ID Cover Page

Summary of WP Student Team

HOW DO AI FOCUS AND MANAGEMENT STANCE INFLUENCE THE VALUATION OF PUBLICLY LISTED AMERICAN COMPANIES

Group constitution:

Student Name	Program	Individual Title
Thomas Kanzian 64010	Management	Impact of AI related disclosures on firm value - Ai Intensity and the Effects on Company Valuation
Maximilian Klotzke 63844	Management	Impact of AI related disclosures on firm value – ai sentiment and stock price reactions

Work project carried out under the supervision of:

Advisor: Gonçalo Rocha

Reading observations for jury:

Note: this template is a support document for the jury members before the defense, only.

A Work Project, presented as part of the requirements for the Award of a Master's degree in Management from the Nova School of Business and Economics

HOW DO AI FOCUS AND MANAGEMENT STANCE INFLUENCE THE VALUATION OF PUBLICLY LISTED AMERICAN COMPANIES?

Thomas Kanzian (64010)

Maximilian Klotzke (63844)

Work project carried out under the supervision of:

Gonçalo Rocha

26/01/2026

Abstract

Firms devote increasing attention to AI, yet the extent to which such emphasis is reflected in market valuations is not well established. Using S&P 500 firms, we address this gap by measuring corporate AI commitment through textual analysis of 10-K disclosures. We find that higher relative quantitative AI focus is associated with significantly higher long-term valuations, reflected in Tobin's Q and enterprise-value multiples but not equity-based ratios. In contrast, the qualitative sentiment of AI-related language does not generate short-term abnormal returns. Overall, markets price AI as a long-term intangible asset rather than a short-term market signal.

Keywords

Artificial Intelligence, AI, S&P 500, Sentiment Analysis, Market Value, 10-K

Table of Contents

Lists of Tables and Figures III

Table of Abbreviations IV

1. Introduction 1

2. Literature Review 3

2.1. Artificial Intelligence and Company Productivity 3

2.2. AI Disclosure in Annual Reports and Public Coverage of AI 4

2.3. Textual Analysis 6

3. Data Collection 8

3.1. Database Considerations 8

3.2. Sample Size and Period 11

3.3. AI Anchors 12

4. AI Focus and the Effects on Company Valuation 13

4.1. Hypothesis 13

4.2. Methodology 15

4.3. Results 19

4.4. Discussion 25

5. AI Sentiment and Stock Price Reactions 30

5.1. Hypothesis 30

5.2. Methodology 32

5.3. Results 36

5.4. Discussion 40

6. Conclusion 45

7. Limitations 47

8. Outlook 48

References V

Appendix XV

Lists of Tables and Figures

Table 1: Model Coefficients – Tobin's Q 21

Table 2: Model Coefficients – Robustness Checks 23

Table 3: Model Coefficients – Lagged 2023 AI_z 23

Table 4: Sentiment Classification of 10-K Reports from 2017 - 2024.................................. 33

Table 5: Overall Impact of 10-K Releases on the Stock Price 36

Table 6: Impact of Management Sentiment towards AI on Stock Prices 37

Table 7: Testing Differences between Sentiments.. 38

Table 8: Pairwise Mann-Whitney Test for Sentiment Significance 39

Table 9: Impact of Industry Differences on Stock Prices... 40

Table 10A: Pairwise Welch's t-Test for Proxy Validation XV

Table 11A: Cook's Distance including Nvidia XV

Table 12A: Cook's Distance excluding Nvidia XV

Table 13A: Test for OLS Regression Assumptions XV

Table 14A: AI Sector Dependence XVI

Table 15A: Test for Normal Distribution of CARs XVI

Table 16A: Testing for Heteroskedasticity XVI

Table 17A: Testing for Autocorrelation... XVI

Table 18A: Testing for Multicollinearity... XVI

Table 19A: Pairwise Mann-Whitney U Test for Identifying Industry Differences XVII

Table 20A: AI Keywords... XVIII

Figure 1: Filter Criteria and Population Selection 12

Figure 2: AI Focus Across Sectors 16

Figure 3: Theoretical Model for the Event Study Methodology 33

Table of Abbreviations

AI	Artificial Intelligence
AR	Abnormal Returns
CAPM	Capital Asset Pricing Model
CAR	Cumulative Abnormal Returns
D/E Ratio	Debt-to-Equity Ratio
EDGAR	Electronic Data Gathering, Analysis, and Retrieval System
EV	Enterprise Value
GICS	Global Industry Classification Standard
GPT	General Purpose Technology
KWIC	Keywords-in-Context
MD&A	Management Discussion & Analysis
OLS	Ordinary Least Square
P/B	Price-to-Book Ratio
P/E	Price-to-Earnings Ratio
R&D	Research and Development
SEC	U.S. Securities and Exchange Commission
SG&A	Selling, General, and Administrative
VIF	Variance Inflation Factors

1. Introduction

Over the coming years, advances in artificial intelligence (AI) are expected to reshape how firms organize work and create value. In this regard, AI has been mentioned as a core enabler for the fourth industrial revolution (McKinsey, 2022). Simultaneously, organizations have begun systematically deploying AI tools, including conversational agents, coding assistants, and machine learning models for advanced analytics, across key business functions (Korst et al., 2025, Maslej et al., 2025).

From a macroeconomic perspective, AI is projected to generate substantial economic growth across both companies and economies. Projections by Chui et al. (2023) estimate that the annual economic contribution of AI could range from \$2.6 trillion to \$4.4 trillion across major use cases. Nevertheless, the realization of these gains at the industry level remains contingent on the existence of sector-specific automation capability and the availability of domain-specific use cases for the enhancement of products and services (Acemoglu, 2024).

From a corporate perspective, emerging evidence suggests that, rather than merely automating tasks to replace workers, AI yields the highest returns when used to augment human labor. For example, recent empirical studies show that generative AI can increase worker productivity by roughly 15.0% by complementing human capabilities rather than replacing them (Brynjolfsson et al., 2025). However, the aggregate impact of AI on firm-level productivity growth and the bottom line remains uncertain, due to the often-long-lasting effects of AI on firm productivity. Currently, AI adoption is still early and mostly centered around large companies, requiring significant complementary investments to fully utilize the potential of this technology (Goldman Sachs, 2025). Apotheker et al. (2025) also found a widening gap between “future-built” companies, firms that have put in place critical infrastructure to make use of AI, and laggards, showing a compounding gap in revenue over the coming years.

Existing empirical evidence on firm-level productivity and stock market data is still limited and mixed, but it broadly points towards a positive association between AI focus and operating performance. Studies find increased sales, employment and market valuations in line with structural changes caused by the adoption of the technology (Babina et al., 2024; Brynjolfsson et al., 2025; Mishra et al., 2022). In contrast, prior research has not yet explored sentiment analysis in this context. Moreover, short-term observations yield more mixed results. Lui et al. (2021) find a slightly negative relationship between the announcements of AI investments and stock market reactions, which is contradicted by the results of Eisfeldt et al. (2023). This suggests that short-run market reactions may depend not only on whether firms invest in AI but also on how these activities are framed in corporate communication.

Against this background, this thesis examines how investors value strategic corporate AI engagement as reflected in mandatory financial disclosures. Specifically, it focuses on firms' annual 10-K filings to the U.S. Securities and Exchange Commission (SEC). We employ text-based measures to capture both the quantitative AI focus and the qualitative stance of communication. By linking disclosure-based measures to long-term financial performance and short-term stock-market outcomes, this thesis aims to shed light on whether AI is currently perceived by capital markets as a credible strategic asset or as part of a broader technology hype.

Our work consists of two complementary methodological approaches, both based on the textual analysis of firms' 10-K reports. In the first part, a regression analysis uses the frequency of AI-related terms at the firm-year level as a proxy for strategic AI focus and examines its relation to firms' long-run financial performance, broadly following the work from Mishra et al. (2022). In the second part, we use the same AI anchors to classify the surrounding context of each AI-related keyword into distinct categories of managerial stances toward the technology. Based on these stance measures, an event study analyzes short-run stock-market reactions around 10-K filing dates. Both analyses, therefore, rely on the same underlying text but extract different dimensions of information: how much firms talk about AI as a proxy for AI engagement versus how they talk about it.

2. Literature Review

2.1. Artificial Intelligence and Company Productivity

While a rapidly expanding body of empirical research documents substantial productivity gains associated with the adoption of AI, these studies also highlight significant implementation challenges and variations in outcomes across firms. Using micro-level manufacturing data, Gao and Feng (2023) show that a one percent increase in AI penetration increases total factor productivity by 14.2% and trace this effect to value-added enhancement, skill-biased upgrading, and technology upgrading mechanisms. At the global firm level, Damioli et al. (2022) exploit a panel of AI-patenting companies and find that AI patent applications generate an 'extra-positive' impact on labor productivity, with effects concentrated in SMEs and services, using a system generalized method of moments to address endogeneity and dynamic adjustment. Babina et al. (2024) move closer to corporate finance by constructing a human-capital-based measure of AI investment from resumes and job postings. Their results imply that an increase in AI-skilled employment predicts roughly 20.0% higher sales, employment, and market valuation, with product innovation as the main channel. These studies rely on rich microdata and plausible causal designs but largely do not address how capital markets price AI-related productivity improvements.

Other work examines narrower tasks or organizational capabilities. In an online experiment, Noy and Zhang (2023) randomly grant professionals access to ChatGPT and show sizable, short-run productivity gains while noting external validity limits for longer, context-rich tasks and general-equilibrium outcomes. For public organizations, Mikalef et al. (2023) use survey data and partial least squares structural equation modeling to show that AI capabilities improve perceived organizational performance through process automation and cognitive insights. However, they find that cognitive engagement applications can even reduce performance, highlighting that not all AI deployments translate into efficiency gains. Text-based measures based on mandatory reports are closest to the present setting: Mishra et al. (2022) construct a 10-K "AI focus" score from AI-related terms and find improvements in net operating efficiency and profitability. Chen and avasan (2022), on the other hand, use digital word counts (including AI) and analyze a potential increase in firm performance. They find higher market-to-book ratios and some increases in ROA and asset turnover for companies with more digitalization mentions, but no significant improvements in margins and even sales growth. Collectively, existing studies emphasize productivity, operating efficiency, and innovation, often using patents, AI labor, or survey capabilities, and only indirectly touch on valuation.

Despite this, there is limited evidence on how capital markets price firm-level exposure to AI as reflected in mandatory corporate disclosures. Existing textual measures either bundle AI into broad "digital" activity or treat AI focus as a scalar intensity, without exploiting section-specific AI focus or management tone, nor examining how such disclosures map into valuation multiples and announcement-window Cumulative Abnormal Returns (CAR). The analysis of firm-level AI focus in 10-Ks and management stance toward AI, therefore, fills an important gap by linking documented productivity implications of AI to capital-market outcomes.

2.2. AI Disclosure in Annual Reports and Public Coverage of AI

A rapidly expanding body of evidence documents the growing strategic relevance of AI for firms and, indirectly, its salience in corporate reporting. Global surveys and policy reports show that AI adoption, while still far from universal, is increasingly viewed as central to core business processes. According to a report by BCG, INSEAD and the OECD (2025), the uptake of AI in production and services remains modest in aggregate. The report indicates that larger firms and industries that rely heavily on information and communication technologies are the main drivers of AI adoption. Among AI adopters, however, a substantial majority regard AI as "critical" or "very important" to core operations, especially when AI accounts for a sizeable share of research and development (R&D) spending.

Consistent with this, de Bellefonds et al. (2024) find that 98.0% of surveyed companies are at least experimenting with AI, although only about a quarter have advanced beyond proof-of-concept to generate measurable value and a small minority operate at the technological frontier. Maslej et al. (2025) similarly document a sharp acceleration in enterprise AI use and private AI investment, with business usage rising from just over half to nearly four-fifths of surveyed organizations within a single year. Singla et al. (2025) further highlight that firms reporting the largest AI-related EBIT contributions tend to embed AI into scaled business processes and to have senior leadership explicitly committed to AI initiatives.

Against this backdrop, only a small number of studies examine how firms communicate AI-related activities in mandatory filings. The "AI focus" metric by Mishra et al. (2022) finds that these comprehensive, regulated filings provide a credible signal of strategic emphasis but they treat AI text as context-free counts and do not study market-based outcomes. Chen and Srinivasan (2022) document that an increase in digital activity, as proxied by disclosure intensity, leads to higher market-to-book ratios and stronger earnings–returns relations. However, their measure is not specific to AI and focuses on the business description rather than the Management Discussion & Analysis (MD&A), and abstracts from the tone or framing of technology discussions. More generally, prior textual research suggests that linguistic features of MD&A sections can predict future performance and returns, but this work has not isolated AI content or examined investors' reactions to AI narratives in real time (Li, 2010).

Taken together, existing studies establish that AI is increasingly central to firms' strategies and that technology-related disclosures in 10-Ks contain value-relevant information. Yet they leave open how investors price the intensity and qualitative stance of AI-related discussion in mandatory filings. Prior work provides little direct evidence linking AI-specific disclosure metrics, viewed here as proxies for strategic commitment, to valuation multiples Furthermore, there is virtually no large-sample evidence on how management's tone towards AI in MD&A shapes short-window CARs around 10-K filings. These gaps are important for corporate finance because AI-related investments are highly intangible, risky, and difficult to observe outside of narrative disclosure. They motivate research that combines AI keywords-in-context (KWIC) measures in 10-Ks with sentiment-sensitive characterization of AI discussion in MD&A to test whether increased AI focus is reflected in valuation multiples (H1) and whether a more positive management stance toward AI is rewarded in CARs following 10-K releases (H2).

2.3. Textual Analysis

Keywords in Context

Textual analysis has become an established tool in corporate finance literature for systematically extracting information from unstructured disclosures such as annual reports, earnings announcements and financial news. It allows for the quantification of qualitative constructs, such as a company's operational or strategic focus and management sentiment. Earlier works document the rapid growth of this literature arm, emphasizing textual methods as an empirical tool beneath traditional numerical measures.

Fundamental for the sake of our textual analysis is the so-called KWIC approach. Luhn (1958) first coined this term in his work on keyword indexes for literature. In that sense, the method assigns each word an index and a defined span of surrounding words, which support the retrieval of contextual information. Biber et al. (1998) and Baker (2008) further refined this approach when analyzing the use of language based on context using the corpus-based approach. Rather than just aggregating counts, this approach inspects how a focal concept is embedded in sentences by analyzing whether language usage is predominantly positive or negative in our case.

Dictionary-based Approach

To determine whether language is used positively or negatively, we combine the KWIC approach with a dictionary-based approach. Dictionary-based textual analysis relies on pre-defined word lists that classify terms into theoretically meaningful categories. The most widely used are the Harvard IV-4 and the Loughran and McDonald dictionaries. In a typical application, researchers tokenize and clean the text. They count the frequency of words belonging to each category and then scale it by document length to obtain measures such as the proportion of negative words in a filing. Early studies in finance apply general-purpose sentiment dictionaries, such as the Harvard IV-4, to financial text. Researchers do this to measure media pessimism and link it to stock returns or to capture the tone of disclosure reports by management, analysts and news reporters to analyze their effects on cost of capital, stock return volatility and analyst forecasts (Kothari et al., 2009; Tetlock, 2007).

Based on this, Loughran and McDonald (2011) show that such generic dictionaries perform poorly in a corporate reporting context: in their large sample of 10-K filings, they document that most words flagged as negative by the Harvard IV-4 list are not actually negative in financial usage. In response, they develop finance-specific dictionaries tailored to corporate analysis, providing separate word lists for negative and positive tone, uncertainty, litigiousness, and strong and weak modal language. Corporate finance and accounting researchers have adopted these domain-adjusted lists as the standard and use them to construct quantitative proxies of sentiment, perceived risk, or disclosure style that can be linked to stock market reactions, cost of capital, or other outcome variables.

A growing body of evidence documents that the Loughran and McDonald dictionaries systematically relate to firm risk, valuation and real corporate policies. Del Gaudio et al. (2020) show that a more negative tone in corporate disclosures relates to higher levels of financial distress, highlighting increased insolvency risk. Jegadeesh and Wu (2013) find that the tone of

10-K filings, as measured with the dictionary lists, can trigger significant stock market reactions in both directions. Ahmad et al. (2016), who studied the relationship between media pessimism and stock returns, report a similar result. In the IPO context, Loughran and McDonald (2013) study the tone of S-1 filings and show that uncertain language in these prospectuses impairs investors' ability to assess the information, resulting in higher price volatility and larger first-day returns. García et al. (2023) further demonstrate that specific dictionaries capture priced sentiment effects, but their incremental explanatory power for stock returns is quantitatively small, underlining both their usefulness and their possible limitations. Other studies link tone directly to the cost of capital and real decisions. Berns et al. (2022) document that improvements in the net positive MD&A tone predict higher future investment expenditures, acquisition activity and more aggressive investment policies even after controlling for environmental variables. Taken together, these studies confirm that dictionary-based tone measures are conceptually appealing and easy to implement but also economically meaningful in explaining cross-sectional differences in risk, valuation and corporate policies.

3. Data Collection

3.1. Database Considerations

EDGAR for Annual 10-K Filings

The empirical analysis relies on the Electronic Data Gathering, Analysis, and Retrieval system (EDGAR), the SEC electronic platform through which publicly listed firms submit mandatory filings. EDGAR provides machine-readable access to a broad universe of corporate disclosures and can be queried programmatically via an application programming interface, which we use in this study to systematically download all 10-K filings for the sample firms and years. The SEC requires firms to file 10-Ks as comprehensive annual reports, containing audited financial statements, a detailed description of the firm's business, disclosures of risk factors, and MD&A of financial condition and performance (Loughran & McDonald, 2011). The richness of these narrative sections, in particular the business description and MD&A, enables the construction of firm-level text-based measures. Specifically, we process the textual content of 10-K filings to count AI-related keywords and assess the sentiment.

Standardizing 10-Ks

To ensure consistency and comparability in textual analysis across firms and industries, we implement a standardized preprocessing pipeline for 10-K filings that follows best practices in financial textual analysis (Babina et al., 2024; Kang et al., 2018). Our objective is to isolate and structure narrative content from these regulatory documents in a way that facilitates robust keyword-based measurement of AI focus and management stance.

We begin by extracting and converting the full textual content from all narrative sections of annual 10-K filings. Following Kang et al. (2018), we exclude financial statements, tables, charts, and all structured numerical disclosures that do not contribute meaningfully to managerial tone or thematic content. Additionally, we omit Item 4, “Mine Safety Disclosures”, to preserve comparability across sectors, as this section is industry-specific and disproportionately affects extractive firms. The result is a standardized set of firm-year narrative disclosures that includes the MD&A, Business Overview, Risk Factors, and other unstructured text segments, which prior research has shown to carry rich forward-looking and strategic information content.

Next, we normalize all extracted text to lowercase. This step ensures that token recognition is case-insensitive and prevents duplicate entries from capitalized variants of the same word. Subsequently, we remove numerals, hyphens, apostrophes, and other non-alphabetic characters to further reduce noise in the tokenized corpus. These elements, while potentially useful in financial contexts, are not relevant for detecting semantic references to AI-related themes.

To further refine the textual input, we remove stop words using a standard English stop-word list, as implemented in Babina et al. (2024). This step eliminates high-frequency, low-information words such as articles, prepositions, and auxiliary verbs (e.g., “the,” “is,” “at”), thereby improving the salience of substantively meaningful content in subsequent keyword counts.

Finally, to avoid double-counting when both an acronym and its expanded form appear in proximity, we implement a rule-based filter: if the term “artificial intelligence (AI)” appears, we record only one instance in the AI term frequency count. This method avoids redundantly capturing compound phrases and their abbreviations, thus preserving the interpretive integrity of the AI focus measure.

Through this multi-stage standardization procedure, we obtain a clean, comparable textual dataset across firms and years, allowing us to quantify the prevalence and framing of AI-related disclosures in a rigorous and replicable manner.

Timeline Considerations: Fiscal Year versus Filing Year

For tractability, we do not explicitly align the fiscal year referenced in 10-K filings with the calendar year in which firms file the report. Any resulting timing misalignment is unlikely to materially affect our analysis, as the set of filings collectively spans the same underlying economic period. Moreover, because our primary outcomes are firm valuation measures that reflect investors’ forward-looking expectations, the precise fiscal year label attached to a filing is of secondary importance. A 10-K released in the early calendar year is relevant for contemporaneous market valuation regardless of whether it reports outcomes for fiscal year t or t-1. Imposing a stricter alignment between fiscal and calendar years would substantially increase data-handling complexity while yielding limited incremental gains in measurement accuracy. What matters for the empirical design is that each observation consistently reflects a comparable annual reporting period, not the exact calendar label assigned to that period. Using the filing year as the reference year satisfies this requirement to a sufficient degree for the research questions examined here.

3.2. Sample Size and Period

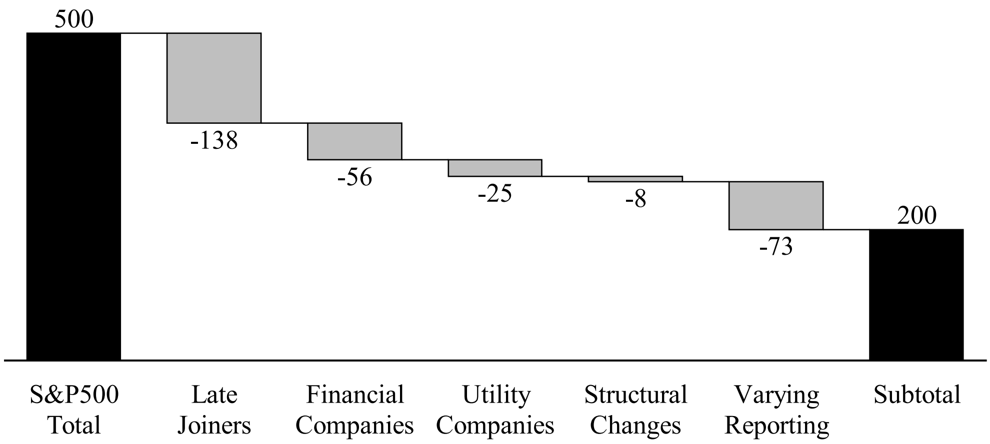
The focus for our population is standardization. To the best of our knowledge, American companies produce the most standardized quarterly and annual financial reports, which firms file and make accessible through the EDGAR database. Consequently, we focus on firms within the S&P 500 Index. These companies account for approximately 80.0% of the market capitalization of American companies and represent most of the country's economy, thus making the sample representative of the American industry (S&P Global, 2025).

The analysis spans the period from 2017 to 2024. This window captures the rapid rise in corporate attention to AI while also encompassing earlier years prior to the recent surge in interest. Importantly, the sample period includes distinct market environments: a pre-pandemic phase (2017–2019), the COVID-19 period, which marks heightened uncertainty and market disruption (2020–2022), and a post-pandemic phase (2023–2024). This structure allows the analysis to account for potential differences in disclosure behavior and valuation dynamics across major economic regimes, similar to the approach in Zhu et al. (2024).

Starting with the 500 constituents, our focus is to only include companies that have been part of the index from 2017 to 2024. That guarantees stable operational and financial performance of the companies while increasing comparability among the constituents (Chen et al., 2004; Lynch & Mendenhall, 1997). Therefore, we remove 138 companies from the dataset because they joined the index only in 2017 or later, leaving us with 362 companies. Based on previous research, we omit 56 financial companies with the Global Industry Classification Standard (GICS) code 40 and 25 utilities companies with the code 55. Financial companies tend to be more highly leveraged and risk falsifying our results since they are based on asset multiples (Fama & French, 1992). Similarly, utilities companies operate in highly regulated environments, and regulation has been shown to affect the financial performance and risk profile of public utilities (Alexander & Irwin, 1996).

This leaves us with 281 companies, from which we omit an additional eight companies because of organizational or structural changes. They either changed their name or ticker, thus, we cannot retrieve consistent historical data, decreasing the sample to 273 companies. Last, although regulators standardize the substantive content of 10-K filings, formatting varies across firms. Because our analysis relies on automated textual extraction, we exclude 73 firms whose filings embed material portions of narrative text within images or non-machine-readable formats. Such formatting prevents reliable text retrieval and would compromise the accuracy of the textual measures. After applying these data-quality filters, the final sample consists of 200 firms that meet all inclusion criteria. Figure 1 summarizes the sample selection process.

Figure 1: Filter Criteria and Population Selection



3.3. AI Anchors

To quantify firm-level exposure to AI, we construct a robust, composite dictionary grounded in three authoritative sources: the European Commission's Joint Research Centre report "AI Watch" (Samoili et al., 2021), the OECD's report on identifying and measuring developments in AI (Baruffaldi et al., 2020) and the specific AI dictionary developed by Chen and Srinivasan (2022). Rather than relying on a single definition, which might be tailored to specific empirical settings or suffer from unique biases, we seek a consensus-based measure that balances technical precision with broad applicability across the S&P 500.

We select the source dictionaries for their complementary strengths. The Joint Research Centre derives the AI Watch taxonomy through a multi-step process that combines semi-automated text mining of academic Scopus articles, desk research and validation by a panel of AI experts. Similarly, the OECD list proposes keyword-based indicators designed for macroeconomic measurement, while Chen and Srinivasan (2022) provide a dictionary that prior work has specifically validated within the context of corporate financial disclosures. While these lists provide extensive coverage independently, they also contain terms that may be too generic (e.g. broad 'digital') or too niche for a general valuation study. To enhance precision and minimize noise, we employ a systematic intersection approach. We aggregated terms from the three source dictionaries and retained only those keywords appearing in at least two of the three lists. This method identifies a consensus vocabulary for AI, excluding terms that are tangential or broadly technological. Additionally, we supplement this consolidated list with common abbreviations for the retained terms (e.g., "AI," "ML"). The resulting dictionary consists of 35 terms and serves as the basis for capturing firm-level AI focus.

4. AI Focus and the Effects on Company Valuation

4.1. Hypothesis

Recent work shows that AI commitment improves firm outcomes. The AI focus by Mishra et al. (2022) shows that AI-oriented firms exhibit higher marketing performance and superior financial outcomes. Similarly, Babina et al. (2024) have developed a measure of firm-level AI investment using worker resumes and job postings. They document that AI-investing firms experience faster growth in sales, employment and market valuations, with increased product innovation primarily driving the effect. These findings support the idea that sustained AI investment expands firms' growth opportunities and intangible capital.

The asset-market consequences of AI focus are further highlighted in Eisfeldt et al. (2023), where the authors construct portfolios made of firms with different AI exposure and show that, following the release of ChatGPT, high-exposure firms earned substantially higher abnormal returns (AR). Their results suggest that, in response to significant AI breakthroughs, investors should revise their expectations of future cash flows for AI-exposed firms upwards.

This AI-specific evidence aligns with broader literature showing that technology capability and digital transformation are positively priced. Brynjolfsson and Hitt (2000) argue that technology complements organizational changes and innovation, leading to higher productivity and business performance. Bharadwaj (2000) demonstrates that firms with superior IT capability achieve higher profitability and market-based performance. Taken together, this literature supports three key mechanisms that directly motivate the present research.

First, AI-related investments and capabilities foster productivity, innovation, and growth, which markets should capitalize into higher valuation multiples. Second, markets react positively to strategically important IT and digitalization announcements, indicating that investors anticipate long-run benefits from such technologies. Third, text-based measures extracted from corporate communications and regulatory filings are informative about both firm fundamentals and investor perceptions. These measures systematically link to stock price reactions and longer-run valuation outcomes.

Based on this reasoning, we expect AI focus to operate as a priced firm characteristic. Greater emphasis on AI in 10-K disclosures reflects firms' strategic commitment to AI rather than transitory communication choices. Accordingly, investors are expected to perceive firms with higher relative AI focus as possessing superior AI-driven growth opportunities and greater intangible capital, which markets should capitalize into higher firm valuation. Following the growing relevance of AI and the impact of company technology exposures to its valuation, this research gap leads to our first hypothesis:

H1) An increase in sector-standardized AI focus is associated with a relative increase in firm valuation

4.2. Methodology

To measure a firm's AI focus, we use an ordinary least squares (OLS) regression for financials from calendar year 2024. We run different models to maximize robustness and improve the quality of the results. In each iteration, a valuation multiple serves as the dependent variable while we use the firm's AI focus as the main independent variable. To improve robustness, we introduce a set of control variables.

AI Focus

To measure a firm's strategic emphasis on AI, we use the adjusted set of 200 current S&P500 companies. From those companies, we analyze the most recent filings for a full year, i.e. 2024, as well as 2023 for robustness. We construct the sample to ensure correspondence with the concurrent sentiment-analysis component of this paper, enabling direct comparability between word-count–based and sentiment-based measures of AI focus on the same firms. To then measure a firm's AI focus, we follow the approach from Mishra et al. (2022) and construct the metric for firm i as follows

$$(1)\ AI\_Focus_i = \frac{No.of\ AI\ Related\ Keywords_i}{No.of\ total\ words\ in\ 10-K_i} * 100$$

This measure represents the number of AI anchor matches per 100 words of narrative text. Scaling by document length addresses the heterogeneity in 10-K length across firms of different sizes and sectors. Without such normalization, large firms with longer 10-Ks would mechanically record higher AI hit counts, conflating firm size with AI attention.

Proxy Validation

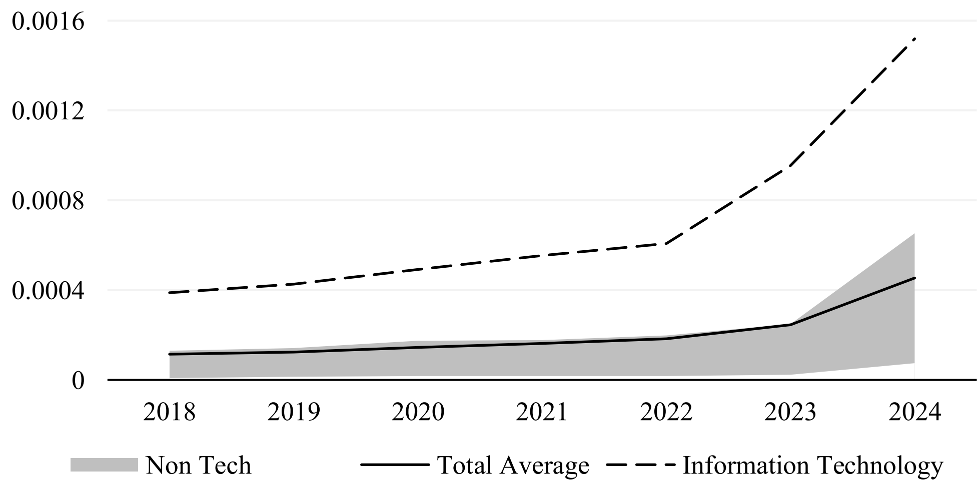
To establish the validity of the AI focus measure, we examine its relationship with tangible accounting inputs, adopting the quintile-based validation approach of Mishra et al. (2022). However, to control for cross-sectional heterogeneity in asset intensity and business models, we employ sector-standardized metrics (AI_z and SG&A_z) rather than raw accounting ratios.

Given the predominant role of firms as adopters rather than developers of AI technology, firms are more likely to reflect relevant investments in selling, general, and administrative (SG&A) expenses, capturing implementation and talent costs, than in capitalized R&D. Consistent with this hypothesis, the results in Table 10A indicate that firms in the highest quintile of AI focus exhibit significantly higher sector-relative SG&A expenses compared to those in the lowest quintile (0.018**)[1]. This confirms that firms with high narrative AI focus allocate significantly greater sector-relative resources to operational implementation, validating the measure as a proxy for substantive commitment.

Industry Differences

Our analysis shows that the average AI focus substantially varies across industries. Companies in the Information Technology sector experience by far the most mentions of AI-related keywords, illustrated in Figure 2. Business models that are often highly dependent on the technology drive this pattern and therefore require greater attention to the topic.

Figure 2: AI Focus across Sectors



[1] *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

To account for this, we standardize AI focus within each GICS sector to isolate the relative positioning of firms within their peer group. For each sector s, we compute the mean and standard deviation of raw AI focus across all firms in that sector:

$$(2)\ \mu_s = mean(AI\_Focus \mid s)$$

$$(3)\ \sigma_s = std(AI\_Focus \mid s)$$

The sector-standardized AI focus for firm i in sector s is then computed as:

$$(4)\ AI\_z_i = \frac{AI\_Focus - \mu_s}{\sigma_s}$$

This standardization has several advantages and follows the approach from Fresard (2010), where they used the z-score to make cash holdings comparable among different industries. It focuses on the relative, within-sector AI Focus of firms and mitigates the risk that sector-level differences in technology adoption distort correlations between AI focus and valuation.

Financial Data

The primary outcome variables are the logarithms of five widely used equity valuation multiples, which we compute for the same fiscal year as the 10-K. Tobin's Q captures the ratio of firm’s market value to accounting book value, reflecting investor expectations of future profitability and growth. Practitioners most frequently employ EV/EBIT and EV/EBITDA multiples in practice. They control for capital-structure and depreciation and amortization policies respectively, making it useful for comparing firms with different asset bases and leverage. The price-to-earnings ratio (P/E) measures the value of a firm’s equity relative to its current earnings. It captures how much investors are willing to pay for one unit of current earnings and is often interpreted as a summary statistic of growth expectations and perceived risk. The price-to-book ratio (P/B) measures the value of a firm’s equity relative to the book value of its shareholders’ equity. It reflects how the market values the firm’s net assets as recorded on the balance sheet.

After initially sourcing the data for the set of firms, we find that individual companies have artificially high or low multiples, which are caused by short-term valuation effects on the due date. To account for such outliers, we winsorize the dataset at the 1st and 99th percentiles, as in the existing research of Gaio and Raposo (2010). Furthermore, in the regression analysis, we enter all five multiples as natural logarithms, which helps reduce skewness and allows for the interpretation of the results as percentage differences in valuation across firms, rather than absolute differences. We exclude three firms from the final sample because they reported negative EBIT, which would render EV/EBIT valuation multiples difficult to interpret. Furthermore, we remove an additional four companies with a negative P/B because liabilities exceed the firm's assets. Moreover, we remove Nvidia from the sample because Cook's distance diagnostics indicated a value of 1.58, which exceeds the commonly used cutoff of 1.00 for identifying highly influential observations (Fox, 2015, see Table 11A).

Control Variables

To account for potential confounding influences, we use control variables that capture standard determinants of firm valuation, following the research of Chen and Srinivasan (2022). We include these control variables in the main model, with slight adjustments for the profitability metric in additional models for robustness.

Firm size, measured as the natural logarithm of market capitalization, controls for cross-sectional differences in scale and access to capital. We capture profitability using different metrics for each valuation multiple. Firms with higher profitability typically generate more cash flow per unit of sales and therefore often achieve higher enterprise value (EV) based multiples. In the robustness specifications, we adapt the profitability proxy to match the respective dependent variable. We use the debt-to-equity ratio (D/E) to capture firms' capital structure choices and financial risk by relating total interest-bearing debt to shareholders' equity. We capture revenue growth using the cumulative annual growth rate of sales from 2021 to 2024, deliberately starting after the peak COVID-19 disruptions in 2020 to avoid transitory pandemic effects on the levels and volatility of revenues.

A Tech × AI_z interaction term captures whether the valuation effect of AI focus differs between tech and non-tech firms. While AI_z measures the firm's overall AI focus, tech firms may already be more digital and intangible-intensive, so an additional unit could have a different marginal impact on valuation than in traditional sectors.

Regression

We use a cross-sectional OLS regression for fiscal year 2024 as the primary analysis. The baseline specification is:

$$(5)\ \ln(valuation_i) = \alpha + \beta_1 * AI\_z + \beta * controls + \varepsilon$$

where $\ln(\text{valuation}_i)$ is, in separate regressions, ln(Tobin's Q), ln(EV/EBITDA), ln(EV/EBIT), ln(P/B), or ln(P/E). Following the research from Lie and Lie (2002), who suggest that for nonfinancial large-cap firms, book value-based multiples lead to the most precise valuations, we use Tobin's Q for our main models. The analysis aims to test whether firms with higher AI focus trade at systematically different valuation multiples after controlling for fundamentals, not to forecast cross-sectional variation. Hence, we prioritize sign, magnitude, and statistical significance of β over the overall R^2, which is often modest even in well-specified cross-sectional financial regressions.

4.3. Results

In this section, we report the regression estimates linking changes in firms' AI focus to valuation multiples, using various dependent variables and model specifications. We begin with a baseline cross-sectional regression using ln(Tobin's Q) as the dependent variable, then test robustness across alternative valuation measures, and finally assess the impact of using a lagged AI focus measure. All regressions include key controls, and we conduct thorough diagnostic tests to ensure the validity of our results.

Baseline Regression: Tobin's Q

Table 1 presents the baseline cross-sectional regression results for the fiscal year 2024, utilizing ln(Tobin's Q) as the dependent variable.

The sector-standardized AI focus (AI_z) exhibits a positive association with firm valuation (0.067**). In the log-linear specification, this implies that, holding other factors constant, a one-standard-deviation increase in a firm's AI focus relative to sector peers is associated with an approximate 6.7% increase in Tobin's Q. The magnitude of this effect indicates an economically meaningful valuation premium for firms that place greater emphasis on AI than their industry peers.

Regarding the control variables, profitability and firm size emerge as the most significant determinants of valuation in this sample. Return on Assets (ROA) displays the largest effect magnitude (5.786***), confirming that profitability is the primary driver of the market-to-book ratio. Firm size, measured by the natural logarithm of market capitalization, also shows a positive and statistically significant coefficient (0.132***). Conversely, the coefficients for Revenue Growth (0.048), the Debt-to-Equity ratio (0.006), and the binary Tech sector indicator (0.030) are statistically insignificant, implying that these factors do not explain the residual variation in valuation after accounting for profitability and size in this specific model specification.

Finally, we test whether the valuation effect of AI is specific to the tech industry by examining the interaction term between AI_z and the tech sector dummy. The coefficient on the interaction term is -0.106 with a standard error of 0.079. This result is statistically insignificant, meaning we fail to reject the null hypothesis that the slopes are equal. Consequently, the results indicate that the positive association between AI focus and Tobin's Q does not differ significantly between tech and non-tech firms. Instead, investors assign a structurally similar marginal valuation benefit to increased AI focus across both tech and non-tech sectors.

Table 1: Model Coefficients – Tobin's Q

Predictor	***Estimate***	***SE***	***t***	***p***
Intercept	-2.945	0.624	-4.721	<.001***
AI_z	0.067	0.030	2.193	0.030**
ln_MarketCap	0.132	0.026	5.083	<.001***
Tech (1–0)	0.0303	0.074	0.650	0.516
Growth	0.048	0.369	0.387	0.699
ROA	5.786	0.449	12.888	<.001***
D/E	0.006	0.004	1.486	0.139
AI_z × Tech (1–0)	-0.106	0.079	-1.335	0.183

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Overall, the significance of key controls supports the validity of the model and indicates that the AI_z effect reflects valuation differences above and beyond established determinants such as firm size and profitability. Regarding model fit, the Tobin's Q regression yields $R^2 = 0.64$ and adjusted $R^2 = 0.66$ with N = 194 observations. Thus, the included regressors explain over 60.0% of the cross-sectional variation in Tobin's Q, indicating a strong explanatory capacity for a corporate-finance valuation model.

We use multiple diagnostic tests to confirm that the model satisfies the OLS assumptions. In Table 12A, Cook's distance values are small, suggesting no influential observations. Further tests in Table 13A confirm the assumptions. The Durbin–Watson test shows a statistic of 2.15, with a high p-value, indicating no evidence of autocorrelation, consistent with the cross-sectional nature of the dataset. Variance inflation factors (VIF) for all regressors fall between 1.01 and 1.32, far below conventional thresholds, showing that multicollinearity is minimal and coefficient estimates are stable. Ultimately, the application of the Shapiro-Wilk test proves normal distribution. Collectively, these diagnostics support the adequacy of the linear specification and provide confidence in the reliability of the inference drawn from the Tobin's Q regression.

Robustness: Alternative Valuation Metrics

To assess whether the valuation premium associated with AI focus is robust, we re-estimate our baseline model using different dependent variables. Table 2 presents regressions using ln(EV/EBITDA), ln(EV/EBIT), ln(P/E), and ln(P/B) as outcomes. All models retain the same set of predictors, and we adjust profitability measures to align with each multiple.

Across all valuation models, the coefficient on AI_z remains positive, although the degree of statistical significance varies. In the EV/EBITDA regression, AI_z is positive ($\beta \approx 0.078^{*}$), implying that a one-standard-deviation increase in AI focus corresponds to an approximate 7.8% increase in the EV/EBITDA multiple. In the EV/EBIT regression, the effect is slightly larger ($\beta \approx 0.117^{***}$), corresponding to an 11.7% increase. This consistency follows from the structural similarity between these EV-based multiples, both of which abstract from capital structure differences and focus on operating performance. By contrast, the evidence for P/E and P/B is more muted, as shown in Table 2. Although AI_z enters positively in both specifications, it is statistically insignificant for P/E ($\beta \approx 0.052$) and for P/B ($\beta \approx 0.049$).

The AI_z × Tech interaction term exhibits a consistent pattern across specifications but changes in significance. The interaction coefficient is negative in all models, indicating that the incremental valuation effect of AI focus tends to be smaller for tech firms than for non-tech firms. However, the strength of this effect varies. In the EV/EBITDA and EV/EBIT regressions, the interaction is slightly statistically significant, suggesting a difference between tech and non-tech firms in these EV-based multiples.

Taken together, these robustness checks support the main conclusion that AI focus is value relevant. AI_z is consistently positive and significant in the EV/EBITDA and EV/EBIT specifications, while it is insignificant in P/E and P/B ratios. This pattern suggests that markets mainly capitalize AI-related activities in valuation metrics that (i) abstract from short-term accounting noise and capital structure (EV-based multiples) or (ii) capture broader expectations about long-run intangible capital (Tobin's Q).

Table 2: Model Coefficients – Robustness Checks

Variable	EV/EBITDA	EV/EBIT	P/E Ratio	P/B Ratio
R^2	0.204	0.082	0. 125	0.553
Adjusted R^2	0.233	0.048	0. 092	0.535
AI_z Estimate	0.078	0.117	0. 052	0.049
AI_z SE	0.036	0.044	0. 041	0.064
AI_z t	2.148	2.629	1.284	0.760
AI_z p	0.033**	0.009***	0. 201	0.448

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

Robustness: Lagged AI Focus (2023)

An additional robustness check examines whether the timing of the AI focus measure affects our results, as the main analyses only use AI_z measured in 2024. To validate the hypothesis that 2024 was not subject to unusually high AI focus, we re-ran the baseline regression with Tobin's Q using the prior year's AI focus (AI_z for 2023) instead of 2024. The relationship remains qualitatively similar and significant as shown in Table 3.

Table 3: Model Coefficients – Lagged 2023 AI_z

Predictor	Estimate	SE	T	p
Intercept	-2.953	0.623	-4.740	<.001***
AI_z	0.066	0.030	2.194	0.030**
ln_MarketCap	0.133	0.026	5.117	<.001***
Tech (1–0)	0.048	0.074	0.650	0.516
Growth	0.145	0.368	0.401	0.689
ROA	5.764	0.448	12.878	<.001***
D/E	0.006	0.004	1.477	0.141
AI_z × Tech (1–0)	-0.105	0.079	-1.334	0.184

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

Control Variables

Across all regression specifications, the control variables behave mostly as expected, and a series of diagnostic checks indicates that the models are well-specified. Firm size is a consistently positive predictor of valuation multiples in every model. This implies that larger firms have higher valuation ratios even after accounting for performance metrics. The EBITDA Margin is also positively associated with valuations. A higher margin signals efficiency and pricing power, which investors reward with higher EV/EBITDA and EV/EBIT multiples (Novy-Marx, 2013). Growth remains a strong and significant control: firms with higher growth rates uniformly trade at higher Tobin's Q and P/E, reflecting that growth opportunities are a critical component of valuation.

The interaction term between AI focus and tech captures whether the valuation effect of AI_z differs across sectors. This interaction term is negative and non-significant, suggesting that the marginal valuation premium associated with AI focus does not differ meaningfully between tech and non-tech firms. In effect, while tech firms generally trade at higher multiples, the AI-related valuation effect appears to be broad-based.

AI Focus by Sector

As a descriptive robustness check, we compare raw AI focus (non-weighted) between tech firms and non-tech firms. We use an independent-samples t-test with the tech dummy (Tech = 1 for Information Technology firms, 0 otherwise) as grouping variable and report Welch's t and a Mann–Whitney U test to account for unequal variances and non-normality.

Results in Table 14A show that tech firms mention AI substantially more often than non-tech firms. The mean AI focus for non-tech firms is 2.47 bps, compared with 13.9 bps for tech firms, roughly a six-fold difference. The median intensity is about ten times higher for tech firms (13.2 bps vs. 1.28 bps). Welch's t-test indicates that this difference is statistically significant

(0.001***), with a very large effect size (Cohen's d ≈ 1.48). The non-parametric Mann–Whitney U test confirms this result ($U = 606$, $p < 0.001$).

Overall, these findings validate the expectation that AI-related language is much more prevalent in tech firms' disclosures. At the same time, our main regressions rely on the sector-standardized AI_z measure, which captures a firm's AI focus relative to its own sector. Taken together, the tests show that tech firms lead in absolute AI focus, while AI_z allows us to study within-sector variation without conflating it with broad sectoral differences.

4.4. Discussion

This analysis empirically examines the relationship between corporate AI focus and firm valuation among S&P 500 companies. Our results provide robust evidence that the extent to which firms strategically emphasize AI, measured by the AI focus in mandatory filings, is a priced characteristic in capital markets and reflects long-term expectations.

First, we document a statistically significant, positive association between sector-standardized AI focus (AI_z) and firm valuation. Specifically, a one-standard-deviation increase in AI_z corresponds to an approximate 6.7% increase in Tobin's Q, suggesting that capital markets price AI engagement as a determinant of future firm value.

Second, we observe a distinct divergence between valuation metrics. The valuation premium is robust across EV and asset-based multiples (Tobin's Q, EV/EBIT, EV/EBITDA), which capture long-term growth options and intangible capital. Conversely, the relationship is statistically insignificant for equity-based multiples (P/E, P/B), indicating that AI focus is priced as a strategic asset rather than a driver of immediate residual equity earnings.

Third, our results demonstrate sector neutrality. While tech firms naturally exhibit a higher absolute AI focus, the marginal valuation benefit of relative AI focus does not differ significantly between tech and non-tech sectors. This is in line with Babina et al. (2024) and supports the view of AI as a general purpose technology (GPT) where strategic differentiation yields valuation premiums regardless of the industry (Calvino et al., 2025).

This study therefore makes several contributions to the literature on the economic value of intangible assets, information technology, and corporate disclosure.

First, this paper extends research on the market valuation of intangible technological assets. Prior literature has extensively documented that firms with high levels of intangible capital, often unrecorded on balance sheets, exhibit elevated market-to-book ratios (Lev & Sougiannis, 1996; Peters & Taylor, 2017). While previous studies have relied on R&D expenditures or patent counts as proxies for innovation, this measure often fails to capture the strategic adoption of technologies that do not generate proprietary intellectual property but enhance operational capability. We address this gap by establishing narrative AI focus as a valid proxy for strategic intangible capital. Our finding that AI_z predicts Tobin's Q but not P/E aligns with Gu and Lev (2016), who argue that current earnings are increasingly noisy measures of value for innovation-intensive firms. Furthermore, direction and significance of the coefficients across models are consistent with prior work showing that markets increasingly reflect intangible investments in firm valuations, while also highlighting the importance of using multiple valuation measures and accounting for sectoral context when studying intangible-driven effects (Basso et al., 2014). Adding on to this, EV-based multiples can show stronger relations than equity multiples because EV is less distorted by leverage. Consequently, EV-based multiples often track operating-value variation more cleanly, as they are not affected by financing (Lie & Lie, 2002). However, the insignificance of P/E still implies a disconnect between market capitalization and current net income. Since P/E is a function of current profitability, our result suggests that AI adoption in the current cycle does not immediately translate into the bottom-line efficiency gains observed in prior cycles.

This stands in contrast with the findings of Mishra et al. (2022), who document a positive association between AI focus and return on sales, indicating that AI adoption was accretive to margins during their sample period from 2005 to 2019. This empirical divergence likely reflects a structural shift in the nature of the technology adoption. Mishra et al. (2022) captured an era driven by simple use cases defined by Agrawal et al. (2018) as "prediction machines", technologies primarily used to automate discrete prediction tasks to drive operational efficiency. In contrast, AI functions as a GPT requiring significant "co-invention" of new business processes before returns are realized (Brynjolfsson et al., 2021).

Consequently, our 2024 sample captures firms in a later stage with more complex implementations, elevating current operating expenses without yielding realized profit improvements. In addition, firm size stands out as the single control variable that behaves differently from prior findings. While earlier research from Fama and French (1992) documents a negative relationship between firm size and valuation, we observe a positive association in our sample. This divergence may be driven by the composition of the S&P 500, which consists exclusively of large-cap firms and is disproportionately dominated by tech companies, particularly large ones that tend to exhibit a stronger focus on AI than peers.

Second, this study extends the literature on the business value of IT and industry heterogeneity. We extend this strain by examining the cross-sectional applicability of AI valuation premiums across industry verticals. We find a difference in raw AI focus, i.e. non industry weighted, between firms classified as Information Technology within GICS sectors and companies that are not. This indicates that tech companies focus on AI a lot more than non-tech firms at an effect size of -1.47 and aligns with the assumption that tech firms, whose business model often is AI related, engage with the technology more often. Existing research from Bardhan et al. (2013) as well as Bharadwaj (2000) finds similar results, suggesting that increased investments and adoption of IT go in line with higher Tobin's Q due to increased intangible assets. While our descriptive statistics confirm that tech firms discuss AI significantly more than non-tech firms, our regression analysis reveals a non-significant interaction term between AI focus and the tech sector. In practical terms, a manufacturing or retail firm with high AI focus sees a similar valuation premium as a tech firm with comparable AI_z, all else equal.

This sector-neutral result underscores that AI as a long-term value driver is widespread and not confined to the traditional technology industry. Instead, it appears that within any given sector, firms that focus more on AI compared to their industry peers tend to be valued more highly, suggesting a broad-based market belief in the value of AI-related strategy. We contribute to the literature by providing empirical evidence that the capital market rewards AI focus symmetrically across industries, rather than being driven solely by the tech sector.

Third, this research contributes to the literature on corporate disclosure and the valuation of intangible capital. Under current accounting standards, most internally generated intangibles and therefore AI-related investments are expensed rather than capitalized, leading financial statements to understate firms' technological capital (Gu & Lev, 2016). Prior work shows that markets nonetheless capitalize such intangible investments into firm value, as reflected in elevated market-to-book ratios and Tobin's Q (Peters and Taylor, 2017). Our findings extend this literature by demonstrating that a simple, dictionary-based measure of AI focus extracted from 10-K filings carries incremental explanatory power for firm valuation beyond standard R&D measures and financial controls. This suggests that narrative AI focus provides a tractable proxy for AI-related intangible capital that is not observable in balance-sheet data, consistent with the growing use of textual analysis to quantify otherwise unmeasured firm characteristics (Bardhan et al., 2013).

These findings offer critical, actionable guidance for corporate executives and investment professionals regarding the strategic management of AI focus. The most profound implication for practitioners is that capital markets actively price organizational commitment to AI, though investors specifically realize this valuation premium in EV-based multiples, suggesting that they interpret AI commitment as a structural driver of total firm value and long-term growth potential. This interpretation aligns with the theoretical framework of Lev and Sougiannis (1996), who establish that the market capitalizes intangible investments into elevated market-to-book ratios. Consequently, placing strategic emphasis on AI within 10-K narratives is a critical imperative for companies aiming to maximize their competitive standing, regardless of industry. Because the validation that this narrative proxy correlates with higher SG&A expenses confirms that the market prices a real allocation of resources toward implementation and talent, investment professionals should interpret high relative AI focus as a signal of organizational transformation that standard financial statements fail to capitalize.

Managers should therefore recognize that firms exhibiting lower relative AI focus, interpreted as weaker strategic commitment to AI, are priced at a valuation discount, which can translate into a competitive disadvantage in capital markets. This creates a substantial opportunity for legacy firms to unlock shareholder value by clearly signaling a deviation from the sector norm, thereby allowing the market to price future efficiency gains and innovation premiums.

5. AI Sentiment and Stock Price Reactions

5.1. Hypothesis

As discussed in the previous part, the current state of academic research focuses on long-term benefits from AI that firms achieve through an increased operational efficiency through cost reductions and enhanced decision-making (Agrawal et al., 2019; Mishra et al., 2022; Shiyyab et al., 2023). However, AI implementation can also be costly, complex and thus risky to implement. The process can pose reputational risks or result in revenue losses if systems fail, which then results in negative CARs (Lui et al., 2021). This ambiguity mirrors the state of the broader IT investment literature. While selected research finds a positive relationship between IT investments and firm value (Dehning et al., 2003; Subramani & Walden, 2001), other studies find adverse market reactions (Bose et al., 2011). Since most previous research focuses on long-term effects, which are measurable only with a certain lag, confounding factors that accumulate over time may affect the observed relationship. This motivates the development of hypotheses on investors' short-term assessment of AI-related information, where price reactions occur when such information becomes available to the market. For that reason, the event study methodology reflects a complementary approach to observing investors' immediate assessment of newly available information and thereby provides an appropriate empirical basis for testing our hypotheses on the short-term market reactions.

In one of the most influential works on sentiment analysis and the tone of information, Tetlock (2007) finds that pessimism in news articles exerts downward pressure on stock prices. Loughran and McDonald (2011) further improve this fundamental work by developing a finance-specific bag-of-words or dictionary approach widely used to analyze reports for sentiment. While this extends to earnings announcements, press releases or news reports, regulatory filings play an important role as well. Researchers predominantly use 10-K filings, particularly the MD&A section which helps inform investors about going-concern considerations and is directly linked to both a firm's profitability and thus also stock price reactions (Brown & Tucker, 2011; Mayew et al., 2015; Sun, 2010). These reports are regulatorily mandated and standardized, thus providing an excellent source for the sake of our analysis, allowing us to capture the specific tone towards AI. For that reason, we analyze both Item 1 (Business) and Item 7 (MD&A). The MD&A reflects the current opinion of top-level management, which is important for expectations of the future. The Business section adds to that by explaining the business model and environment.

Lastly, relevant literature focuses on managerial optimism and its implications on market reactions. Some parts of the literature view this management stance as risky because it can lead to biased decisions and an elevated risk of failure because optimism does not always materialize (Jiang & Liu, 2019; Malmendier & Tate, 2005). Others find a correlation between optimism and strategic choices as well as subsequent corporate performance induced by improving operational efficiency and reducing analyst forecast dispersion (Wang et al., 2024). From an agency-theoretic viewpoint (Jensen & Meckling, 1976), sentiment is important because it forms the expectations of shareholders. In the context of AI, where projects are complex and outcomes might be uncertain, managerial optimism, as found in the 10-K reports, is extremely important for the informed investor.

Overall, while AI implementation entails substantial costs and risks, existing evidence suggests that the net effect on firm value is positive in the short run. Prior work further indicates that investors use managerial sentiment in textual reports to price information and finance-specific dictionaries allow this tone to be measured in standardized reports. Related to that, the literature on managerial optimism suggests that an optimistic management tone can shape firm success and thus market expectations. Therefore, we derive our second hypothesis:

H2) A positive management stance towards AI in 10-K reports is positively accepted by investors, as can be seen by subsequently observed positive CARs

5.2. Methodology

Sentiment Analysis

Our research approach is divided into two different parts. First, gathering the relevant information on financial statements and management tone requires intensive data mining and collection. Our data consists of 200 S&P companies that have been part of the index continuously from 2017 to 2024. In total, we collect 1600 unique data points. Three of those instances reported unexpected errors. Thus, we cannot fetch the 10-K data for companies with the following tickers and years: "HWM" in 2017, "KMX" in 2019 and "MAR" in 2021.

Second, to filter for sentiment but solely regarding AI, we combine the dictionary approach (Loughran & McDonald, 2011) with the KWIC approach, as introduced by Luhn (1958). More concretely, our script examines all those 1597 reports for so-called AI-anchors (see Table 20). If our script finds such an anchor word in Item 1 or Item 7 of the 10-K report, it then analyzes the sentiment ten words before and ten words after the anchor by matching it with the list of positive and negative words that Loughran and McDonald (2011) define. Thus, each anchor word is associated with several positive and negative sentiment words, in case such are found in the defined perimeter. The labeling of the AI trigger word, as positive or negative, is based on the absolute frequency of sentiment words. Meaning an anchor is categorized as positive if more positive than negative sentiment words are found in the vicinity of an anchor and vice versa. If the script finds an AI anchor word but no sentiment trigger, it labels the stance toward AI as neutral. If it finds no AI words in the first place, the script stops and labels the company as "No Stance". Based on the number of those categorized AI anchors, we label companies' stances on an annual basis. So, if more AI anchors are labeled positive than negative, we label the company's stance for that corresponding year as positive and vice versa. Based on that, we divide our sample into four categories, which Table 4 shows below. For the summary statistics in Table 4, the “No Stance” category has declined almost continuously from 182 instances in 2017 to 119 in 2024, while all other categories grew.

Table 4: Sentiment Classification of 10-K Reports from 2017 - 2024

	2017	2018	2019	2020	2021	2022	2023	2024	Total
Positive Stance	10	14	16	19	32	41	43	53	228
Neutral Stance	2	5	8	13	10	9	8	11	66
Negative Stance	5	6	8	6	10	15	13	17	80
No Stance	182	175	167	162	147	135	136	119	1,223
Total	199	200	199	200	199	200	200	200	1,597

The data in general is based on the results of our text mining operations for Item 1: Business and Item 7: MD&A of the annual 10-K Report

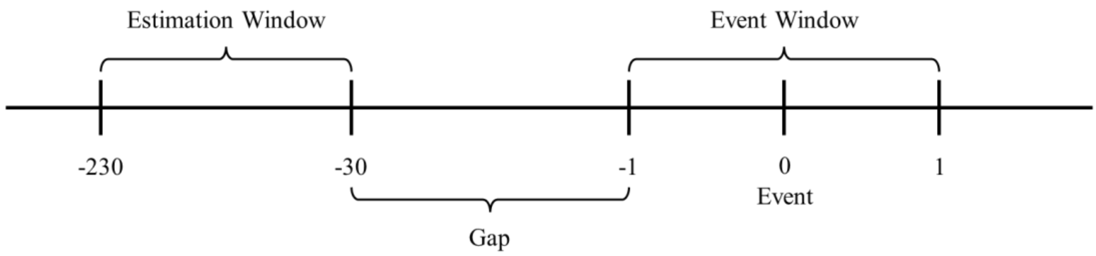
Especially the “Positive Stance” equaled or exceeded the sum of the “Neutral Stance” and “Negative Stance” categories in every year we observe.

This indicates an increased importance of AI in relation to the filing of annual reports. Additionally, it supports the importance of introducing sentiment analysis to the field of AI and corporate valuation, as there seems to be a positive inclination regarding the management’s stance. This positive inclination could also be evidence for the optimism bias.

Event Study Methodology

To measure the direct and short-term impact of those filings, we conduct a classic event study, where we use the reaction of the share price to an event to determine the immediate effects on company valuation (Kothari & Warner, 2007; MacKinlay, 1997). The model for the prediction can be seen below in Figure 3.

Figure 3: Theoretical Model for the Event Study Methodology



Source: based on MacKinlay (1997)

The event itself is the annual filing of a company's 10-K report. Our estimation period starts 230 days before the defined event and goes until 30 days prior to the event. We use the gap between the estimation window and the event window to ensure that both are separate observations and that no cross-correlation influences the significance of our results (Luoma, 2011). To the best of our knowledge, this is the most widely used estimation window in academics, which is also aligned with MacKinlay (1997) and Lui et al. (2021). To evaluate the magnitude of events, we compute the $AR_{i\tau}$. As equation (6) shows, we compute the ARs by deducting the expected return $E(R_{i\tau}|X_{\tau})$ from the actual observed return $R_{i\tau}$, where i stands for the individual company on event day τ (MacKinlay, 1997).

$$(6)\ AR_{i\tau} = R_{i\tau} - E(R_{i\tau}|X_{\tau})$$

The expected return, on the other hand, we compute by using a model such as the constant return model or another statistical or economic model. In our case, we choose the Capital Asset Pricing Model (CAPM), which, according to Lui et al. (2021), is the most widespread market model, in order to estimate the AR. The formula for the CAPM can be seen in equation (7).

$$(7)\ E(R_i) = R_f + \beta_i(E(R_m - R_f))$$

Utilizing the EasyEventStudy library in Python, we derive all market data, such as the risk-free rate and the market return, from the Center for Research in Security Prices. To measure the stock price response over the whole window, we cumulate the individual $AR_{i\tau}$, resulting in equation (8), which portrays the formula for computing the CAR_i (MacKinlay, 1997).

$$(8) CAR_i(\tau_1, \tau_2) = \sum_{\tau = \tau_1}^{\tau_2} AR_{i\tau}$$

Based on these results, we evaluate whether the CARs are statistically different from the null hypothesis, meaning expected CARs of zero.

Data Validation

To assess which statistical tests are appropriate for evaluating the null hypothesis, we examine whether the CARs satisfy the normality assumptions. First, we utilize the Shapiro-Wilk, Kolmogorov-Smirnov and Anderson-Darling tests to check for normality and validate our results (Table 15A). Across all tests, we reject the null hypothesis at conventional significance levels, with p-values below 0.001, indicating that the ARs and CARs of our dataset are not normally distributed.

Second, we test the data for heteroskedasticity, autocorrelation and collinearity as can be seen in Table 16A-18A in the appendix. Since we are only able to reject multicollinearity[2], the data does not support all the necessary assumptions for a regression analysis. However, the central limit theorem justifies the applicability of the parametric z-test statistic, and earlier works by Dos Santos et al. (1993) and Lui et al. (2021) also apply this approach. We use both the nonparametric Wilcoxon signed-rank test and the sign p-test to improve the quality of our results. Especially the latter is ideally suited for event study methodologies, considering that it specifically tests the distribution of signs. Additionally, neither test requires normality and both are more robust to volatility compared to parametric tests such as the t-test or z-test (Campbell et al, 1998).

To determine whether the different management stances and control groups differ statistically from each other rather than from the null hypothesis alone, we consider two tests. First, the Kruskal-Wallis test analyzes whether one of the groups is statistically different from all others, allowing for a simple first indication (Table 7). It requires the observations from both groups to be independent from each other and at least ordinally scalable (Kruskal & Wallis, 1952). Second, we utilize a pairwise Mann-Whitney U test to determine whether the actual individual

[2] The p-value of the Durbin-Watson test statistic is statistically significant (0.036**), yet leads to believe that the data still suffers from a mild autocorrelation of residuals

groups are different from one another (Table 8 and Table 19A). This test has the same requirements as the Kruskal-Wallis (Nachar, 2008).

5.3. Results

Event Window

Table 5 shows the CARs for all 200 companies over the span of eight years for different lengths of event windows.

Table 5: Overall Impact of 10-K Releases on the Stock Price

CAR	Mean	Median	z-test p	Wilcoxon p	Sign p	N
[-1]	0.060	0.031	0.168	0.193	0.515	1,597
[0]	0.080	0.055	0.183	0.191	0.031**	1,597
[1]	0.045	0.057	0.323	0.223	0.072*	1,597
[-1,0]	0.140	0.081	0.053*	0.059*	0.121	1,597
[0,1]	0.124	0.079	0.107	0.132	0.133	1,597
[-1, 0, 1]	0.185	0.091	0.030**	0.030**	0.064*	1,597

Mean and Median are displayed in basis points
**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

In relation to Lui et al. (2021), we decide to include single-day windows and multiple-day windows to analyze whether a significant impact occurs on standalone days or over an interval of days.

Analyzing the single-day ARs, both median ARs of 0.055 bp (0.031**) for day [0] and 0.057 bp (0.072*) for day [1] show significance for the nonparametric Sign p-test, while day [-1], previous to the event, shows no statistical significance at all. Consequently, the mean ARs for all three single-day windows fail to show significance for the parametric z-test and the nonparametric Wilcoxon signed-rank test. Observing the CARs over the different multiple-day windows, we find no statistical significance for the [0, 1] window, while the [-1, 0] window shows a light significance for the parametric z-test (0.053*) and also the nonparametric Wilcoxon signed-rank test (0.059*). The window [-1, 0, 1] shows the strongest statistical significance as stated by both the parametric and nonparametric tests for the mean CAR of

0.185 bp and the median CAR of 0.091 bp. Especially, the z-test shows a statistical significance below five percent (0.030**) and so does Wilcoxon's test (0.030**), while the Sign p-value (0.064*) shows a lighter statistical significance.

Taken on their own, these results do not yet allow us to draw any conclusions about the validity of our hypothesis, but they show that CARs exist around the filing date and serve to identify a suitable event window, on which our further analysis, which includes the management sentiment, should focus.

Management Sentiment

Therefore, when analyzing the CARs for the individual management stances, our focus relies on that single window covering all the event days [-1, 0, 1], as can be seen in Table 6.

Table 6: Impact of Management Sentiment towards AI on Stock Prices

CAR	Mean	Median	z-test p	Wilcoxon p	Sign p	N
Positive Stance	0.347	0.313	0.229	0.163	0.098*	228
Neutral Stance	-0.194	-0.097	0.583	0.383	0.435	59
Negative Stance	-0.192	0.025	0.600	0.791	1.000	87
No Stance	0.199	0.100	0.031**	0.039**	0.123	1,223

Mean and Median are displayed in basis points
**** p < 0.01, ** p < 0.05, * p < 0.10*

At first glance, it might seem that the main part of the statistical significance of CARs in general is driven by the "No Stance" category. Both the z-test (0.031**) and Wilcoxon's p-value (0.039**) are only significant for the "No Stance" group. However, the Sign p-test statistic shows a slight significance (0.098*) for the median CAR of 0.313 bp for the "Positive Stance". In this case, the Sign p-value shows that for the "Positive Stance", the probability of positive and negative CARs is not equal. Instead, the "Positive Stance" category is significantly more likely to be followed by positive CARs than negative CARs. On the other hand, the results for both the "Negative Stance" and "Neutral Stance" categories reveal no such significance, especially the Sign p-test for "Negative Stance" indicates that we cannot reject the null hypothesis of equal probability of positive and negative CARs. A similar result holds for the "Neutral Stance", where the null hypothesis cannot be rejected as well.

So far, we have shown that the CARs for the "No Stance" and "Positive Stance" groups are partially but significantly different from zero. Even so, this indicates just the potential attractiveness of a "Positive Stance" towards AI, considering that it might still be more attractive not to mention AI at all, than to take a neutral or negative stance. We conduct all significance tests in the context of the null hypothesis, meaning that CARs are centered around zero.

To compare the management stances, we conduct a Kruskal-Wallis test to determine whether the CARs are significantly different across the categories. The test result (0.392) leads us to conclude that no single management stance has a CAR distribution that differs from those of all other management stances. Consequently, we cannot reject the null hypothesis that the CARs across all sentiments are equal. Yet, it is still possible that pairs of particular management stances differ from each other.

Table 7: Testing Differences between Sentiments

Kruskal-Wallis Statistic	df	p-value
2.999	3	0.392

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

To give our simplified results further robustness, we conduct a pairwise Mann-Whitney U test (see Table 8). Again, the results show that neither sentiment is significantly different from the other. The comparison of the "Positive Stance" and "Negative Stance" yield no statistical significance (0.262) and neither does the "Positive Stance" and "No Stance" pair (0.465). Thus, we cannot demonstrate a statistical difference between having a "Negative Stance" and a "Neutral Stance" toward AI. The lowest, yet not statistically significant p-value is between the "Positive Stance" and the "Neutral Stance" (0.151), while the highest p-value is shown for the "Negative Stance" and "Neutral Stance" pair (0.495).

Table 8: Pairwise Mann-Whitney Test for Sentiment Significance

	Group 1	Group 2	N1	N2	Statistic	p-value
1	Positive Stance	Negative Stance	228	87	9,107	0.262
2	Positive Stance	Neutral Stance	228	59	5,910	0.151
3	Positive Stance	No Stance	228	1,223	135,177	0.465
4	Negative Stance	Neutral Stance	87	59	2,738	0.495
5	Negative Stance	No Stance	87	1,223	50,549	0.437
6	Neutral Stance	No Stance	59	1,223	30,570	0.194

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

At first glance, these results indicate that there is no short-term difference between taking a management stance towards AI or taking none in broader terms, since we do not observe any differences between taking a positive, neutral or negative management stance towards AI. For that reason, we reject our initial hypothesis. We cannot say that a positive management stance leads to favorable investor reactions, as can be seen by subsequently observed positive CARs. Nonetheless, it is important to note at this point that the absence of evidence is not equal to the evidence of absence.

Industry Differences

These results could also indicate that important variables for determining the CARs have been omitted so far and should be accounted for.

For that matter, we conduct a further event study to examine the industry variable as a control variable. Table 9 shows the results for our final event study, based on the [-1, 0, 1] event window and categorized by the different industries according to GICS. We observe significant CARs for both the Consumer Staples and the Materials industries. For Consumer Staples, the parametric z-test (0.077*), the nonparametric Wilcoxon's signed rank test (0.045**) and the nonparametric Sign p-test (0.065*) show a slight significance. In comparison, Materials show significance for the nonparametric test results of Wilcoxon's test (0.043**) and the Sign p-test (0.025**). We do not conduct an additional Kruskal-Wallis test for industries because our objective is to identify pairwise differences between the individual industry pairs.

Table 9: Impact of Industry Differences on Stock Prices

GICS	Mean	Median	z-test p	Wilcoxon p	Sign p	N
Communication Services	0.929	0.525	0.258	0.198	0.260	64
Consumer Discretionary	-0.233	-0.055	0.425	0.635	0.835	206
Consumer Staples	0.428	0.329	0.077*	0.045**	0.065*	184
Energy	0.153	0.121	0.732	0.707	1.000	80
Health Care	0.227	0.206	0.224	0.150	0.295	264
Industrials	0.049	0.063	0.764	0.948	0.507	327
Information Technology	0.217	0.041	0.214	0.201	0.755	256
Materials	0.351	0.556	0.195	0.043**	0.025**	72
Real Estate	0.245	-0.199	0.325	0.902	0.279	144

Mean and Median are displayed in basis points
**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

An omnibus Kruskal-Wallis test would indicate whether any differences exist across industries overall and could therefore obscure more granular industry variation. Therefore, to evaluate whether we observe systematic differences between the different industries, we conduct another pairwise Mann-Whitney U test (Table 19A). This test compares each industry with every other industry individually to analyze whether the CARs significantly differ from each other. The evidence points to systematic differences among industry CARs. For Materials, the CARs differ significantly from the Consumer Discretionary CARs (0.094*) and Industrials CARs (0.058*). For Consumer Staples, on the other hand, we see a significant difference to the Industrials CARs as well (0.082*). These results indicate that industry differences across short-term observations exist.

5.4. Discussion

This research offers initial evidence for the relevance of management stances regarding the communication of AI. First, we find statistically significant CARs around the annual filing date for 10-K reports. When analyzing the management stances, our results show that most of the statistical significance is driven by the "No Stance" sentiment, while the "Positive Stance" also shows significant CARs for one of the nonparametric tests.

Second, we cannot observe statistical differences among the different management stances. Both the Kruskal-Wallis test and the Mann-Whitney U test fail to reject our null hypothesis. Thus, we cannot conclude that investors positively accept a positive management stance, because we observe no significant differences.

Third, we explore the importance of industry differences. Here, we see that the Materials and Consumer Staples industries had significant CARs around the filings of 10-Ks. Additionally, these results are robust because the distributions of these industries differed significantly from others. Although we reject our initial hypothesis, our findings point to the relevance of control variables.

In general, several results need to be highlighted. In Table 6, we see that apart from the "No Stance" category, only the "Positive Stance" of management on AI observed significant results for the nonparametric Sign p-test (0.098*), but lacking robustness, we cannot reject our null hypothesis. Yet, this hints at a potential attractiveness of positive management positioning towards AI in reporting. In combination with our results from the industry analysis, this underlines the meaningfulness of our research.

Regarding our theoretical contribution, we shift the focus to short-term reactions from capital markets that reflect how investors process corporate disclosures. Previous research predominantly examines long-term links between AI focus and firm value, while we focus on immediate CARs around 10-K filing dates. This is important because managers need to consider the signaling quality of AI towards market valuation.

In this regard, our results contradict those of Lui et al. (2021), who find a negative relationship between the announcement of AI investments and stock price reactions. Analyzing 119 announcements from 62 companies, they find that stock prices decreased by an average of approximately 1.8% upon announcement. They largely attribute their findings to contextual factors such as the IT capability, credit rating and type of industry. Comparing their study with ours, the different results can be due to many factors. Our sample size is much larger and thus more diverse. Additionally, they focus on 2015 to 2019, while our sample corresponds to the years from 2017 to 2024. As our summary statistics show, the importance of AI disclosure, which we measure using management stances, increased rapidly over time and remained very small in the early years.

Yet, our findings align with the broader positioning of academics on the impact of AI on corporate valuation, attesting a positive relationship between AI announcements and market value (Feldman et al., 2009). They also align with findings from IT investment literature, stating a positive relationship between IT investments and market reactions (Dehning et al., 2003; Subramani & Walden, 2001). Our findings contribute to the literature by showing that short-term market reactions can occur around mandated reporting dates in relation to AI. Additionally, our study provides novel empirical evidence based on a large sample of firms over a recent time period.

The second part of our research focuses on the importance of textual information and managerial tone by analyzing the explanatory power of management stances. Prior research predominantly finds that a positive management tone is associated with positive market reactions (Feldman et al., 2009; Yekini et al., 2016). Similar to both, we focus on regulatorily mandated reports that provide incremental information (Brown & Tucker, 2011). Building on this, we test whether differentiated management stances towards AI are systematically associated with CARs around the filing dates. Although we observe significant CARs around filing dates, the stance-based analysis yields no robust results, indicating that investors do not consistently price a positive management stance toward AI as more favorable than other stances.

Solely the "Positive Stance" yields marginally significant CARs, which aligns with the broader literature (Feldman et al., 2009; Yekini et al., 2016). The absence of statistical robustness has several plausible reasons. First, 10-K reports and especially the MD&A section yield relevant information, but they might just be incremental to previous announcements, such as press conferences. This becomes especially relevant considering that a company could hypothetically maintain a positive stance over several years without revealing any new information. As Brown and Tucker (2011) highlight, the same report with only slightly altered information does not yield much additional information, which gives little reason for potential stock market reactions. This aligns with the findings of Wang et al. (2024) regarding managerial optimism. Their findings state that the MD&A section can contain information relevant for decreasing investor forecasting dispersion if managers provide new information.

By contrast, our results are not entirely consistent with Tetlock's (2007) evidence on the predictive power of negative tone, at least in the specific context of AI-related MD&A language. However, Tetlock (2007) also states that downward price pressure is followed by a reversion to normality, which helps explain the insignificance of that sentiment. Additionally, Loughran and McDonald (2011) find that negative news is often framed by using positive words, while framing positive news using negative connotation is rarely the case. As argued by Jiang and Liu (2019) in their game-theoretical model, managerial optimism has limited power if the market as a whole has a positive stance already. Indeed, our results demonstrate that the "Positive Stance" does outweigh the "Neutral Stance" and "Negative Stance" in terms of observations, indicating a positive market bias, which would then result in weaker effects of a positive stance. This contributes to the current literature by suggesting that the valuation relevance of management tone regarding AI depends on the novelty and credibility of the information. Thus, it is important for companies to portray relevant information in a market that already has a positive bias towards a certain topic that is covered within the management communication.

Another potential reason for not finding any differences between the sentiment groups is the omission of relevant control variables. Potential variables are the IT capabilities, firm size or type of industry (Im et al., 2001; Lui et al., 2021; Subramani & Walden, 2001). For that reason, we explore the importance of industry variables as potential moderators for our analysis of the relationship between management tone towards AI and stock market reactions. Our results present evidence for the importance of industry factors. Considering the nature of Materials and Consumer Staples as posed by the GICS, these industries contain sub-industries such as Chemicals, Construction Materials, Metals & Mining and Food Products, Beverages and Household Products, which are highly associated with manufacturing. Lui et al. (2021) find that non-manufacturing CARs are significantly more negative than manufacturing CARs, which aligns with our findings for the Materials and Consumer Staples industry groups.

Here we see the CARs of Materials and Consumer Discretionary being significantly different, as well as Consumer Staples and Industrials and Materials and Industrials. While additional evidence for the individual industry differences is harder to find, more general trends can be identified to explain the industry differences. The GICS “Industrials” also considers “Commercial & Professional Services”. For those services, the applicability of typical AI use cases, such as process automation or monitoring, may be more limited because human interaction requires a lot of individuality (Mukherjee, 2022). Similar applies to “Consumer Discretionary”, considering this category includes both “Consumer Services” and “Consumer Discretionary Distribution & Retail”, which are more service-oriented and rely on individual customer interaction. However, it must be noted that, given the heterogeneity within these GICS categories and the limited ability to clearly separate manufacturing from non-manufacturing activities, any inferences drawn from this industry comparison should be interpreted with caution. Yet, our results suggest the valuation implication of AI-related management stances is context-dependent rather than uniform across firms from different industries.

This supports a view in which industry characteristics shape the perceived credibility and economic interpretability of AI disclosure, consistent with prior research on manufacturing and non-manufacturing differences (Lui et al., 2021).

In sum, our work has two implications for practitioners. For managers, capital markets react to 10-K filings, implying that AI-related disclosure is relevant for valuation and should be viewed strategically. But because the management stance differences are not robust, managers should not rely on optimistic framing alone. Instead, they should focus on increasing the informativeness of their AI disclosures, providing investors with new and relevant information. For investors, our results suggest that they should focus on incremental information and not the standalone management stance. Additionally, investors need to interpret AI disclosure through an industry lens. Market reactions show that industries are heterogeneous and AI communication yields different results based on the industries. This indicates that the perceived value of AI interactions is different depending on the industry.

6. Conclusion

This study examines whether capital markets capitalize corporate focus on AI into firm value. The first hypothesis addresses that increases in sector-standardized AI focus correlate with higher firm valuation. Using a cross-sectional sample of S&P 500 firms, the analysis constructs a text-based measure of AI focus from firms' 10-K filings and standardizes it within GICS sectors to capture relative AI positioning among peers. OLS regressions relate this measure to multiple valuation metrics while controlling for standard determinants of value, including firm size, profitability, leverage, and growth. The empirical results provide consistent support for the hypothesis. Sector-standardized AI focus is positively and statistically significantly associated with Tobin's Q: a one-standard-deviation increase in AI focus corresponds to an economically meaningful increase of approximately 6.7% in firm value. This valuation premium is robust to alternative EV based measures (EV/EBIT and EV/EBITDA) but is not statistically significant for equity-based multiples (P/E and P/B). The findings indicate that markets primarily capitalize AI engagement as a long-run intangible asset. Importantly, interaction tests show no significant difference between tech and non-tech firms, implying that the valuation effect of AI focus is broad-based rather than confined to the tech sector.

These results contribute to the finance literature on intangible capital, information technology, and corporate disclosure by demonstrating that narrative AI intensity in mandatory filings functions as a priced firm characteristic. In contrast to traditional proxies such as R&D, AI focus captures strategic adoption and implementation that are largely expensed under current accounting standards. The evidence suggests that investors interpret relative AI emphasis as a credible signal of future growth opportunities and organizational capability, thereby extending our understanding of how unstructured disclosure informs asset prices and firm valuation.

In the second part of this thesis, we aim to expand the first part by adding the component of qualitative communication to the quantitative approach. While we cannot prove that a "Positive Stance" consistently leads to positive CARs, we observe that the "Positive Stance" increased by more than five times during our observation period. This aligns and adds to the literature on general managerial optimism, stating only incremental value that provides little value for investors. This is meaningful evidence of the growing importance of AI in reporting, especially from a management perspective. Theoretically, this evidence suggests that capital markets react more to the presence of AI or the general AI focus than to the nuanced tone of managerial communication about AI itself.

In a last robustness check, we analyze the importance of industry factors. Here, we saw that certain industries, such as Consumer Staples and Materials industries, do experience significant CARs in relation to AI and in contrast to other sectors. A subsequent Mann-Whitney U test verifies these results, which highlights a statistically significant difference compared to other industries. This study makes a significant contribution to the extant literature by underscoring the imperative for industry-specific heterogeneity considerations when processing AI communication within a short-term window. Furthermore, it interrogates the incremental value of annual mandated reports, in which firms and investors can utilize information repeatedly over multiple years. Thus, this does not negate the lack of significance observed in the technology versus non-technology comparison in the first part. This offers a starting point for future empirical research.

7. Limitations

First, the external validity of our findings is limited to the geographical location and the size focus of our sample. While selecting filtered S&P 500 companies ensures consistency with the concurrent sentiment analysis part and reflects a large part of global markets, this approach restricts our findings both geographically and by firm size. The focus on large-cap equities means that our findings might not be applicable to smaller firms or those outside the USA. This is underscored by the fact that AI adoption remains highly driven by large-cap companies. Therefore, although we control for firm size in the model, the impact on valuation for smaller firms could be significantly different.

Second, our measurement of AI focus relies on mandated 10-K filings as a source of information. Rather than using annual mandated filings, news articles, or press conferences could be used to enable a broader information base, especially considering that mandated filings might only provide incremental value. For the calculation of AI focus, we use a dictionary-based approach. Although the methodology normalizes for document length to prevent conflation of firm size with AI focus and we validated AI_z with SG&A spending, it should be noted that this metric mainly captures strategic orientation rather than the magnitude of actual capital deployment. Consequently, although the data shows a correlation with valuation, we cannot confirm that a greater focus on narratives leads to a greater tangible investment.

Third, the event study methodology used in this study entails minor additional limitations. For the sake of this study, we define the event as the annual filing of the 10-K report. Given the incremental information nature of these reports, this choice implies that the absolute volume of disclosed information may fog the already marginal AI-related content that is relevant for investors. In doing so, we focus on levels rather than changes. Consequently, the measure does not capture incremental disclosures that would represent new information to investors. In addition, we see that not just the choice of proxy and database but also the selection of control variables plays an important role. In the event study, we solely account for the industry variable. Despite these minor caveats and actively accounting for those limitations, our research still produces meaningful insights for understanding the relationship between quantitative AI focus and qualitative AI communication. More importantly, combining our insights with the outlined limitations opens concrete inspiration for further research.

8. Outlook

Our work opens several avenues for future scholars to extend these findings. A primary opportunity lies in integrating the two methodological approaches. Currently, the impact of AI focus on long-term valuation (H1) and the impact of management stance on short-term returns (H2) are treated as distinct phenomena. It is plausible that the positive association between AI focus and firm value is conditional on the qualitative framing. That is, high AI focus might only accrete value when accompanied by a confident, optimistic management tone. Investigating whether "Positive Stance" amplifies the valuation premium of AI focus would help distinguish between substantive strategic commitment and mere "AI-washing".

Furthermore, while 10-K filings provide a standardized and regulatory-mandated dataset, they represent only one channel of corporate communication. Future inquiries could expand the dictionary-based analysis to more communication channels, such as earnings calls or press releases, where management sentiment may be less constrained and more predictive of market reactions. Examining whether the "Positive Stance" observed in a 10-K is consistent with the tone taken in subsequent earnings calls could offer a measure of narrative consistency.

Finally, the S&P 500's geographic scope limits the generalizability of the findings to the U.S. regulatory and economic environment. As AI regulation diverges globally, comparative studies could investigate whether regulatory headwinds dampen the valuation premium of AI focus in different jurisdictions. Additionally, validating the narrative AI focus measure against tangible inputs, such as AI-specific capital expenditures or human capital data from job postings, would further clarify the transmission mechanism between corporate disclosure, actual investment, and firm value.

References

Acemoglu, D. (2024). The simple macroeconomics of AI. *Economic Policy*, *40*(121), 13–58. https://doi.org/10.1093/epolic/eiae042

Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction Machines: The Simple Economics of Artificial Intelligence. In Google Books. https://books.google.at/books/about/Prediction_Machines.html?id=wJY4DwAAQBAJ&redir_esc=y

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Artificial intelligence: The ambiguous labor market impact of automating prediction. *The Journal of Economic Perspectives*, *33*(2), 31–50. https://doi.org/10.1257/jep.33.2.31

Ahmad, K., Han, J., Hutson, E., Kearney, C., & Liu, S. (2016). Media-expressed negative tone and firm-level stock returns. Journal of Corporate Finance, 37, 152–172. https://doi.org/10.1016/j.jcorpfin.2015.12.014

Alexander, I., Irwin, T., & The World Bank Group. (1996). Price caps, Rate-of-Return regulation, and the cost of capital. In *The World Bank Group*. https://documents1.worldbank.org/curated/en/289361468779073062/pdf/multi-page.pdf

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics, 151, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Baker, P. (2008). Corpus Methods in Linguistics. In Research Methods in Linguistics (pp. 167–192). https://doi.org/10.5040/9781350043466.ch-008

Bardhan, I., Krishnan, V., & Lin, S. (2013). Research Note—Business Value of Information Technology: Testing the Interaction Effect of IT and R&D on Tobin's Q. Information Systems Research, 24(4), 1147–1161. https://doi.org/10.1287/isre.2013.0481

Baruffaldi, S. et al. (2020), “Identifying and measuring developments in artificial intelligence: Making the impossible possible”, OECD Science, Technology and Industry Working Papers, No. 2020/05, OECD Publishing, Paris, https://doi.org/10.1787/5f65ff7e-en

Basso, L. F. C., De Oliveira, J. a. S., Kimura, H., & Braune, E. S. (2014). The impact of intangibles on value creation: Comparative analysis of the Gu and Lev methodology for the United States software and hardware sector. Investigaciones Europeas De Dirección Y Economía De La Empresa, 21(2), 73–83. https://doi.org/10.1016/j.iedee.2014.09.001

Apotheker, J., Beauchene, V., De Bellefonds, N., Forth, F., & Forth, M. (2025). The Widening AI Value Gap. In https://media-publications.bcg.com/The-Widening-AI-Value-Gap-October-2025.pdf. Boston Consulting Group.

Berns, J., Bick, P., Flugum, R., & Houston, R. (2022). Do changes in MD&A section tone predict investment behavior? Financial Review, 57(1), 129–153. https://doi.org/10.1111/fire.12280

Bharadwaj, A. S. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical investigation1. MIS Quarterly, 24(1), 169–196. https://doi.org/10.2307/3250983

Biber, D., Conrad, S., Reppen, R., & Precht, K. (1998). Corpus Linguistics: Investigating Language structure and use. TESOL Quarterly, 32(4), 789. https://doi.org/10.2307/3588017

Bose, I., Lui, A. K. H., & Ngai, E. W. T. (2011). The Impact of RFID adoption on the market Value of Firms: An Empirical analysis. Journal of Organizational Computing and Electronic Commerce, 21(4), 268–294. https://doi.org/10.1080/10919392.2011.614184

Brown, S. V., & Tucker, J. W. (2011). Large-Sample Evidence on Firms’ Year-over-Year MD&A Modifications. Journal of Accounting Research, 49(2), 309–346. https://doi.org/10.1111/j.1475-679x.2010.00396.x

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: information technology, organizational transformation and business performance. *The Journal of Economic Perspectives*, *14*(4), 23–48. https://doi.org/10.1257/jep.14.4.23

Brynjolfsson, E., Rock, D., & Syverson, C. (2020). The Productivity J-Curve: How intangibles complement General Purpose Technologies. American Economic Journal Macroeconomics, 13(1), 333–372. https://doi.org/10.1257/mac.20180386

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. https://doi.org/10.1093/qje/qjae044

Calvino, F., Haerle, D., Liu, S., & OECD Directorate for Science, Technology and Innovation. (2025). Is Generative AI a General-Purpose Technology? OECD ARTIFICIAL INTELLIGENCE PAPERS. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/is-generative-ai-a-general-purpose-technology_6c76e7b2/704e2d12-en.pdf

Campbell, J. Y., Lo, A. W., & MacKinlay, A. (1998). The econometrics of financial markets. In *Princeton University Press eBooks*. https://doi.org/10.2307/j.ctt7skm5

Chen, H., Noronha, G., & Singal, V. (2004). The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation. *The Journal of Finance*, *59*(4), 1901–1930. https://doi.org/10.1111/j.1540-6261.2004.00683.x

Chen, W & Srinivasan, S. (2022). Going Digital: Implications for Firm Value and Performance. Review of Accounting Studies (Forthcoming). http://dx.doi.org/10.2139/ssrn.4177947

Damioli, G., Van Roy, V., Vertesy, D., & Vivarelli, M. (2022). AI technologies and employment: micro evidence from the supply side. Applied Economics Letters, 30(6), 816–821. https://doi.org/10.1080/13504851.2021.2024129

De Bellefonds, N., Charanya, T., Franke, M. R., Apotheker, J., Forth, P., Grebe, M., Luther, A., De Laubier, R., Lukic, V., Martin, M., Nopp, C., & Sassine, J. (2024). *Where's the Value in AI?* https://media-publications.bcg.com/BCG-Wheres-the-Value-in-AI.pdf

Dehning, N., Richardson, N., & Zmud, N. (2003). The Value Relevance of Announcements of Transformational Information Technology Investments. MIS Quarterly, 27(4), 637. https://doi.org/10.2307/30036551

Del Gaudio, B. L., Megaravalli, A. V., Sampagnaro, G., & Verdoliva, V. (2020). Mandatory disclosure tone and bank risk-taking: Evidence from Europe. Economics Letters, 186, 108531. https://doi.org/10.1016/j.econlet.2019.108531

Dos Santos, B. L., Peffers, K., & Mauer, D. C. (1993). The impact of information technology investment announcements on the market value of the firm. Information Systems Research, 4(1), 1–23. https://doi.org/10.1287/isre.4.1.1

Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2023). Generative AI and firm values. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4436627

Fama, E. F., & French, K. R. (1992). The Cross-Section of expected stock returns. The Journal of Finance, 47(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2009). Management's tone change, post earnings announcement drift and accruals. Review of Accounting Studies, 15(4), 915–953. https://doi.org/10.1007/s11142-009-9111-x

Fox, J. (2015). Applied regression analysis and generalized linear models. SAGE Publications, Incorporated. ISBN: 9781452205663

Fresard, L. (2010). Financial strength and product market behavior: The real effects of Corporate cash holdings. ideas.repec.org. https://ideas.repec.org/a/bla/jfinan/v65y2010i3p1097-1122.html

Gaio, C., & Raposo, C. (2010). Earnings quality and firm valuation: international evidence*. Accounting and Finance, 51(2), 467–499. https://doi.org/10.1111/j.1467-629x.2010.00362.x

Gao, X., & Feng, H. (2023). AI-Driven Productivity Gains: Artificial Intelligence and Firm Productivity. *Sustainability*, *15*(11), 8934. https://doi.org/10.3390/su15118934

García, D., Hu, X., & Rohrer, M. (2023). The colour of finance words. Journal of Financial Economics, 147(3), 525–549. https://doi.org/10.1016/j.jfineco.2022.11.006

Goldman Sachs. (2025). How Will AI Affect the Global Workforce? In *https://www.goldmansachs.com/insights/articles/how-will-ai-affect-the-global-workforce*.

Gu, F. & Lev, B. (2016). The end of accounting and the path forward for investors and managers. https://doi.org/10.1002/9781119270041

Im, K. S., Dow, K. E., & Grover, V. (2001). Research report: A reexamination of IT investment and the market value of the firm. *Information Systems Research*, *12*(1), 103-117.

Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. Journal of Financial Economics, 110(3), 712–729. https://doi.org/10.1016/j.jfineco.2013.08.018

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305–360. https://doi.org/10.1016/0304-405x(76)90026-x

Jiang, B., & Liu, C. (2019). Managerial optimism in a competitive market. Production and Operations Management, 28(4), 833–846. https://doi.org/10.1111/poms.12952

Kang, T., Park, D., & Han, I. (2018). Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics. Telematics and Informatics, 35(2), 370–381. https://doi.org/10.1016/j.tele.2017.12.014

Korst, J., Puntoni, S., & Tambe, P. (2025). Accountable Acceleration: Gen AI Fast-Tracks Into the Enterprise. In *Knowledge at Wharton*. https://ai.wharton.upenn.edu/wp-content/uploads/2025/10/2025-Wharton-GBK-AI-Adoption-Report_Full-Report.pdf

Kothari, S. P., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. The Accounting Review, 84(5), 1639–1670. https://doi.org/10.2308/accr.2009.84.5.1639

Kothari, S., & Warner, J. B. (2007). Econometrics of Event Studies. In Elsevier eBooks (pp. 3–36). https://doi.org/10.1016/b978-0-444-53265-7.50015-9

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in One-Criterion variance analysis. Journal of the American Statistical Association, 47(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Lev, B., & Sougiannis, T. (1996). The capitalization, amortization, and value-relevance of R&D. Journal of Accounting and Economics, 21(1), 107–138. https://doi.org/10.1016/0165-4101(95)00410-6

Li, F. (2010). The information Content of Forward-Looking Statements in Corporate Filings—A naïve Bayesian Machine Learning approach. Journal of Accounting Research, 48(5), 1049–1102. https://doi.org/10.1111/j.1475-679x.2010.00382.x

Lie, E., & Lie, H. J. (2002). Multiples used to estimate corporate value. Financial Analysts Journal, 58(2), 44–54. https://doi.org/10.2469/faj.v58.n2.2522

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2013). IPO first-day returns, offer price revisions, volatility, and form S-1 language. Journal of Financial Economics, 109(2), 307–326. https://doi.org/10.1016/j.jfineco.2013.02.017

Luhn, H. P. (1958). The automatic creation of literature abstracts. IBM Journal of Research and Development, 2(2), 159–165. https://doi.org/10.1147/rd.22.0159

Lui, A. K. H., Lee, M. C. M., & Ngai, E. W. T. (2021). Impact of artificial intelligence investment on firm value. Annals of Operations Research, 308(1–2), 373–388. https://doi.org/10.1007/s10479-020-03862-8

Luoma, T. (2011). Nonparametric event study tests for testing cumulative abnormal returns. In *Osuva (University of Vaasa)*. http://osuva.uwasa.fi/handle/10024/7648

Lynch, A. W., & Mendenhall, R. R. (1997). New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index. *The Journal of Business*, *70*(3), 351–383. https://doi.org/10.1086/209722

MacKinlay, A. C. (1997). Event studies in economics and finance. ideas.repec.org. https://ideas.repec.org/a/aea/jeclit/v35y1997i1p13-39.html

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. The Journal of Finance, 60(6), 2661–2700. https://doi.org/10.1111/j.1540-6261.2005.00813.x

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., . . . Oak, S. (2025). Artificial Intelligence Index Report 2025. https://arxiv.org/abs/2504.07139

Mayew, W. J., Sethuraman, M., & Venkatachalam, M. (2015). MD&A Disclosure and the Firm's Ability to Continue as a Going Concern. The Accounting Review, 90(4), 1621–1651. https://doi.org/10.2308/accr-50983

McKinsey. (2022). What are Industry 4.0, the Fourth Industrial Revolution, and 4IR?. In https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-are-industry-4-0-the-fourth-industrial-revolution-and-4ir?

McKinsey. (2023). The economic potential of generative AI. In *https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/*.

Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, *164*, 113998. https://doi.org/10.1016/j.jbusres.2023.113998

Mishra, S., Ewing, M. T., & Cooper, H. B. (2022). Artificial intelligence focus and firm performance. Journal of the Academy of Marketing Science, 50(6), 1176–1197. https://doi.org/10.1007/s11747-022-00876-5

Mukherjee, A. N. (2022). Application of artificial intelligence: benefits and limitations for human potential and labor-intensive economy – an empirical investigation into pandemic ridden Indian industry. *Management Matters*, *19*(2), 149–166. https://doi.org/10.1108/manm-02-2022-0034

Nachar, N. (2008). The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. Tutorials in Quantitative Methods for Psychology, 4(1), 13–20. https://doi.org/10.20982/tqmp.04.1.p013

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. Journal of Financial Economics, 108(1), 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187–192. https://doi.org/10.1126/science.adh2586

OECD/BCG/INSEAD. (2025). *The Adoption of Artificial Intelligence in Firms*. OECD Publishing, Paris. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/05/the-adoption-of-artificial-intelligence-in-firms_8fab986b/f9ef33c3-en.pdf

Peters, R. H., & Taylor, L. A. (2017). Intangible capital and the investment-q relation. Journal of Financial Economics, 123(2), 251–272. https://doi.org/10.1016/j.jfineco.2016.03.011

S&P Global Inc. (2025). S&P 500 The Gauge of the U.S. Large-Cap Market. www.spglobal.com. https://www.spglobal.com/spdji/en/documents/additional-material/sp-500-brochure.pdf

Samoili, S., Lopez Cobo, M., Delipetrev, B., Martinez-Plumed, F., Gomez Gutierrez, E. and De Prato, G., AI Watch. Defining Artificial Intelligence 2.0, EUR 30873 EN, Publications Office of the European Union, Luxembourg, 2021, https://data.europa.eu/doi/10.2760/019901

Shiyyab, F. S., Alzoubi, A. B., Obidat, Q. M., & Alshurafat, H. (2023). The impact of artificial intelligence disclosure on financial performance. International Journal of Financial Studies, 11(3), 115. https://doi.org/10.3390/ijfs11030115

Singla, A., Sukharevsky, A., Yee, L., & Chui, M. (2025). The state of AI in 2025. In McKinsey. McKinsey. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/

Subramani, M., & Walden, E. (2001). The impact of E-Commerce announcements on the market value of firms. Information Systems Research, 12(2), 135–154. https://doi.org/10.1287/isre.12.2.135.9698

Sun, Y. (2010). Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases? The Accounting Review, 85(4), 1411–1440. https://doi.org/10.2308/accr.2010.85.4.1411

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Wang, C., Xin, J., Sun, F., Shi, Y., & Du, Y. (2024). The effects of manager sentiment in financial disclosure: Perspectives of operational efficiency and market reaction. Finance Research Letters, 64, 105425. https://doi.org/10.1016/j.frl.2024.10542

Yekini, L. S., Wisniewski, T. P., & Millo, Y. (2015). Market reaction to the positiveness of annual report narratives. The British Accounting Review, 48(4), 415–430. https://doi.org/10.1016/j.bar.2015.12.001

Zhu, M., Wen, S., & Song, Y. (2024). Impact of COVID-19 on jump occurrence in capital markets. Humanities and Social Sciences Communications, 11(1). https://doi.org/10.1057/s41599-024-03357-y

Appendix

Table 10A: Pairwise Welch's t-Test for Proxy Validation

Quintile A	Quintile B	p-value
1	2	0.602
1	3	0.195
1	4	0.056*
1	5	0.018**
2	3	0.071*
2	4	0.017**
2	5	0.005***
3	4	0.475
3	5	0.203
4	5	0.557

Mean and Median are displayed in basis points

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

Table 11A: Cook's Distance including Nvidia

			Range	
Mean	Median	SD	Min	Max
0.0128	0.00171	0.1141	$1.28*10^{-8}$	1.58

Table 12A: Cook's Distance excluding Nvidia

			Range	
Mean	Median	SD	Min	Max
0.0050	0.0016	0.1110	$1.28*10^{-8}$	0.08

Table 13A: Test for OLS Regression Assumptions

Durbin-Watson		Shapiro-Wilk		VIF	
Statistic	p-value	Statistic	p-value	Min	Max
2.15	0.300	0.911	0.281	1.01	1.32

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$*

Table 14A: AI Sector Dependence

Student's t		Welch's t		Mann-Whitney U	
Statistic	p-value	Statistic	p-value	Statistic	p-value
-11.5	<.001***	-6.1	<.001***	606	<.001***

**** p < 0.01, ** p < 0.05, * p < 0.10*

Table 15A: Test for Normal Distribution of CARs

	Shapiro-Wilk		Kolmogorov-Smirnov		Anderson-Darling	
	Statistic	p-value	Statistic	p-value	Statistic	p-value
CAR	0.892	<0.001***	0.097	<0.001***	36.040	<0.000***

**** p < 0.01, ** p < 0.05, * p < 0.10*

Table 16A: Testing for Heteroskedasticity

Breusch-Pagan Statistic	Df	p-value
10.601	3	0.014**

The heteroskedasticity is tested using the sentiment as an explanatory variable
**** p < 0.01, ** p < 0.05, * p < 0.10*

Table 17A: Testing for Autocorrelation

Durbin-Watson Statistic	p-value
1.910	0.036**

**** p < 0.01, ** p < 0.05, * p < 0.10*

Table 18A: Testing for Multicollinearity

	Sentiment	GICS	Year	EV
GVIF	1.294	1.305	1.120	1.146
Df	3	8	7	2
GVIF^(1/(2*Df))	1.044	1.017	1.008	1.035

Table 19A: Pairwise Mann-Whitney U Test for Identifying Industry Differences

GICS 1	GICS 2	N1	N2	p-value
Communication	Consumer Discretionary	64	206	0.157
Communication	Consumer Staples	64	184	0.640
Communication	Energy	64	80	0.461
Communication	Health Care	64	264	0.359
Communication	IT	64	256	0.371
Communication	Industrials	64	327	0.165
Communication	Materials	64	72	0.860
Communication	Real Estate	64	144	0.266
Consumer Discretionary	Consumer Staples	206	184	0.111
Consumer Discretionary	Energy	206	80	0.548
Consumer Discretionary	Health Care	206	264	0.240
Consumer Discretionary	IT	206	256	0.242
Consumer Discretionary	Industrials	206	327	0.770
Consumer Discretionary	Materials	206	72	0.094*
Consumer Discretionary	Real Estate	206	144	0.730
Consumer Staples	Energy	184	80	0.620
Consumer Staples	Health Care	184	264	0.476
Consumer Staples	IT	184	256	0.419
Consumer Staples	Industrials	184	327	0.082*
Consumer Staples	Materials	184	72	0.586
Consumer Staples	Real Estate	184	144	0.159
Energy	Health Care	80	264	0.934
Energy	IT	80	256	0.906
Energy	Industrials	80	327	0.632
Energy	Materials	80	72	0.491
Energy	Real Estate	80	144	0.784
Health Care	IT	264	256	0.900
Health Care	Industrials	264	327	0.245
Health Care	Materials	264	72	0.209
Health Care	Real Estate	264	144	0.300
IT	Materials	256	72	0.222
IT	Real Estate	256	144	0.309
Industrials	IT	327	256	0.311
Industrials	Materials	327	72	0.058*
Industrials	Real Estate	327	144	0.864
Materials	Real Estate	72	144	0.106

Mean and Median are displayed in basis points

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 20A: AI Keywords

Keyword	Abbreviation
Artificial Intelligence	AI
Artificial General Intelligence	AGI
Artificial Neural Network	
Autonomous Agent	
Chatbot	
Computer Vision	
Convolutional Neural Network	
Deep Learning	
Embedding	
Entity Recognition	
Facial Recognition	
Generative Adversarial Network	
Image Processing	
Inferencing	
Intent Classification	
Large Language Model	LLM
Machine Learning	ML
Machine Translation	
MLOps	
Model Training	
Named Entity Recognition	
Natural Language Processing	NLP
Natural Language Understanding	
Neural Network	
Object Recognition	
Pattern Recognition	
Predictive Analytics	
Recurrent Neural Network	
Reinforcement Learning	
Semantic Web	
Speech Recognition	
Supervised Learning	
Tensor Processing Unit	
Transfer Learning	
Unsupervised Learning	

We hereby confirm that this work was written independently by Thomas Kanzian (64010) and Maximilian Klotzke (63844). No sources other than those cited were used, and all passages and ideas taken from other sources are cited accordingly.