-concave, we have already shown in the baseline model that, in the proof of Corollary 1, $\tau^* - P_h(\emptyset)$ is increasing in the threshold. The right-hand side is the product of two positive decreasing functions, and thus must be decreasing. Hence, the solution is unique. Conditional on uniqueness, the implicit function theorem yields

$$\frac{\partial \tau^*}{\partial p} = \frac{-\frac{\rho\big(1-F(\tau^*)\big)}{1-(1-\rho)F(\tau^*)}(\mu - \tau^*) - (1-q)(\tau^* - P_h(\emptyset))}{-\Gamma'(\tau^*)} < 0, \quad \text{(A15)}$$

so that the probability of disclosure increases in human processing p.

Second, we consider the formulation that the AI yields a non-disclosure with probability $1-\rho$ and, in this case,

$$P_a(v) = \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau) + \frac{1-F(\tau)}{1-(1-\rho)F(\tau)}(v - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau))$$

which can be readily verified to imply the same H as in (A12). The indifference condition in this setting is:

$$pP_h(\emptyset) + (1-p)((1-\rho)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*) + \rho H) = p(qP_h(\emptyset) + (1-q)\tau^*) + (1-p)P_a(\tau^*)$$
$$p(1-q)P_h(\emptyset) + (1-p)\rho\big(H - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*)\big) = p(1-q)\tau^* + (1-p)\big(P_a(\tau^*) - \mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*)\big)$$
$$\frac{(1-p)(1-F(\tau))}{1-(1-\rho)F(\tau)}(\rho\mu + (1-\rho)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*) - \tau^*) = p(1-q)\big(\tau^* - P_h(\emptyset)\big) \quad \text{(A16)}$$

To show existence, we similarly define

$$\Gamma(\tau) = \frac{\rho(1-p)\big(1-F(\tau)\big)}{1-(1-\rho)F(\tau)}(\rho\mu + (1-\rho)\mathbb{E}(\,\tilde{v} \mid \tilde{v} \le \tau\,) - \tau) - p(1-q)\big(\tau - P_h(\emptyset)\big) \quad \text{(A17)}$$

which satisfies $\Gamma(\mu) < 0 < \Gamma(\underline{v})$ and therefore has a solution with $\Gamma'(\tau^*) < 0$. Further, $\Gamma(\tau) < 0$ for any $\tau \ge \mu$. Unfortunately, because the left-hand side is no longer known to be positive or negative, there is no longer a simple characterization of uniqueness via logconcavity (which implies that the left-hand side is increasing in threshold). We assume in what follows that the solution is unique for all p and the implicit function theorem yields

$$\begin{aligned}\frac{\partial \tau^*}{\partial p} &= \frac{-\frac{\rho\big(1-F(\tau^*)\big)}{1-(1-\rho)F(\tau^*)}(\rho\mu + (1-\rho)\mathbb{E}(\tilde{v} \mid \tilde{v} \le \tau^*) - \tau^*) - (1-q)(\tau^* - P_h(\emptyset))}{-\Gamma'(\tau^*)} \\ &= \frac{-\frac{p}{1-p}(1-q)(\tau - P_h(\emptyset)) - (1-q)(\tau^* - P_h(\emptyset))}{-\Gamma'(\tau^*)} \\ &= \frac{(1-q)(\tau^* - P_h(\emptyset))}{\Gamma'(\tau^*)(1-p)}. \end{aligned} \quad \text{(A18)}$$

Defining τ_0^* from (27), there are three cases to consider. First, if $\tau_0^* = \tau_d$, then the solution $\tau^* = \tau_d$ satisfies $\Gamma(\tau^*) = 0$ for all p. Second, if $\tau_0^* > \tau_d$, we know from the minimum principle that $\tau_0^* > P_h(\emptyset)$ and therefore τ^* initially decreases in p for p small from until $\tau^* = \tau_d$ at $p = 1$. Third, the case with $\tau_0^* < \tau_d$ is a mirror image and implies that τ^* increases until $\tau^* = \tau_d$ at $p = 1$.