AI Disclosures as Signal or Shield

Risk, Investment, and the Informativeness of 10-K Filings

Hyeon Kwon Park

Hyungchul Choi

Master Thesis

Stockholm School of Economics

2026

AI Disclosures as Signal or Shield: Risk, Investment, and the Informativeness of 10-K Filings

Abstract

This thesis examines whether section-specific artificial intelligence (AI) disclosure in U.S. 10-K filings is associated with filing-window market reactions and future operating performance. Using 2017-2024 filings for firms in a retrospective early-2026 S&P 500 constituent snapshot, the study links SEC EDGAR text with Capital IQ accounting data, CRSP returns, and CRSP/Compustat identifiers. AI disclosure is measured separately in strategy-oriented sections, risk factors, and financial statement footnotes.

The results are mixed. Filing-window tests do not show a consistently positive market reaction to AI disclosure. By contrast, strategy-oriented AI disclosure is significantly associated with one-year-ahead operating performance, especially operating income to revenue, while risk-oriented and accounting-note AI disclosure shows weaker and less stable associations. Strict observed R&D backing does not strengthen the strategy-disclosure relation. Overall, the findings suggest that AI disclosure is most informative when interpreted by section location, timing, and outcome type. The results are conditional associations rather than causal effects, and they point to strategy-oriented AI disclosure as more closely related to future operating outcomes than to immediate filing-date market reactions.

Keywords:

AI disclosure, 10-K filings, textual analysis, event study, operating performance

Authors:

Hyeon Kwon Park (42756)

Hyungchul Choi (42821)

Tutor:

Irina Gazizova, Assistant Professor, Department of Accounting

Master Thesis

Master Program in Accounting, Valuation and Financial Management

Stockholm School of Economics

Acknowledgements

We would like to express our sincere gratitude to our supervisor, Irina Gazizova, for her valuable guidance and feedback throughout the process of writing this thesis. Her encouragement, patience, and support have been extremely helpful in shaping our work and guiding us through the thesis process.

Stockholm, May 2026

Hyeon Kwon Park Hyungchul Choi

List of Abbreviations

Acronym / term	Full wording	Use in the thesis
10-K	10-K annual report	Annual SEC filing used for text construction and filing-date event measurement.
AccountingNoteAI	Accounting-note AI disclosure	AI disclosure in financial statement footnotes; included in H1 and H2 and retained as a non-interacted H3 control.
AI	Artificial intelligence	Core disclosure topic measured in 10-K text.
Capital IQ Pro	S&P Capital IQ Pro	Source of accounting and financial statement variables.
CAR	Cumulative abnormal return	Filing-window or post-filing abnormal return outcome.
CCM	CRSP/Compustat Merged database	Identifier-linking source used to connect Compustat/issuer records to CRSP returns.
CIK	Central Index Key	SEC issuer identifier used to anchor filing retrieval and issuer-level consolidation.
CRSP	Center for Research in Security Prices	Source of daily stock returns and security identifiers.
CUSIP	Committee on Uniform Securities Identification Procedures identifier	Security identifier used in linkage and validation checks.
EDGAR	Electronic Data Gathering, Analysis, and Retrieval system	SEC filing system used for 10-K text retrieval.
FE	Fixed effects	Firm and year effects used to absorb persistent firm differences and year shocks.
LPERMNO	Linked PERMNO	CRSP-linked permanent security identifier from CCM records.
MCAR	Market-adjusted cumulative abnormal return	Main market-reaction outcome based on returns net of the market benchmark.
MD&A	Management's Discussion and Analysis	Item 7 narrative section included in the StrategyAI block.
PERMNO	Permanent security number	CRSP security identifier used for daily return measurement.
R&D	Research and development	Observed investment-backing context used only for the H3 moderation test.
RDStockTA	R&D stock scaled by total assets	Asset-scaled R&D-stock proxy used as the H3 backing moderator.
RiskAI	Risk-oriented AI disclosure	AI disclosure in Item 1A risk factors; included in H1 and H2 and retained as a non-interacted H3 control.
S&P	Standard & Poor's / S&P Dow Jones Indices	Source of the early-2026 S&P 500 constituent snapshot.
SEC	U.S. Securities and Exchange Commission	Source authority for 10-K filings.
StrategyAI	Strategy-oriented AI disclosure	AI disclosure in Item 1 business description and Item 7 MD&A; main disclosure variable for H2 and H3.
Strict weighted RDStockTA	Strict observed R&D weighted stock scaled by total assets	H3 moderator requiring current and four lagged observed R&D inputs and positive disclosure-year assets.
WRDS	Wharton Research Data Services	Platform used to access CRSP and CCM data.

Contents

1. Introduction

1.1. Motivation and Research Question

AI disclosure has become increasingly visible in corporate reporting, and recent research uses 10-K text to measure firm-level AI engagement (Ante & Saggu, 2025). At the same time, explicit AI language in annual reports remains difficult to interpret. Managers may use AI-related language to communicate strategic initiatives, but such language may also reflect investor attention, litigation risk, or competitive pressure. This ambiguity places AI disclosure within a broader disclosure theory problem. Corporate reporting can reduce information asymmetry, but disclosure choices are also shaped by proprietary costs and managerial incentives (Verrecchia, 1983; Healy & Palepu, 2001; Beyer et al., 2010).

The 10-K setting is useful because annual reports are mandatory, recurring, and legally salient, but they may not be the first channel through which investors learn about corporate AI activity. This creates a distinction between filing-date informativeness and operating informativeness: a filing-window return test asks whether the 10-K provides incremental news at the filing date, while a future operating-performance test asks whether the same disclosure contains information about later firm outcomes.

The empirical problem is also a measurement problem. AI terms do not have a stable meaning across all sections of a 10-K filing. In the Item 1 business description and Item 7 Management's Discussion and Analysis sections, AI language may describe products, services, analytics capabilities, automation systems, customer-facing technology, or internal process improvements. In Item 1A risk factors, the same language may describe cybersecurity exposure, regulatory uncertainty, model error, competitive disruption, or legal risk. In financial statement footnotes, AI language appears in a more formal accounting context rather than in narrative strategy or risk discussion. Treating these contexts as identical would weaken the economic interpretation of the disclosure measure.

AI disclosure is also difficult to interpret because underlying AI activity is heterogeneous. One firm may use AI to automate customer service, another may use it to improve internal analytics, and another may mention AI mainly when it creates cybersecurity, regulatory, or competitive risk. A single aggregate AI word count would collapse these settings into one measure. The empirical design therefore separates AI disclosure into three domains rather than treating all AI mentions as a single signal. These domains are strategy-oriented AI language (StrategyAI), risk-oriented AI language (RiskAI), and accounting-note-oriented AI language (AccountingNoteAI). The strategy domain is constructed from the Item 1 business description and Item 7 MD&A text, while Item 1A risk factors and financial statement footnotes are retained as separate disclosure domains.

The thesis asks whether AI disclosure in 10-K filings is more consistent with an informative signal, a strategic shield, or a combination of both. It does not assume in advance that one disclosure domain is the only relevant channel. Instead, the empirical

design first separates strategy-oriented, risk-oriented, and accounting-note-oriented AI language, and then tests whether each domain is associated with filing-window market reactions or later operating outcomes.

The empirical tests are organized around three questions. First, is section-specific AI disclosure associated with filing-window or post-filing abnormal returns? Second, are strategy-oriented, risk-oriented, and accounting-note-oriented AI disclosures associated with one-year-ahead operating outcomes through domain-specific coefficient patterns? Third, is the StrategyAI-performance relation moderated by strict observed R&D backing?

These three tests move from a short-window market test to a future outcome test to a narrower R&D-backed moderation test. H1 asks whether the three section-specific AI disclosure domains are associated with filing-window or post-filing abnormal returns. H2 applies section-specific design by testing whether strategy-oriented, risk-oriented, and accounting-note-oriented AI disclosures show domain-specific coefficient patterns in one-year-ahead operating-outcome regressions. This test reflects the premise that AI language in strategy, risk, and accounting-note contexts may have different economic meanings. H3 then narrows the analysis to the StrategyAI relation in H2 and asks whether that relation is stronger in the strict observed R&D subsample measured with strict weighted RDStockTA. Weak, mixed, or domain-specific H1 evidence would not invalidate H2, because the 10-K may repeat information already known to investors and because filing-date informativeness differs from operating informativeness. Similarly, a weak H3 result would not invalidate H2, because H3 tests a narrower observed R&D moderation mechanism rather than the full operating-content claim.

The signal interpretation is therefore not that every AI mention represents favorable news. Rather, it is that section-specific disclosure may contain information that is reflected either in filing-window market associations or in future operating outcomes when the language appears in economically relevant sections and survives controls. The shield interpretation is not that firms necessarily misstate AI activity. It is that broad AI language may also manage uncertainty, legal exposure, investor expectations, or competitive positioning. These interpretations are not mutually exclusive, since firms can disclose substantive AI-related activity while also using broad language to manage uncertainty around implementation and outcomes.

1.2. Contribution

This thesis makes several contributions. First, it evaluates AI disclosure through both capital-market and accounting outcome channels rather than relying only on AI mention counts. This distinction matters because a 10-K filing may contain economically meaningful language even when it does not generate a clear filing-window market reaction. Second, the thesis separates strategy-oriented AI language in Item 1 and Item 7 MD&A, risk-oriented AI language in Item 1A, and accounting-note AI language in

financial statement footnotes. This section-specific approach follows textual-analysis evidence that financial text requires domain-specific measurement and that 10-K meaning depends on reporting context, document structure, and disclosure location (Loughran & McDonald, 2011, 2016; Dyer et al., 2017; Akyol & Shekhar, 2026). Third, the thesis treats R&D evidence conservatively by restricting the H3 backing test to the observed R&D subsample and by not recoding missing R&D as zero. This treatment follows evidence that missing R&D does not necessarily indicate the absence of innovation activity (Koh & Reeb, 2015). Fourth, the thesis makes the text pipeline auditable by reporting the dictionary scope, exclusion rules, section construction, CRSP/CCM linking logic, and reproducibility documentation. Empirical results based on textual analysis are credible only when the researcher can explain how raw filings become regression variables. The appendix therefore documents the main construction choices so that the empirical measures can be checked against the filing text and sample selection rules.

The empirical scope is narrow by design. The thesis does not claim to measure internal AI use. It measures explicit AI-related disclosure in 10-K sections. It also does not claim causal effects. Instead, it tests associations using timing discipline, control variables, firm fixed effects, year fixed effects, and documented text construction. These boundaries are important because AI-related activity can occur through vendor contracts, cloud services, acquired software, internal process redesign, or other investments that may not be separately reported as R&D.

The sample is built from an early-2026 S&P 500 constituent snapshot. The results should therefore be interpreted as evidence from a retrospective large-firm panel, rather than as evidence for all public firms or for historical S&P 500 membership.

1.3. Structure of the Thesis

The thesis proceeds as follows. Section 2 reviews the relevant disclosure, textual-analysis, AI-disclosure, and intangible-investment literature. Section 3 describes the data, variables, controls, and empirical specifications. Section 4 reports the results, Section 5 discusses interpretation, credibility, limitations, and the divergence between market-reaction and operating-performance evidence, and Section 6 concludes. The appendix reports full coefficient tables, supplementary diagnostics, and reproducibility notes.

2. Literature Review

2.1. Disclosure Theory and the 10-K Information Environment

Disclosure theory implies that managers have reasons to be selective in corporate reporting. Proprietary-cost models explain why managers may withhold information when disclosure would impose competitive or other costs (Verrecchia, 1983). Broader disclosure research links reporting choices to information asymmetry, managerial incentives, litigation risk, and proprietary costs (Healy & Palepu, 2001). Information-environment research treats mandatory disclosure, voluntary disclosure, and information intermediaries as interacting sources of firm information (Beyer et al., 2010). The interaction between mandatory and voluntary disclosure is important because annual reports do not exist in isolation. Managers may disclose related information through multiple channels, and one disclosure can affect how another disclosure is interpreted (Einhorn, 2005).

Applied to AI disclosure, this means that a 10-K statement should be interpreted as one part of a broader information environment rather than as the only channel through which investors learn about corporate AI initiatives. The filing may inform, confirm, or formalize prior information while still omitting implementation details that managers have incentives not to reveal. This is particularly relevant for AI because proprietary-cost disclosure theory implies that firms may disclose some information while withholding details that could be competitively sensitive, such as implementation plans, customer targets, data infrastructure, or product timing (Verrecchia, 1983; Healy & Palepu, 2001).

This broader information environment is directly relevant to annual report AI disclosure. Investors do not necessarily wait for the 10-K to learn about a firm's AI initiatives. Information may already have appeared through earnings announcements, conference calls, investor presentations, product releases, press coverage, or earlier filings. Conference-call evidence shows that earnings-related calls can provide information to market participants and help investors process earnings news (Frankel et al., 1999; Kimbrough, 2005), while the business press can act as an information intermediary by collecting, processing, and disseminating firm information (Bushee et al., 2010). The annual report can still be useful because it consolidates information in a regulated document, but the filing event may not create a clear price reaction if investors have already processed related information through other channels.

This timing issue shapes the interpretation of H1. Event-study methods estimate abnormal returns around a defined event window to assess whether a specific event contains incremental information for investors (Brown & Warner, 1985; MacKinlay, 1997). A filing-window CAR design should therefore be interpreted as a test of incremental filing-date information rather than as a test of the total economic importance of the disclosed activity. It cannot detect information that investors already incorporated into prices

through earlier channels. For this reason, weak, mixed, or domain-specific H1 evidence would not by itself imply that AI disclosure lacks operating content. It would instead indicate that the 10-K filing does not provide consistently positive or stable incremental price-relevant news across disclosure domains at that point in time.

Capital-market interpretation follows this information-content tradition. Classic accounting studies show that earnings announcements and accounting numbers are associated with security-price or trading responses when they convey information to investors (Ball & Brown, 1968; Beaver, 1968). Although Fama et al. (1969) examine stock splits rather than 10-K filings, their evidence on price adjustment to public information is consistent with the broader event-study logic that market reactions can be measured around dated information events. Applied to this thesis, the market-reaction test asks whether AI disclosure provides incremental price-relevant information at or shortly after the 10-K filing date. It does not ask whether AI activity is economically important in a broader operating sense.

2.2. Textual Disclosure and Section-Specific Measurement

Textual-analysis research supports the view that annual report language can contain economically relevant information, but it also shows that measurement design matters. Annual report readability is associated with earnings and earnings persistence (Li, 2008), and forward-looking MD&A language can contain information about future performance (Li, 2010). Financial text also requires domain-specific dictionaries and careful interpretation because ordinary-language word lists can misclassify corporate disclosures (Loughran & McDonald, 2011, 2016). This point is central for AI disclosure because several terms that may appear AI-related can also have non-AI meanings in business, engineering, accounting, or risk contexts.

10-K textual disclosure also changes substantially over time. Dyer et al. (2017) document marked changes in 10-K disclosure, including increases in length, boilerplate, stickiness, and redundancy and decreases in specificity, readability, and the relative amount of hard information. This evidence supports caution about raw document-level AI counts, which may be affected by filing length, boilerplate, and broad changes in 10-K disclosure style over time (Dyer et al., 2017). The same logic supports the use of year fixed effects in a setting where AI salience changes over the sample period.

The readability literature also reinforces the need for context-specific measurement. Loughran and McDonald (2014) argue that conventional readability measures can be poorly specified in financial disclosure settings and propose measuring readability in ways that better reflect the communication of valuation-relevant information. Bonsall et al. (2017) similarly develop a financial-reporting readability measure designed for plain-English assessment. Although this thesis does not make readability the main variable of interest, this literature supports the broader principle that financial-text measures must be

adapted to the reporting setting rather than imported mechanically from general-language applications (Bonsall et al., 2017; Loughran & McDonald, 2014).

The section location of AI disclosure is also important. Akyol and Shekhar (2026) show that MD&A and footnotes are disclosure channels with different regulatory oversight, discretion, and reporting purposes, while also being used by managers in a complementary way. Their setting does not study AI disclosure directly, but it supports the broader section-aware logic of this thesis. Extending this logic to AI, the thesis separates strategy-oriented AI language from Item 1 and Item 7 MD&A, risk-oriented AI language from Item 1A, and accounting-note AI language from financial statement footnotes.

The economic reason for this separation is that identical AI terminology may convey different information depending on where it appears. In Item 1 and MD&A, AI language may describe products, services, analytics capabilities, automation systems, customer-facing technology, operating initiatives, or process improvements. In Item 1A, the same terminology may refer to cybersecurity exposure, regulatory uncertainty, implementation risk, competitive disruption, or legal risk. In financial statement footnotes, AI language appears in a more formal accounting context. Treating these locations as interchangeable would weaken the interpretation of the disclosure measure because the same word can carry different economic meaning across reporting sections.

This section-specific design is also consistent with the broader textual-disclosure principle that financial-text measures should be adapted to the reporting setting and disclosure channel being studied. Prior evidence shows that MD&A language can contain information about future performance, while dictionary-based financial-text measures are sensitive to word choice, context, and implementation design (Li, 2010; Loughran & McDonald, 2011, 2016). For this thesis, the relevant mechanism is not simply whether a firm mentions AI. The relevant mechanism is whether AI language appears in strategy, risk, or accounting-note sections and whether those section-specific domains show different market-reaction or operating-performance associations.

2.3. AI Disclosure and Emerging-Technology Reporting

Recent research shows that AI language in 10-K filings can be measured systematically. Ante and Saggu (2025) construct firm-level AI engagement measures from SEC 10-K filings and use disclosure-based AI measures to build data-driven AI stock indices. Their approach supports the idea that AI-related 10-K language can be treated as a measurable empirical object, while also highlighting that AI measurement requires transparent dictionary construction and contextual interpretation.

At the same time, AI disclosure should not be treated as direct proof of internal AI use. A firm can discuss AI because it has AI-related products, analytics systems, automation initiatives, or internal process improvements. It may also discuss AI because the topic

attracts investor attention, creates risk exposure, affects competitive positioning, or helps management frame the firm's strategy. This ambiguity is consistent with the broader disclosure theory, which views corporate reporting as shaped by information asymmetry, litigation risk, and proprietary costs (Verrecchia, 1983; Healy & Palepu, 2001; Beyer et al., 2010).

The AI setting is also relevant because corporate disclosures are increasingly processed by both human readers and machine readers. Cao et al. (2023) show that firms adjust disclosure when machine readers and algorithmic processing become part of the information environment. This reinforces the need for auditable and context-aware textual measurement.

The AI disclosure setting therefore requires a cautious signal-or-shield framing. AI disclosure may be informative when it reflects substantive business activity and is associated with later operating outcomes. It may operate as a shield when it manages uncertainty, legal exposure, investor expectations, or competitive positioning. This framing is consistent with disclosure theory and with the empirical caution required in textual-analysis research (Verrecchia, 1983; Healy & Palepu, 2001; Loughran & McDonald, 2011, 2016).

2.4. Intangibles, R&D, and Innovation Backing

Research on R&D and intangible investment motivates the use of observed R&D as a backing context, while also cautioning against overclaiming the link. R&D-related measures can contain value-relevant information, particularly when R&D is treated as accumulated intangible capital rather than only as a period expense (Lev & Sougiannis, 1996). R&D is also linked to information asymmetry as its value is difficult for outsiders to observe and evaluate (Aboody & Lev, 2000). R&D-intensive firms create valuation challenges for investors because the benefits of R&D are uncertain and may not be fully reflected in stock prices immediately (Chan et al., 2001). Future benefits from R&D are also more uncertain than future benefits from capital expenditure (Kothari et al., 2002). This literature is relevant because AI-related investment often has intangible characteristics, including model development, data infrastructure, software engineering, process redesign, and AI-enabled product development.

Merkley (2014) provides a useful accounting-disclosure analogue. The study examines narrative R&D disclosure in 10-K filings and argues that narrative disclosure can supplement financial statements when the underlying value of R&D investment is difficult to communicate through accounting numbers alone. Merkley also provides evidence that market participants find narrative R&D disclosure informative through analyst behavior, disclosure information content, and information asymmetry (Merkley, 2014). The lesson for this thesis is not that AI disclosure is identical to R&D disclosure. Rather, the lesson is that topic-specific narrative disclosure can be informative when the

underlying activity is intangible, strategically important, and difficult to observe directly from financial statements.

Observed R&D is therefore a useful but incomplete proxy for AI-specific backing. AI-related activity may be supported by internal R&D, but it may also occur through acquired software, licensed models, cloud infrastructure, external vendors, acquisitions, capitalized software, employee training, or operating-process redesign. For this reason, the thesis treats R&D as a conservative backing context rather than as a complete measure of AI-related investment. The H3 test therefore asks whether the strategy-disclosure performance relation is stronger when firms have observed R&D backing. It does not claim that firms without reported R&D lack AI-related activity.

The treatment of missing R&D is especially important. Koh and Reeb (2015) show that missing R&D expenditures do not necessarily indicate the absence of innovation activity, since some firms with missing R&D still file and receive patents. Missing R&D is therefore not recoded as zero. Recoding missing R&D as zero would mix firms with truly no reported R&D and firms whose innovation activity is not separately reported as R&D.

The R&D literature therefore gives H3 a narrow role. It supports a mechanism test of whether strategy-oriented AI disclosure is more informative when accompanied by observed innovation capacity. It does not make R&D a necessary condition for AI disclosure to be informative, and it does not turn the thesis into a direct test of internal AI use. The R&D-backing mechanism is applied only to StrategyAI because the proposed mechanism concerns whether strategy-oriented AI language is more credible when accompanied by observed innovation capacity. RiskAI and AccountingNoteAI remain relevant controls, but they do not represent the same strategic investment-backing mechanism.

2.5. Hypothesis Development

H1 is a section-specific market-reaction test. If investors view AI disclosure as incremental filing-date information, at least one of the section-specific AI disclosure coefficients may be statistically associated with filing-window or short post-filing abnormal returns. The expected sign may differ by disclosure domain. Strategy-oriented AI disclosure may be positive if investors interpret it as evidence of business opportunity or operating capability. It may be muted if the filing confirms information already known through other channels. It may also be negative if investors associate it with implementation costs, credibility concerns, hype, or execution uncertainty. Risk-oriented AI disclosure may be negative if it reveals uncertainty or exposure, but it may also be positive if investors value more concrete risk-specific transparency. Accounting-note AI disclosure is left as an empirical question because footnote language may reflect formal accounting detail, implementation complexity or reporting uncertainty.

The interpretation follows two related literatures. Classic accounting information-content studies show that accounting announcements can be associated with security-price or trading responses (Ball & Brown, 1968; Beaver, 1968). Event-study research provides the method for testing whether a dated event contains incremental information by measuring abnormal returns around the event window (Brown & Warner, 1985; MacKinlay, 1997). H1 therefore tests incremental filing-window and post-filing market reaction, not the full economic importance of the underlying AI activity.

H1: Strategy-oriented, risk-oriented, and accounting-note AI disclosures are associated with filing-window or short post-filing abnormal returns.

H2 predicts that the relation between AI disclosure and one-year-ahead operating outcomes shows section-specific coefficient patterns. Strategy-oriented AI disclosure is expected to have the strongest directionally favorable association with later operating outcomes, positive for profitability outcomes and negative for expense-intensity outcomes, because Item 1 and Item 7 MD&A language is more likely to describe business activity, operating initiatives, products, analytics capability, automation, and process redesign. Risk-oriented AI disclosure may have a weaker or different relation because Item 1A language may reflect uncertainty, legal exposure, cybersecurity risk, regulatory exposure, or implementation risk. Accounting-note AI disclosure is expected to be narrower because footnote language is tied more closely to accounting context and implementation details. The one-year-ahead design matches the operating mechanism more closely than a filing-window return test because AI-related operating changes may appear in future margins without producing a favorable filing-date return response.

H2: Strategy-oriented, risk-oriented, and accounting-note AI disclosures are associated with one-year-ahead operating outcomes in section-specific ways.

H3 predicts that the strategy-disclosure performance relation is stronger when strategy-oriented AI disclosure is accompanied by observed R&D backing. The citation support should be read as background rather than direct evidence of an AI-specific R&D interaction. Narrative R&D disclosure can supplement financial statements and provide useful information to market participants (Merkley, 2014). Missing R&D does not necessarily mean the absence of innovation or R&D-type activity (Koh & Reeb, 2015). The H3 test therefore treats observed R&D backing as one visible innovation channel that may make strategy-oriented AI disclosure more credible. It is an incremental mechanism test rather than a necessary condition for H2. RiskAI and AccountingNoteAI remain in the H3 regressions as non-interacted section-specific disclosure controls, but the R&D interaction is applied only to StrategyAI because the proposed mechanism concerns strategic AI language backed by observed innovation capacity.

H3: The association between strategy-oriented AI disclosure and one-year-ahead operating outcomes is stronger when the disclosure is accompanied by observed R&D backing.

The expected empirical pattern is not a simple hierarchy where H1, H2, and H3 must all point in the same direction. Corporate disclosure can occur through regulated reports, voluntary communication, and information intermediaries (Healy & Palepu, 2001; Beyer et al., 2010). Conference-call and business-press studies show that firm information can reach investors outside or around formal filing events (Frankel et al., 1999; Kimbrough, 2005; Bushee et al., 2010). Event-study logic therefore limits H1 to an incremental filing-window return test (Brown & Warner, 1985; MacKinlay, 1997). Broader economic relevance is better assessed through separate operating and longer-horizon evidence. Prior textual-disclosure research shows that annual report language can contain information about future performance or market participants' information environment (Li, 2010; Merkley, 2014). This provides the background for interpreting H2 as an operating-performance association rather than a filing-window reaction test. R&D evidence supports treating observed R&D as economically meaningful but incomplete for innovation-related interpretation (Lev & Sougiannis, 1996; Koh & Reeb, 2015). H3 is therefore interpreted as a strict observed R&D boundary test rather than a full AI-capability mechanism test.

The design is intentionally associational. The hypotheses test whether AI disclosure is statistically related to market reactions, future operating outcomes, and observable innovation backing. They do not claim that disclosure language causes stock returns or operating performance. This interpretation preserves the distinction between disclosure as an information signal and disclosure as a strategic narrative shaped by incentives, uncertainty, and proprietary costs.

3. Method

3.1. Data and Sample

The empirical design combines four data-source families. SEC EDGAR supplies the 10-K filing texts (U.S. Securities and Exchange Commission, n.d.). S&P Capital IQ Pro supplies accounting and financial statement variables (S&P Global Market Intelligence, 2026). CRSP supplies daily return data and permanent security identifiers (Center for Research in Security Prices, 2026). WRDS-hosted CRSP/Compustat Merged records support the CRSP-Compustat linkage step (Wharton Research Data Services, 2026).

The local analysis panel begins from the early-2026 large-cap constituent snapshot, retains fiscal years 2017-2024 with usable 10-K AI-related text and matched accounting variables. Fiscal-year 2025 accounting data are used only to construct one-year-ahead outcomes for fiscal-year 2024 disclosure observations when available. A single 10-K/A observation is excluded from the main frame so that amended filing text is not mixed with the original annual filing corpus and to keep the disclosure measure anchored to the original annual filing event. This leaves 3,510 firm-year observations in the main firm-year analysis frame before outcome-specific filters. The 3,510-observation text-accounting frame reflects the intersection of EDGAR section recovery, Capital IQ accounting availability, identifier harmonization, and the exclusion of the single 10-K/A observation.

The raw constituent snapshot was downloaded from the S&P Dow Jones Indices / S&P Global S&P 500 index page (S&P Dow Jones Indices, 2026) and saved locally on March 1, 2026. The file contains 503 constituent rows. Because several issuers appear through multiple share classes, the raw file was collapsed from 503 rows to 500 unique CIK-level issuers before the CIK, ticker, filing, accounting, and CRSP linkage steps. This snapshot defines the retrospective large-cap issuer universe and determines which firms are eligible for inclusion in the sample. It is not used to measure outcomes.

The early-2026 constituent snapshot defines a retrospective large-cap issuer universe rather than historical S&P 500 membership during 2017-2024. The empirical claims are therefore limited to firms that were present in the early-2026 large-cap snapshot that also had usable EDGAR filing text, matched Capital IQ accounting data, and valid CRSP/CCM return links where required. Firms that exited the index universe before the snapshot date, firms without recoverable filing sections, and firms without required accounting or return data are not represented in the main estimation frame. This sample boundary is reported as a limitation rather than treated as evidence for historical S&P 500 membership as a whole.

The main sample retains financial firms, insurers, and real-estate firms when they satisfy the text-accounting and return-link requirements. This choice follows the retrospective large-cap issuer-universe design rather than an industry-restricted operating sample

design. Because revenue-scaled operating ratios may have different economic meanings in financial firms, insurers, and REITs, the H2 and H3 findings are interpreted as large-cap cross-sector associations rather than sector-specific production-function estimates.

The event-study sample is constructed separately at the filing-event level. Each filing event is anchored to the EDGAR CIK and filing date from the local EDGAR manifest, then linked through WRDS CCM records to CRSP LPERMNO/PERMNO for daily return measurement (Wharton Research Data Services, 2026; Center for Research in Security Prices, 2026). The implemented event date is the EDGAR filing date mapped to the first CRSP trading day on or after that date. Intraday SEC acceptance timestamps are not used, so the H1 estimates are interpreted as filing-date window associations rather than precise intraday announcement effects. The audited implementation retains ticker, CUSIP, CCM link type, CCM link date, and event-date-gap checks before calculating CRSP filing-window returns. This procedure is stricter than ticker-only matching because it verifies the issuer-security link and the alignment between the SEC filing event and the traded security. The event-study design follows the standard requirement that the event date, event window, abnormal-return measure, and return benchmark be defined before estimation (Brown & Warner, 1985; MacKinlay, 1997).

The two-sample structure is necessary because the thesis combines a filing-event market-reaction design with a fiscal-year operating-performance design. H1 event-study sample requires an observed filing date, a return window, and a CRSP-linked security identifier. H2 and H3 operating-performance samples instead require fiscal-year accounting variables and one-year-ahead outcomes. H3 further restricts the operating-performance sample to firm-years with strict observed R&D coverage required for the StrategyAI interaction test. A firm-year can therefore be useful for H2 or H3 even when it does not have a clean return-window observation, and a filing event can remain useful for H1 even when later accounting outcomes are missing. The sample construction keeps each test aligned with its own unit of observation rather than forcing the market and accounting designs into one merged sample.

Because the disclosure panel covers fiscal years 2017 through 2024, the lead-outcome design requires accounting outcomes through fiscal year 2025 for fiscal-year 2024 disclosure observations. Fiscal-year 2024 firm-years are retained in the H2 and H3 accounting-performance regressions only when the corresponding fiscal-year 2025 outcome and required fiscal-year 2024 controls are available. Observations without the necessary t+1 accounting outcome remain available for H1 when they have a valid filing event and CRSP return window, but they are not used in the lead-outcome operating-performance regressions. This treatment preserves the timing distinction between disclosure-year explanatory variables and one-year-ahead operating outcomes. The interpretation of narrative disclosure as potentially informative is consistent with disclosure and R&D narrative-disclosure research (Healy & Palepu, 2001; Merkley, 2014).

The main firm-year analyses use the 2017-2024 unbalanced panel after excluding the single 10-K/A observation.

A balanced text-coverage frame is retained only as a pipeline diagnostic and is documented in Appendix B.1-B.4. It is not used for the main H1-H3 regressions.

The sample period from 2017 through 2024 has a substantive role because it spans both pre-ChatGPT years and the subsequent rise in investor and corporate attention to generative AI. Prior 10-K textual-disclosure research shows that filing language changes over time (Dyer et al., 2017). Recent AI-disclosure research treats AI-related 10-K language as a measurable corporate disclosure object (Ante & Saggu, 2025). The regressions therefore include the relevant year fixed effects, disclosure-fiscal-year fixed effects for H1 and fiscal-year fixed effects for H2 and H3, to absorb common reporting-period, disclosure-cycle, and AI-salience shocks. For H1, broad filing-window market movement is addressed through the market-adjusted CAR construction. The AI coefficients are interpreted as conditional within-firm, year-adjusted associations rather than as a simple time trend in AI language.

The text corpus is built around four 10-K regions with different reporting roles: Item 1 business description, Item 1A risk factors, MD&A, and financial statement footnotes. Consistent with those roles, the empirical blocks combine Item 1 and MD&A into strategy-oriented disclosure, retain Item 1A as risk-oriented disclosure, and retain footnotes as accounting-note disclosure. Section 3.2 describes the section-level measurement logic in detail.

Table 1 summarizes the sample construction and keeps the observation-count path visible before the variable definitions are introduced. Cleaned firm-panel identifiers are used for fixed effects, clustering, and table reporting after CIK, ticker, share-class, and ticker-history harmonization. CIK anchors EDGAR filing retrieval, while CRSP PERMNO/LPERMNO anchors return measurement. The difference between eligible issuer-level firms and cleaned firm-panel identifiers reflects identifier cleaning and harmonization after share-class and ticker-history checks.

Table 1. Sample construction and analysis frame

Step	Rule	Result
Issuer universe	Collapse the early-2026 large-cap constituent snapshot to issuer level using CIK and ticker checks.	500 unique issuer-level firms
Accounting panel	Merge S&P Capital IQ Pro annual accounting data for disclosure years 2017-2024, with fiscal-year 2025 accounting observations used only to construct t+1 outcomes where available.	447 eligible issuer-level firms; 446 cleaned firm-panel identifiers in the cleaned accounting panel
Text match	Retain firm-years with usable SEC EDGAR 10-K section text and exclude the one 10-K/A observation.	3,510 main firm-year observations across 439 cleaned firm-panel identifiers
H1 event frame	Retain filing events from the main text-accounting frame with EDGAR filing dates, valid CCM/CRSP security links, return-window availability, and required regression controls.	3,462 candidate filing events; 3,377 matched event-window return observations; up to 3,376 H1 regression observations
H2 lead-outcome frame	Require fiscal-year t disclosure variables, fiscal-year t controls, and non-missing fiscal-year t+1 operating outcome.	Up to 3,497 H2 regression observations across up to 439 cleaned firm-panel identifiers
H3 strict R&D	Require observed current and lagged R&D inputs, positive disclosure-year assets, non-missing lead outcome, and required controls. Do not recode missing R&D as zero.	1,367 strict weighted RDStockTA coverage observations; 1,362 H3 regression observations after lead-outcome and control filters

Note: Counts are reported at the issuer, cleaned firm-panel identifier, firm-year, and filing-event levels as relevant to each construction step. H1 uses filing-event observations with CRSP return-window availability, while H2 and H3 use firm-year observations with fiscal-year accounting outcomes. H2 N varies by outcome because lead-outcome and SG&A availability differ across firm-years. The difference between matched event-window observations and H1 regression observations reflects the additional requirement for non-missing disclosure variables, controls, fixed-effect identifiers, and cluster identifiers. Cleaned firm-panel identifiers are used for fixed effects, clustering, and table reporting after CIK, ticker, share-class, and ticker-history harmonization. CIK anchors EDGAR filing retrieval, while CRSP PERMNO/LPERMNO anchors event-return matching. The difference between eligible issuer-level firms and cleaned firm-panel identifiers reflects identifier cleaning and harmonization after share-class and ticker-history checks.

3.2. Text Construction and AI Disclosure Measurement

3.2.1. 10-K Section Extraction

The textual disclosure measures are constructed from section-level 10-K text. The extraction design focuses on four filing regions with distinct reporting functions. Item 1 describes the business, including products, services, markets, competition, regulation, operating costs, and how the firm operates. Item 1A reports significant risk factors.

MD&A provides management's discussion of operating results, liquidity, capital resources, and known trends or uncertainties. The financial statement notes explain information presented in the audited financial statements. These section roles motivate the thesis's distinction between strategy-oriented, risk-oriented, and accounting-note AI disclosure.

The section-level design follows the broader textual-disclosure literature, which shows that 10-K language is not homogeneous across time, topics, or reporting locations. Dyer et al. (2017) document substantial changes in 10-K textual disclosure over time, including increases in length, boilerplate, and redundancy. Prior disclosure research also treats MD&A and footnotes as distinct reporting locations with different roles in communicating managerial discussion and accounting-detail information. This thesis therefore does not treat the 10-K as one undifferentiated document. It recovers and measures AI language separately across economically meaningful filing regions.

Section recovery combines automated parsing with audit checks. Annual reports differ in heading style, ordering, HTML structure, table placement, and exhibit material. Section extraction is therefore treated as a measurement step rather than as a purely mechanical download step, because 10-K headings, HTML structure, and table of contents references vary across issuers and years. The pipeline searches for canonical item headings, excludes table-of-contents matches, applies section-cleaning rules, and formats the extracted text into sentence-level files. Filing manifests, written-section counts, error logs, test outputs, and validation files are retained to support reproducibility. The extraction procedure is not assumed to be error-free; the validation files are used to document recoverable cases, remaining failures, and known parsing limitations rather than to claim perfect section identification.

3.2.2. AI Context Identification

AI disclosure is measured with a section-aware dictionary system. The default dictionary contains conservative explicit-AI and AI-adjacent terms, while broader automation terms are excluded from the main measure to reduce false-positive risk. The dictionary is intentionally conservative: the objective is not to measure every technology-adjacent activity or broad automation discussion, but to identify explicit AI-related language that can plausibly be interpreted as a deliberate disclosure choice. This precision-oriented design follows evidence that generic dictionaries can misclassify corporate and financial text when word meanings differ from ordinary-language usage (Loughran & McDonald, 2011).

The conservative dictionary is also consistent with broader textual-analysis guidance that word lists should be evaluated in relation to the research question and the language environment in which they are applied (Loughran & McDonald, 2016). AI language in annual reports can refer to core AI technologies, machine-learning systems, applied business uses, data-driven products, and AI-related governance or risk language. The

resulting measures should therefore be interpreted as explicit AI disclosure intensity rather than as complete AI-use intensity.

The main AI measure uses 27 default include patterns covering core AI terms, model-family terms, applied-business AI phrases, and governance or risk terms, where a pattern may be a single term, acronym rule, or multiword expression. Excluded broad automation terms include generic references to automation, digitization, analytics, technology, software, and algorithms when they do not appear in an explicit AI context. The acronym "AI" is counted only under explicit boundary and context rules described in Appendix C.2 and Appendix C.3, so ordinary character strings or non-AI abbreviations are not counted as AI disclosure. Appendix C.1-C.4 report the dictionary scope, exclusion rules, matching treatment, accepted/rejected examples, and manual validation protocol.

3.2.3. Construction of StrategyAI, RiskAI, and AccountingNoteAI

The thesis constructs three section-level AI disclosure blocks. The first block, StrategyAI, combines AI language from Item 1 and MD&A. Item 1 contributes business-model and product-market context, while MD&A contributes management's interpretation of operations, liquidity, capital resources, and known trends. StrategyAI is therefore designed to capture AI language in sections where firms discuss operations, products, business models, management interpretation, and known trends; it is not designed to measure internal AI use directly. The second block, RiskAI, captures AI language in Item 1A risk factors. This block is kept separate because risk-factor disclosure identifies significant risks rather than ordinary business description language. The third block, AccountingNoteAI, reported in tables as Accounting-note AI, captures AI language in financial statement footnotes. This block is retained separately because footnotes are part of the financial statement reporting environment and provide accounting-detail context distinct from narrative business, MD&A, and risk-factor disclosure (U.S. Securities and Exchange Commission, n.d.).

For each block, AI disclosure is measured using both word-intensity and sentence-intensity ratios. For the word-intensity specification, the numerator is the count of accepted matched word tokens within the relevant section block after exclusion rules are applied. Multiword expressions are counted by the word tokens covered by the implemented dictionary match, so "AI" contributes covered word tokens rather than an inferred one-phrase count. Overlapping matches are resolved through token coverage before ratio construction, so the same word token is not counted twice. The denominator is the cleaned word count for the same section block. For the sentence-intensity specification, the numerator is the number of section-block sentences containing at least one accepted AI dictionary match after exclusions, and the denominator is the cleaned sentence count for the same section block. Ratios are constructed only when the relevant section-block denominator is positive after section recovery and cleaning. Unrecovered or zero-length section blocks are not treated as zero AI disclosure unless the relevant block is otherwise confirmed as usable recovered text with no accepted AI matches.

The word-ratio and sentence-ratio specifications are treated as parallel measurement specifications. Agreement across word and sentence specifications is interpreted as stronger measurement evidence than a result that appears under only one scaling choice. When support is concentrated in only one specification, the thesis reports this limitation rather than treating the alternative specification as a failed robustness check. The sentence-intensity measure is less sensitive to repeated terminology within the same sentence, while the word-intensity measure captures repeated explicit emphasis. The two measures therefore differ by design rather than by error alone. This approach is appropriate because explicit AI disclosure is sparse, and raw disclosure ratios are small and unevenly distributed across firms and years.

Before the regressions are estimated, the relevant continuous disclosure variables are standardized within the analysis frame used for each test family before outcome-specific filters are applied. For H1, AI disclosure variables are standardized once within the broad H1 filing-event estimation frame before window-specific return availability filters are applied. As a result, differences in H1 coefficient magnitudes across event windows are not driven by re-standardizing the AI variables within each window-specific subsample. For H2, word- and sentence-intensity AI variables are standardized within the 2017-2024 10-K/A-excluded text-accounting frame before lead-outcome filters are applied. For H3, StrategyAI, RiskAI, AccountingNoteAI, and strict weighted RDStockTA are standardized within the strict observed R&D coverage frame before regression-specific outcome and control filters are applied. The interaction term is then constructed by multiplying standardized StrategyAI by standardized strict weighted RDStockTA. RiskAI and AccountingNoteAI enter H3 as non-interacted section-specific disclosure controls. H3 coefficients are therefore interpreted within the strict observed R&D subsample and are not directly comparable to the broader H2 coefficients. A coefficient therefore describes the association with a one-standard-deviation change in the relevant constructed variable rather than with an arbitrary raw ratio unit.

3.2.4. Measurement Safeguards

The measurement design addresses three validation concerns. First, it reduces document-length bias by scaling AI matches by section-level denominators. A long filing or long section can mechanically contain more AI terms even when AI is not central to the firm's disclosure. Scaling by section words and section sentences reduces this mechanical length effect and improves comparability across firms and years.

Second, the design preserves section context. Treating the full filing as one document would mix business strategy, risk disclosure, MD&A discussion, and accounting-note language. This would make it difficult to interpret whether AI language is presented as a business opportunity, a risk exposure, or an accounting-note issue. The section-aware construction keeps the economic role of each disclosure variable visible in the regression design.

Third, the design avoids post hoc section selection. The section blocks are defined before observing the hypothesis-test results. The main H1 and H2 specifications estimate the three AI disclosure channels together. The H3 test then focuses on the pre-specified StrategyAI interaction with observed R&D backing while retaining RiskAI and AccountingNoteAI as section-specific disclosure controls. The empirical design therefore does not select the reporting section ex post based on which AI coefficient is most favorable.

Appendix B.1-B.4 report the EDGAR download, section recovery, dictionary-count merge, and CRSP/CCM event-link checks. Appendix C.1-C.4 report the AI dictionary scope, exclusion rules, matching treatment, and validation protocol. This thesis treats explicit AI language as a measurable disclosure object, while keeping section context visible in order to reduce generic document-level interpretation.

3.3. Outcome, Investment, and Control Variables

3.3.1. Market-Reaction Variable

H1 uses two return-window outcomes around the 10-K filing date. The first is arithmetic cumulative raw return, defined as the sum of CRSP firm daily returns over event window [a,b]. The second is market-adjusted cumulative abnormal return, defined as the sum of firm daily returns less the CRSP value-weighted market return over the same window. In compact notation, RawRet[a,b] equals the arithmetic sum of firm daily CRSP returns over the event window, while MCAR[a,b] equals the arithmetic sum of firm daily returns minus the CRSP value-weighted market return over the same window. The thesis uses CAR only for the market-adjusted abnormal-return specification and does not label raw cumulative returns as CAR (Brown & Warner, 1985; MacKinlay, 1997).

The purpose of these market-reaction variables is to test whether the market reaction to the 10-K filing differs with the intensity and section location of AI disclosure. A positive StrategyAI coefficient would be consistent with investors treating strategy-oriented AI language as favorable incremental information in the filing window, while a negative StrategyAI coefficient would be consistent with concerns about implementation costs, credibility, or execution uncertainty. RiskAI and AccountingNoteAI are interpreted separately because their economic meanings may differ from strategy-oriented language. H1 is evaluated as a section-specific association test. Evidence can therefore be consistent with the section-specific design even when the signs differ across StrategyAI, RiskAI, and AccountingNoteAI.

3.3.2. Operating-Performance Variables

The H2 and H3 accounting outcomes are one-year-ahead operating income to revenue, net income to revenue, SG&A expense to revenue, and other operating expenses to revenue. These outcomes are measured in fiscal year t+1, while AI disclosure and control variables are measured in fiscal year t. Operating-performance ratios are constructed only

when the relevant numerator is observed and revenue is positive and non-missing. Observations with missing or non-positive revenue are excluded from the corresponding outcome regression before winsorization. This timing separates disclosure-year explanatory variables from future operating realizations and supports the interpretation of H2 as a lead-outcome informativeness test.

The main profitability outcomes are operating income divided by revenue and net income divided by revenue. Operating income to revenue captures operating profitability before below-the-line financing and tax effects. Net income to revenue captures bottom-line profitability after interest, taxes, and non-operating items. Using both measures is useful because AI disclosure may be linked to operating performance while still being attenuated in net income by financing structure, tax effects, or non-operating shocks.

The cost-intensity outcomes are SG&A expense divided by revenue and other operating expenses divided by revenue. These measures are included because AI-related operating changes may appear first in expense structure rather than in bottom-line profitability. For example, AI-related operating changes may increase short-term implementation, personnel, consulting, software, and integration costs, while potential efficiency gains may appear only later. Cost-intensity outcomes therefore provide a separate channel for evaluating whether AI disclosure is associated with future cost structure.

Other operating expenses to revenue uses Capital IQ Pro's "Other Operating Exp., Total" line scaled by total revenue. This measure is broader than SG&A/revenue and should not be interpreted as a cost category that is mechanically separate from SG&A. It is retained as a broader operating-expense intensity measure, while SG&A/revenue captures the narrower selling, general, and administrative expense line. The interpretation therefore focuses on broad operating-expense intensity.

All margin outcomes are scaled by revenue. This scaling choice makes the dependent variables more comparable across firms of different sizes and keeps the analysis focused on profitability and cost intensity rather than raw dollar magnitude. A dollar increase in operating income or expenses has different economic meaning for a small firm and a very large firm. Revenue scaling reduces this comparability problem and is consistent with the thesis objective of testing future operating performance rather than firm scale.

R&D-based variables are not treated as broad H2 outcomes. They are also not included as broad H2 controls. Instead, R&D enters only through the strict observed R&D backing design in H3. This separation is important because H2 asks whether AI disclosure is associated with future operating outcomes, while H3 asks whether the StrategyAI relation is stronger when observed R&D capacity is present.

3.3.3. R&D-Backing Variables

Strict weighted RDStockTA is a conservative observed innovation-capacity proxy. It is not an AI-specific investment account and does not capture software subscriptions, cloud-

computing capacity, vendor AI systems, acquired AI platforms, proprietary data assets, business process redesign, or AI-specific human capital. H3 therefore tests whether the StrategyAI relation is stronger when broad observed R&D backing is present, not whether all AI activity is R&D-backed. H3 uses R&D only as a strict observed backing proxy, not as a general performance adjustment. This interpretation follows R&D-capitalization research that treats accumulated R&D as an economically meaningful form of knowledge capital (Lev & Sougiannis, 1996; Chan et al., 2001) and acknowledges that R&D benefits are uncertain and difficult to assign precisely across periods (Kothari et al., 2002).

Current R&D/assets is retained as an observed R&D coverage and diagnostic measure, not as the main H3 interaction proxy. It is defined as current R&D expenditure divided by total assets:

$$CurrentRDTA_{i,t} = \frac{RD_{i,t}}{Assets_{i,t}} \quad (1)$$

This variable is defined only when current R&D is observed and total assets are positive. Missing R&D is not recoded as zero. This strict treatment follows evidence that missing R&D should not automatically be interpreted as zero innovation or zero R&D-type activity (Koh & Reeb, 2015).

Strict weighted RDStockTA is first constructed as a raw asset-scaled R&D stock measure. The numerator equals observed current R&D expense plus four observed lagged R&D expense values, weighted 1.0, 0.8, 0.6, 0.4, and 0.2. The weights impose a simple five-year linear decay structure: current R&D receives the highest weight, while older R&D is retained but gradually discounted. This approach follows the idea that R&D may contribute to knowledge capital over several years, while avoiding the stronger assumption that older R&D has the same relevance as current R&D (Lev & Sougiannis, 1996; Chan et al., 2001). The weighted R&D numerator is divided by total assets measured in the disclosure fiscal year (t).

Although the disclosure panel begins in fiscal year 2017, the accounting extraction retains pre-2017 R&D observations where available for the sole purpose of constructing lagged R&D stock inputs. The measure is defined only when current R&D, all four lagged R&D values, and positive disclosure-year total assets are observed.

$$RDStockTA_{i,t}^{strict} = \frac{RD_{i,t} + 0.8RD_{i,t-1} + 0.6RD_{i,t-2} + 0.4RD_{i,t-3} + 0.2RD_{i,t-4}}{Assets_{i,t}} \quad (2)$$

The declining weights are used as a transparent approximation of accumulated R&D support, not as an estimated model of how R&D knowledge actually depreciates over time. The thesis does not estimate an optimal R&D depreciation rate and does not interpret the constructed stock as a precise measure of knowledge capital.

In the H3 regression, the constructed raw RDStockTA measure and the H3 disclosure variables are standardized within the strict observed R&D coverage frame before

regression-specific outcome and control filters are applied. The interaction term therefore multiplies standardized StrategyAI by standardized strict weighted RDStockTA. The interaction coefficient should therefore be interpreted as the change in the StrategyAI association with future outcomes for a one-standard-deviation increase in observed strict R&D backing, not for a one-unit increase in the raw asset-scaled R&D stock.

Strict weighted RDStockTA is a moderator, not an accounting adjustment. It does not revise operating income, net income, SG&A, or other operating expenses. It enters the H3 design only to test whether the relation between StrategyAI and future outcomes is conditional on observed R&D support. This keeps the future operating outcomes conceptually separate from the investment-backing mechanism.

3.3.4. Control Variables

The control set is kept small because the main specifications include firm fixed effects and year fixed effects. The thesis includes size, leverage, cash ratio, and sales growth because these variables adjust for time-varying scale, financing, liquidity, and growth differences rather than attempting to saturate the model with variables that may themselves be affected by AI-related strategy. The size, leverage, and cash-ratio controls follow standard finance research on firm scale, capital structure, and liquidity (Fama & French, 1992; Rajan & Zingales, 1995; Opler et al., 1999). Sales growth is included to absorb contemporaneous operating momentum and growth conditions. The control set is not intended to block every possible economic channel between AI disclosure and future performance. It is designed to adjust for major time-varying fundamentals while avoiding controls that may themselves be intermediate outcomes of AI-related strategy. Size controls for firm scale, disclosure capacity, and persistent differences in business complexity. Leverage controls for capital structure and financial risk. Cash ratio controls for liquidity and internal financing capacity. Sales growth controls for operating momentum and growth opportunities.

The controls are measured in fiscal year t. The accounting outcomes are measured in fiscal year t+1. This timing avoids using future controls and keeps the explanatory variables temporally prior to the operating-performance outcomes. The H1 event-study regressions use the same core firm fundamentals, so the market-reaction and operating-performance tests are conditioned on comparable firm characteristics.

The control vector includes size, leverage, cash ratio, and sales growth. Firm and year fixed effects, as well as clustered standard errors, are not defined as ordinary controls in this section because they are features of the empirical specification.

Table 2 summarizes the main variables used in the empirical tests.

Table 2. Variable definitions and measurement timing

Family	Variable	Definition	Use
AI disclosure	Strategy AI	AI dictionary matches in Item 1 and MD&A measured in fiscal year t filing text, scaled by the corresponding strategy-block word or sentence denominator; reported separately as word-intensity and sentence-intensity ratios.	H1/H2 explanatory variable; H3 main disclosure variable and RDStockTA interaction component
AI disclosure	Risk AI	AI dictionary matches in Item 1A risk-factor text measured in fiscal year t filing text, scaled by the corresponding risk-factor word or sentence denominator; reported separately as word-intensity and sentence-intensity ratios.	H1/H2 explanatory variable; H3 non-interacted section-specific disclosure control
AI disclosure	AccountingNote AI	AI dictionary matches in financial statement footnote text measured in fiscal year t filing text, scaled by the corresponding footnote word or sentence denominator; reported separately as word-intensity and sentence-intensity ratios.	H1/H2 explanatory variable; H3 non-interacted section-specific disclosure control
Outcomes	OI/revenue t+1	Operating income divided by positive revenue in fiscal year t+1; defined only when numerator and revenue are observed.	Main H2/H3 margin outcome
Outcomes	NI/revenue t+1	Net income divided by positive revenue in fiscal year t+1; defined only when numerator and revenue are observed.	Main H2/H3 profitability outcome
Outcomes	SG&A/revenue t+1	Selling, general, and administrative expense divided by positive revenue in fiscal year t+1; defined only when numerator and revenue are observed.	Operating cost outcome
Outcomes	Other opex/revenue t+1	Capital IQ Pro “Other Operating Exp., Total” divided by positive revenue in fiscal year t+1; broader than SG&A/revenue and not mechanically separate from SG&A.	Broader operating-expense intensity outcome
R&D backing	Current R&D/assets	Observed current R&D expense in fiscal year t divided by positive disclosure-year total assets; defined only when R&D and positive total assets are observed.	R&D coverage and diagnostic variable, not the main H3 interaction proxy
R&D backing	Strict weighted RDStockTA	Observed current and four lagged R&D expense measured through fiscal year t, weighted 1.0, 0.8, 0.6, 0.4, and 0.2, then scaled by positive disclosure-year total assets; defined only when all R&D expense inputs and positive assets are observed; standardized from observed strict-R&D coverage for H3 regressions.	Main strict H3 backing proxy
Controls	Size	Natural log of total assets measured in fiscal year t.	Scale and disclosure-capacity control

Family	Variable	Definition	Use
Controls	Leverage	Total debt divided by total assets, measured in fiscal year t.	Capital-structure and risk control
Controls	Cash ratio	Cash and equivalents divided by total assets, measured in fiscal year t.	Liquidity and financial-flexibility control
Controls	Sales growth	Current revenue minus lagged revenue, divided by lagged revenue when lagged revenue is available and nonzero; measured in fiscal year t.	Operating-momentum control

Table 3 reports the control variables, fixed effects, and inference structure used in the empirical design.

Table 3. Controls, fixed effects, and inference structure

Design element	Expected role
Size	Captures firm scale and broad disclosure-capacity differences.
Leverage	Captures financing structure, financial risk, and balance-sheet constraints.
Cash ratio	Captures liquidity and internal financing capacity.
Sales growth	Captures operating momentum, demand conditions, and growth opportunities that may affect future margins independently of disclosure.
Cleaned firm-panel identifier firm fixed effects	Absorb time-invariant business model, reporting culture, sector position, and persistent disclosure style at the cleaned firm-panel identifier level.
H1 disclosure-fiscal-year fixed effects	Absorb common disclosure-year reporting conditions and AI-salience shocks in the filing-event design; broad market movement is addressed through market-adjusted CAR.
H2/H3 fiscal-year fixed effects	Absorb common fiscal-year shocks in accounting outcome regressions.
Clustered standard errors	Main inference uses one-way clustering at the cleaned firm-panel identifier level, allowing residual dependence within firms over time.

Accounting ratios are first constructed using the stated denominator rules. Revenue-scaled outcomes require positive revenue, and asset-scaled R&D variables require positive total assets. The resulting continuous accounting ratios, including operating margins, expense ratios, leverage, cash ratio, sales growth, current R&D/assets, and raw strict weighted RDStockTA, are then winsorized at fiscal-year-specific 1st and 99th percentiles before lead-outcome alignment and regression estimation. AI disclosure ratios are not winsorized before standardization because zeros and upper-tail observations are part of the sparse disclosure distribution. Return outcomes are not winsorized in the main H1 specifications. Winsorization is applied as a preprocessing rule to reduce the influence of extreme denominator-driven ratio values, unusual accounting events, and capital-

structure outliers. However, it is not treated as a substitute for robustness testing, because winsorization can alter valid observations and may not fully address influential-observation problems (Leone et al., 2019).

The variables in this section create the measurement base for the empirical specifications in Section 3.4. H1 uses raw cumulative returns and market-adjusted CARs because the hypothesis concerns investor reaction to the 10-K filing. H2 uses one-year-ahead accounting outcomes because the hypothesis concerns future operating informativeness. H3 uses strict weighted RDStockTA because the hypothesis asks whether the StrategyAI relation is stronger when disclosure is supported by observed innovation capacity.

3.4. Empirical Design

This section describes the empirical design used to test whether section-specific AI disclosure is associated with stock-market reactions around 10-K filings and with future operating performance. The empirical design separates the analysis into three tests. H1 examines filing-period and post-filing raw cumulative returns and market-adjusted abnormal returns. H2 examines whether AI disclosure in year t is associated with accounting outcomes in year t+1. H3 examines whether the association between strategy-oriented AI disclosure and future performance is stronger when the disclosure is supported by observable R&D investment.

The specifications estimate conditional associations, not causal treatment effects. This limitation is structural because firm-level AI implementation quality, product-market positioning, managerial ability, and prior external information are not fully observable. Prior textual-disclosure research links financial-text features to market reactions and future performance, while related work shows that 10-K language changes materially over time (Li, 2008; Loughran & McDonald, 2011; Dyer et al., 2017), but the empirical design does not treat AI disclosure as a randomized treatment.

H1 and H2 are estimated separately using the word-intensity and sentence-intensity versions of StrategyAI, RiskAI, and AccountingNoteAI. H3 uses the same word-intensity and sentence-intensity disclosure specifications, but only StrategyAI interacts with strict weighted RDStockTA. RiskAI and AccountingNoteAI remain as non-interacted section-specific disclosure controls.

The empirical design also follows the thesis's section-level logic. MD&A and footnotes are not treated as interchangeable disclosure locations. MD&A provides a more narrative and forward-looking discussion of operations and business conditions, while footnotes are more directly connected to accounting recognition, measurement, and audited financial statement detail. This section-level separation is consistent with setting-aware financial-text measurement and with evidence that MD&A and footnotes operate as distinct disclosure channels (Loughran & McDonald, 2011, 2016; Dyer et al., 2017; Akyol & Shekhar, 2026). The thesis therefore separates strategy, risk, and accounting-

note AI disclosure rather than collapsing all AI references into one generic disclosure count.

3.4.1. H1 Event-Study Specification

The primary short-window H1 tests are [-1,+1], [0,+5], and [0,+20]. The [-1,+1] window is the narrow filing-period test, while [0,+5] and [0,+20] capture short post-filing adjustment. The [0,+60] and [0,+126] windows are reported only as diagnostic post-filing cumulative-return associations and are not interpreted as clean announcement effects. They diagnose whether the short-window pattern reverses or persists over longer post-filing horizons. The implemented H1 event date is the EDGAR manifest filing date mapped to the first CRSP trading day on or after that date. The design does not use intraday SEC acceptance timestamps and does not make intraday claims about investor reaction. The [-1,+1] and [0,+5] windows partly mitigate timing noise from filings submitted outside regular trading hours, so H1 estimates are interpreted as filing-date window associations. The arithmetic cumulative-return design is used for consistency with the market-adjusted CAR construction across windows. Compounded buy-and-hold returns are not used as the main return specification.

The main abnormal-return specification is a market-adjusted CAR rather than an estimation-window market-model CAR. The benchmark return is the CRSP value-weighted market return. No firm-specific estimation-window market model is used in the main reported CAR specification. This choice keeps the benchmark transparent and avoids losing observations from firm-specific estimation-window requirements. The resulting estimates are interpreted as market-adjusted filing-window associations, not as full market-model abnormal returns.

Raw cumulative return is defined as

$$RawReturn_{i,f}[a,b] = \sum_{\tau=a}^{b} R_{i,f+\tau} \tag{3}$$

The daily market-adjusted abnormal return is defined as

$$AR_{i,f+\tau} = R_{i,f+\tau} - R_{m,f+\tau} \tag{4}$$

The cumulative abnormal return is defined as

$$CAR_{i,f}[a,b] = \sum_{\tau=a}^{b} A\,R_{i,f+\tau} \tag{5}$$

The H1 specification is:

$$Return_{i,f,[a,b]} = \beta_1 StrategyAI_{i,t}^{z} + \beta_2 RiskAI_{i,t}^{z} + \beta_3 AccountingNoteAI_{i,t}^{z} + \gamma' X_{i,t} + \alpha_i + \delta_{FY(t)} + \varepsilon_{i,f,[a,b]} \tag{6}$$

The return-window outcome is estimated separately as raw cumulative return and market-adjusted CAR for each event window. In this notation, f indexes the filing event and t indexes the disclosure fiscal year associated with the 10-K text. The disclosure-fiscal-year fixed-effect term denotes the year effect used in the implemented H1 specification.

In the H1 specification, StrategyAI, RiskAI, and AccountingNoteAI are the section-specific AI disclosure measures defined in Section 3.2. The control vector includes size, leverage, cash ratio, and sales growth, measured in the disclosure fiscal year. The dependent variable is measured around the 10-K filing date. The coefficient on StrategyAI tests whether strategy-oriented AI disclosure is associated with filing-window or post-filing returns after controlling for risk-oriented and accounting-note AI disclosure. The coefficients on RiskAI and AccountingNoteAI capture the incremental return associations of AI language in risk-factor and accounting-note contexts. This section-specific design is consistent with evidence that 10-K textual content differs across reporting locations and with the broader principle that financial-text measurement should be tailored to the disclosure setting (Loughran & McDonald, 2011; Dyer et al., 2017; Akyol & Shekhar, 2026).

In the implemented H1 regressions, firm fixed effects are applied at the cleaned firm-panel identifier level and standard errors are clustered at the same level. H1 uses disclosure-fiscal-year fixed effects because the AI disclosure variables are linked to the fiscal year reported in the 10-K. Since many 10-K filings occur in the following calendar year, market-adjusted CAR is used to remove broad market movement around the actual filing window, while the disclosure-fiscal-year fixed effect absorbs common reporting-period effect and AI-salience shocks. However, due to timing limitations, this choice may not fully absorb calendar-year filing-market conditions for firms with non-calendar fiscal years. The H1 estimates should therefore be interpreted as disclosure-fiscal-year-conditioned filing-window associations rather than as a pure filing-calendar-year design. H2 and H3 use fiscal-year fixed effects because their outcomes are annual accounting variables measured by fiscal year.

The coefficients of interest in H1 are β1, β2, and β3 because H1 is a section-specific market-reaction test. β1 captures the incremental filing-window association of StrategyAI, β2 captures the incremental association of RiskAI, and β3 captures the incremental association of AccountingNoteAI. A positive and statistically significant β1 would be consistent with investors treating strategy-oriented AI disclosure as positive incremental information, while a negative β1 would be consistent with concerns about implementation costs, credibility, hype, or execution risk. For β2 and β3, the sign is interpreted by disclosure domain rather than as a single pooled effect. A statistically insignificant coefficient for a given disclosure domain suggests that the filing does not provide a detectable incremental market association for that domain within the specified window.

The RiskAI coefficient is interpreted empirically. A negative coefficient would be consistent with uncertainty, legal exposure, cybersecurity risk, regulatory risk, or

implementation-risk concerns, while a positive coefficient could indicate that investors value risk-specific transparency or more concrete discussion of AI-related operations. The AccountingNoteAI coefficient β3 is interpreted as the market response to AI language embedded in accounting-note contexts, where disclosures may either clarify recognition and measurement implications or raise concerns about accounting complexity.

3.4.2. H2 Operating-Performance Specification

The H2 specification uses one-year-ahead accounting outcomes. For a disclosure observed in fiscal year t, the dependent variable is measured in fiscal year t+1. Because the disclosure panel extends through fiscal year 2024, fiscal-year 2024 observations can enter H2 only when the corresponding fiscal-year 2025 outcome is available. This rule keeps the disclosure variable temporally prior to the accounting outcome.

The omission of a lagged dependent variable is deliberate. H2 is not specified as a dynamic persistence model. It tests whether disclosure-year AI language is associated with one-year-ahead operating outcomes after firm fixed effects, year fixed effects, and contemporaneous firm fundamentals. Including a lagged dependent variable would change the estimand toward a persistence-adjusted forecasting model rather than the thesis's main disclosure-informativeness test. Firm fixed effects, year fixed effects, and controls reduce persistent firm differences and common shocks, but they do not eliminate all time-varying prior-performance channels. Accordingly, H2 coefficients are interpreted as conditional one-year-ahead associations, not as causal effects or persistence-adjusted forecasts.

The H2 specification is:

$$Y_{i,t+1} = \beta_1 StrategyAI_{i,t}^{z} + \beta_2 RiskAI_{i,t}^{z} + \beta_3 AccountingNoteAI_{i,t}^{z} + \gamma' X_{i,t} + \alpha_i + \delta_t + \varepsilon_{i,t+1} \quad (7)$$

The dependent variable in the equation is estimated separately for operating income to revenue, net income to revenue, SG&A expense to revenue, and other operating expenses to revenue. Operating income to revenue and net income to revenue capture profitability and margin outcomes. SG&A expense to revenue and other operating expenses to revenue capture operating cost intensity. This outcome structure is appropriate because the economic consequences of AI may appear either through higher margins or through lower operating cost intensity.

The H2 coefficients of interest are β1, β2, and β3. β1 is the expected primary operating-content coefficient because StrategyAI is predicted to have the strongest directionally favorable relation with later operating outcomes, positive for profitability outcomes and negative for expense-intensity outcomes, but β2 and β3 are also part of the H2 test because H2 evaluates coefficient patterns across strategy, risk, and accounting-note domains. For profitability outcomes, a positive β1 would be consistent with strategy-oriented AI disclosure containing information associated with stronger future margins. For cost-

intensity outcomes, a negative β1 would be consistent with strategy-oriented AI disclosure containing information associated with lower future expense intensity. β2 and β3 test whether risk-factor and accounting-note AI language contain incremental operating-performance associations after the other section-specific AI disclosure domains are controlled for.

This design is consistent with prior narrative-disclosure research showing that textual disclosure can provide information about future firm outcomes. Merkley (2014) examines R&D narrative disclosure and finds that narrative disclosure can be informative to market participants, including through analyst behavior, information content, and information asymmetry.

H2 is not a causal test of whether internal AI use creates performance improvements. The model does not observe project-level deployment, AI capital expenditure, internal model deployment, employee-level use, or the full cost of organizational change. The H2 coefficients are therefore interpreted as associations between section-specific AI disclosure domains and future operating outcomes, not as causal effects of internal AI use. This interpretation is consistent with the broader disclosure-informativeness design of the thesis.

R&D variables are excluded from the broad H2 specification. This choice is deliberate. H2 asks whether the AI disclosure variables themselves are associated with future operating outcomes. Including R&D as a general control could absorb part of the investment channel that H3 is designed to test separately. As a result, H2 estimates the total section-specific association between AI disclosure and future operating outcomes rather than associations net of observed R&D investment. This separation keeps H2 as a broad informativeness test, prevents the H2 section-specific disclosure test from being converted into an R&D-adjusted performance model, and reserves observed R&D backing for the narrower moderation test.

3.4.3. H3 Strict R&D-Backing Specification

H3 is a narrow moderation test. It does not test whether all AI disclosure is backed by investment and it does not test whether reported R&D is AI-specific. It tests whether the StrategyAI relation with future operating outcomes is stronger when the firm-year has higher strict observed R&D backing. Strict weighted RDStockTA is first constructed as a raw asset-scaled stock of observed R&D investment using current R&D, four observed lagged values, and disclosure-year total assets.

The H3 interaction specification is:

$$Y_{i,t+1} = \beta_1 StrategyAI^{z}_{i,t} + \beta_2 RiskAI^{z}_{i,t} + \beta_3 AccountingNoteAI^{z}_{i,t} + \beta_4 RDStockTA^{z,strict}_{i,t} + \beta_5 StrategyAI^{z}_{i,t} \times RDStockTA^{z,strict}_{i,t} + \gamma' X_{i,t} + \alpha_i + \delta_t + \varepsilon_{i,t+1} \quad (8)$$

The narrowness of the R&D backing proxy is important for interpretation. Strict weighted RDStockTA is a conservative observed innovation-capacity moderator, not a direct measure of AI-specific investment. AI-related investment may be embedded in software subscriptions, data infrastructure, cloud services, vendor systems, acquired platforms, proprietary data, organizational redesign, and human capital that are not separately reported in financial statements. In H3, StrategyAI is the thesis's own text-based measure, and observed R&D capacity is used as a conditioning proxy. The motivation is that R&D-related accounting measures and narrative information can be value-relevant, informative, and associated with information asymmetry (Lev & Sougiannis, 1996; Aboody & Lev, 2000; Merkley, 2014). At the same time, reported R&D is an imperfect proxy for innovation activity because missing R&D does not necessarily indicate the absence of innovation-related activity (Koh & Reeb, 2015). The design therefore treats observed R&D as a strict backing context rather than as a complete measure of AI capability.

The interaction term is not an accounting adjustment. It is not used to revise operating income, net income, SG&A expense, or other operating expenses. Its role is to test whether the disclosure-performance relation is conditional on observed R&D backing. This keeps the dependent variables conceptually separate from the moderator and avoids blending the performance outcome with the proposed credibility mechanism.

3.4.4. Fixed Effects, Standard Errors, and Interpretation

All main firm-year and filing-event regressions report one-way clustered standard errors at the cleaned firm-panel identifier level, together with firm and year fixed effects. In the accounting panel regressions, firm fixed effects are implemented at the cleaned firm-panel identifier level used in the constructed firm-year panel, and standard errors are clustered at the same level. CIK is used to anchor EDGAR issuer filings, while CRSP PERMNO/LPERMNO is used for event-study return matching rather than as the fixed-effect unit in the accounting panel. This clustering choice follows finance and accounting panel-inference guidance that standard errors should account for residual dependence within firms over time (Petersen, 2009; Gow et al., 2010).

For the event-study sample, CIK and PERMNO/LPERMNO identify the filing issuer and the matched security return, but firm fixed effects and clustered standard errors are reported at the cleaned firm-panel identifier level used elsewhere in the analysis. Clustering at this level allows residuals to be correlated within firms across years and keeps the H1 inference comparable to the H2 and H3 tables.

The fixed-effect design removes time-invariant firm heterogeneity and common year shocks, but it does not eliminate time-varying omitted variables, reverse causality, or measurement error in AI disclosure. The estimates are therefore interpreted as conditional associations.

Continuous accounting ratios are winsorized at the year-specific 1st and 99th percentiles, as described in Section 3.3.4. The strict H3 design reduces both observations and firm

clusters because it requires observed current and lagged R&D inputs. H3 should therefore be interpreted as a narrower and lower-power moderation test than H2. This design strengthens the observed-data interpretation of RDStockTA but limits generalization to firms whose AI capability is built through unreported software purchases, cloud services, acquired platforms, vendor systems, or organizational process changes.

The compact result tables in section 4 report the AI-disclosure coefficients, observation counts, cluster counts, fixed-effect structure, clustering choice, and dependent-variable definitions. Appendices A.2, A.3, and A.5 report the coefficient-with-controls estimates, standard errors, adjusted R^2, control variables, fixed-effect structure, and clustering details, while Appendix A.4 reports H2 persistence-control robustness diagnostics.

The interpretation of the empirical design is conservative throughout. H1 estimates whether section-specific AI disclosure is associated with filing-period or post-filing returns. H2 estimates whether section-specific AI disclosure is associated with future operating outcomes. H3 estimates whether observed R&D backing moderates the StrategyAI-outcome relation. None of the tests proves that AI disclosure mechanically causes market returns or operating performance. The design instead tests whether section-specific AI disclosure is associated with market returns, future operating outcomes, and R&D-backed moderation under the stated sample and measurement rules.

3.4.5. Summary of Identification Logic

The empirical design separates measurement, timing, and interpretation. The first layer is measurement discipline because AI disclosure is separated into StrategyAI, RiskAI, and AccountingNoteAI rather than pooled into a single AI mention count. The second layer is timing discipline because H1 uses filing-window and post-filing returns, while H2 and H3 use lead operating outcomes. The third layer is interpretation discipline because the regressions are treated as conditional associations rather than causal estimates.

This structure keeps the market response, operating performance, and R&D backing tests separate. H1 uses market returns because the hypothesis concerns investor response to section-specific AI disclosure around the 10-K filing. H2 uses accounting outcomes because the hypothesis concerns the future operating informativeness of strategy-oriented, risk-oriented, and accounting-note AI disclosure domains. H3 uses an interaction model because the hypothesis asks whether observed R&D backing strengthens the informativeness of strategy-oriented AI disclosure. The empirical design therefore separates market-response, future-operating-informativeness, and observed R&D-moderation tests, preventing one coefficient from being interpreted as evidence for all disclosure mechanisms.

4. Empirical Results

4.1. Descriptive Evidence and R&D Coverage

Table 4 reports descriptive statistics for the main firm-year frame before outcome-specific regression filters are applied. The table includes AI disclosure ratios, operating ratios, core controls, and observed R&D variables. AI disclosure is sparse and right-skewed, as the median strategy, risk, and accounting-note ratios are zero. The table therefore reports 90th percentile, 95th percentile, and nonzero-share statistics to show the upper-tail variation that supports the regression analysis. Figure 1 reports the yearly nonzero AI-disclosure shares by section block. Appendix Table 19 reports pairwise Pearson correlations among the word- and sentence-intensity disclosure measures, operating outcomes, controls, and strict R&D-stock variables.

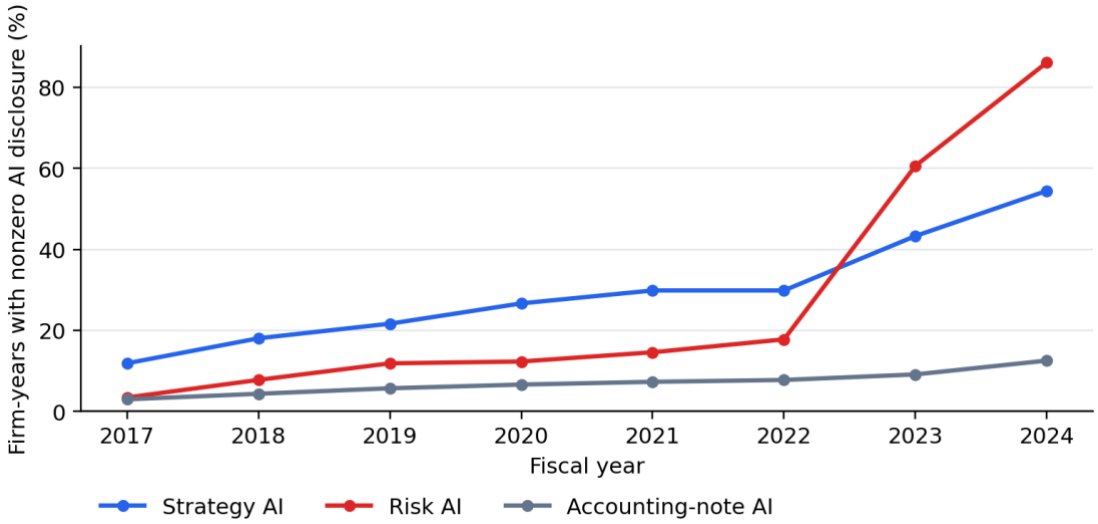
For scale interpretation, a word ratio of 0.0002 corresponds to approximately 0.2 accepted AI word tokens per 1,000 cleaned section-block words. Scaling the ratios this way makes the sparse AI disclosure distribution easier to interpret before standardized regression coefficients are introduced.

R&D coverage is the main sample constraint for H3. The strict weighted RDStockTA sample requires reported current R&D expense, four lagged observed R&D values, and disclosure-year total assets. Missing R&D is retained as missing rather than recoded to zero. The strict R&D design is therefore a narrower observed-data mechanism test rather than a broad test of all AI investment backing. Table 5 and Figure 2 report this coverage constraint.

Table 4. Descriptive statistics for AI disclosure, operating ratios, controls, and R&D variables

Variable	N	Mean	SD	Median	P75	P90	P95	NZ share
Strategy AI word ratio	3510	0.0002	0.0008	0.0000	0.0001	0.0005	0.0011	29.5%
Risk AI word ratio	3510	0.0003	0.0007	0.0000	0.0001	0.0010	0.0017	26.8%
AccountingNoteAI word ratio	3510	0.0000	0.0002	0.0000	0.0000	0.0000	0.0001	7.0%
Strategy AI sentence ratio	3510	0.0025	0.0099	0.0000	0.0012	0.0049	0.0118	29.5%
Risk AI sentence ratio	3510	0.0043	0.0200	0.0000	0.0021	0.0131	0.0265	26.8%
AccountingNoteAI sentence ratio	3510	0.0004	0.0032	0.0000	0.0000	0.0000	0.0013	7.0%
Operating income / revenue	3510	0.2022	0.1375	0.1851	0.2730	0.3797	0.4520	n.a.
Net income / revenue	3510	0.1404	0.1367	0.1245	0.2069	0.2985	0.3600	n.a.
SG&A / revenue	3415	0.1855	0.1520	0.1596	0.2621	0.3965	0.5153	n.a.
Other operating expenses / revenue	3508	0.2824	0.1784	0.2432	0.3703	0.5496	0.6412	n.a.
Size (log assets)	3509	17.0563	1.3781	16.9707	17.8955	18.9058	19.4960	n.a.
Leverage	3509	0.3242	0.1987	0.3179	0.4348	0.5521	0.6558	n.a.
Cash ratio	3509	0.0833	0.0900	0.0537	0.1147	0.1936	0.2677	n.a.
Sales growth	3510	0.0922	0.1928	0.0679	0.1482	0.2602	0.3968	n.a.
Current R&D/assets	1610	0.0498	0.0578	0.0305	0.0715	0.1142	0.1552	n.a.
Strict weighted RDStockTA	1367	0.1331	0.1400	0.0879	0.1940	0.2972	0.3841	n.a.

Note: Disclosure ratios and accounting ratios are rounded to four decimal places. Reported zero values may therefore reflect rounding rather than exact zeros. NZ share is identical across word and sentence rows by construction, because the matching unit is the sentence. Operating-performance ratios in Table 4 are measured in the main disclosure-year firm-year frame before lead-outcome construction. H2 and H3 regression counts differ because those models require t+1 outcome availability and complete control variables.



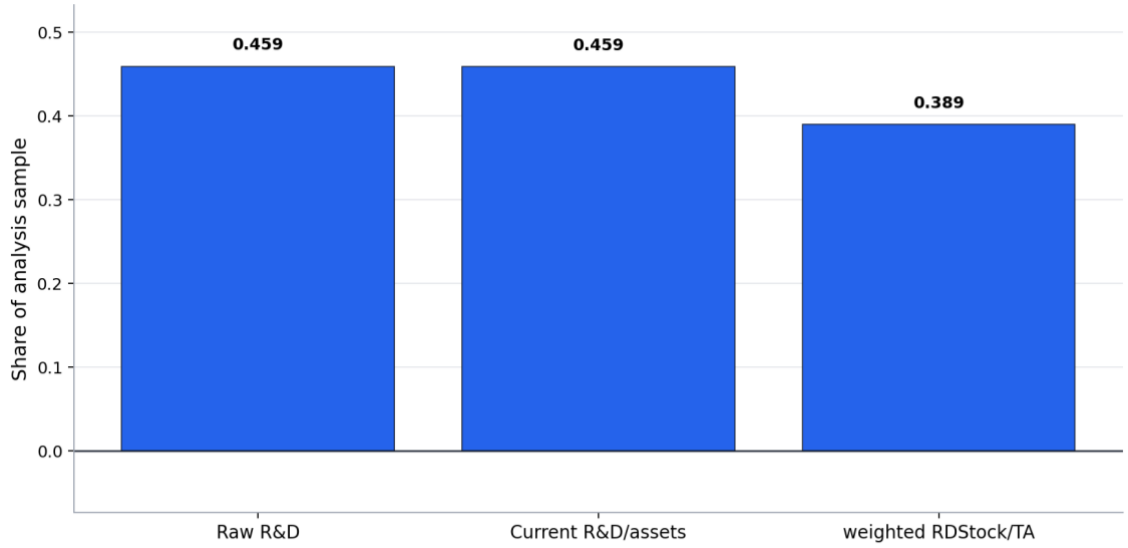
Note: Figure reports the percentage of main-frame firm-years with positive word-intensity AI disclosure ratios in each section block by fiscal year. Ratios are measured in the disclosure-year 10-K text.

Figure 1. Yearly nonzero AI disclosure shares by section block

Table 5. R&D data coverage under strict observed-data rules

Metric	Value	Share
Analysis rows	3510	100.0%
Raw R&D non-missing	1611	45.9%
Raw R&D positive	1611	45.9%
Strict current R&D/assets non-missing	1610	45.9%
Strict weighted RDStockTA non-missing	1367	38.9%

Note: The cleaned strict-R&D input contains positive reported R&D values when R&D is observed. Missing R&D is retained as missing and is not interpreted as zero R&D activity. The equality between raw R&D non-missing and raw R&D positive reflects the cleaned observed R&D input; observed zeros are not recoded from missing values. The difference between raw R&D non-missing and strict current R&D/assets non-missing reflects the positive-assets requirement. Strict weighted RDStockTA additionally requires current R&D, four observed lagged R&D values, and positive current total assets.



Note: Bars show the share of the main analysis sample. Strict H3 tests use observed R&D firm-years only.

Figure 2. Strict observed R&D sample coverage for H3

4.2. H1 Market-Reaction Results

H1 provides mixed evidence of section-specific filing-window market associations but it does not support a consistently positive market reaction to AI disclosure. In the word-intensity specification, StrategyAI is negatively associated with the immediate and short post-filing market-adjusted CAR windows. A one-standard-deviation increase in StrategyAI is associated with a 0.36 percentage point lower market-adjusted CAR in the [-1,+1] window and a 0.72 percentage point lower market-adjusted CAR in the [0,+5] window. This evidence does not support a positive market-reaction interpretation for strategy-oriented AI language.

RiskAI provides an important section-specific contrast. In the [0,+5] window, RiskAI is positive and statistically significant in both the raw-return and market-adjusted specifications. A one-standard-deviation increase in RiskAI is associated with a 0.33 percentage point higher market-adjusted CAR in that window. This pattern suggests that the filing-window market association differs across AI disclosure domains rather than reflecting a single pooled AI-disclosure effect.

The H1 interpretation follows standard event-study logic, where abnormal returns are used to assess market responses around a dated information event (Brown & Warner, 1985; MacKinlay, 1997). The analysis emphasizes market-adjusted CAR because the raw cumulative-return does not control for broad market movement during the filing window. The [0,+60] and [0,+126] windows are reported as longer horizon diagnostic evidence, but they are not interpreted as clean filing-news effects because many non-filing events may affect returns over those periods.

The H1 regressions condition on firm scale, leverage, liquidity, and sales growth and include firm fixed effects, disclosure-fiscal-year fixed effects, and standard errors clustered at the cleaned firm-panel identifier level. The reported non-positive strategy-AI coefficients are therefore conditional within-firm, year-adjusted associations rather than simple raw differences across firms.

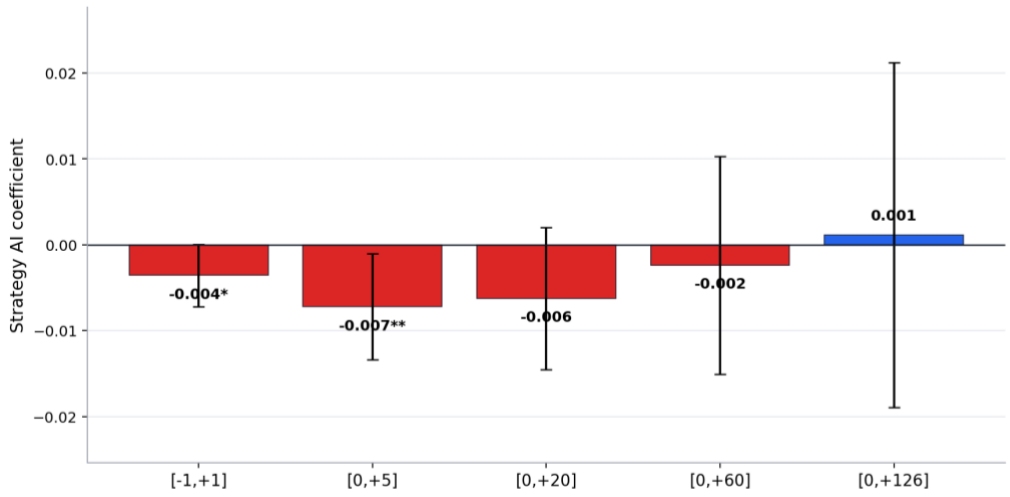
Overall, H1 provides partial evidence consistent with a section-specific market-association hypothesis. StrategyAI shows negative short-window market-adjusted associations, while RiskAI shows a positive short-window association in the [0,+5] word-intensity specification. In the word-intensity results, the statistically significant StrategyAI and RiskAI coefficients move in opposite directions, indicating that filing-window market associations differ across AI disclosure domains. The sentence-intensity estimates mainly reinforce the negative StrategyAI pattern, while the positive RiskAI result is not equally stable across measurement specifications. Hence, the H1 evidence supports a section-specific interpretation of market response, but it does not evaluate internal AI use, reject the economic relevance of AI-related activity, or imply that all AI disclosure channels lack market relevance.

Table 6 reports the H1 word-intensity event-study coefficients. Figure 3 plots the StrategyAI coefficients from market-adjusted CAR specification. The sentence-intensity H1 estimates in Appendix A.1, Table 10 are directionally consistent with the negative short-window StrategyAI market-adjusted CAR pattern, while the positive RiskAI result is not stable across measurement specifications.

Table 6. H1 word-intensity filing-window and post-filing return regressions

Spec	Window	Dependent variable	Strategy AI	Risk AI	AccountingNoteAI	N	Clusters
Word	[-1, +1]	Cumulative raw return	-0.0032 (0.0021)	0.0017 (0.0014)	0.0003 (0.0017)	3301	431
Word	[-1, +1]	Market-adjusted CAR	-0.0036* (0.0018)	0.0016 (0.0013)	-0.0004 (0.0019)	3301	431
Word	[0, +5]	Cumulative raw return	-0.0048 (0.0033)	0.0032* (0.0017)	0.0027 (0.0036)	3376	431
Word	[0, +5]	Market-adjusted CAR	-0.0072** (0.0031)	0.0033** (0.0016)	0.0019 (0.0028)	3376	431
Word	[0, +20]	Cumulative raw return	-0.0033 (0.0040)	-0.0005 (0.0026)	0.0048 (0.0059)	3376	431
Word	[0, +20]	Market-adjusted CAR	-0.0063 (0.0042)	-0.0011 (0.0023)	0.0022 (0.0038)	3376	431
Word	[0, +60]	Cumulative raw return	-0.0016 (0.0074)	0.0001 (0.0037)	0.0006 (0.0044)	3374	431
Word	[0, +60]	Market-adjusted CAR	-0.0024 (0.0065)	-0.0008 (0.0036)	-0.0006 (0.0037)	3374	431
Word	[0, +126]	Cumulative raw return	0.0024 (0.0106)	-0.0066 (0.0064)	0.0052 (0.0039)	3373	431
Word	[0, +126]	Market-adjusted CAR	0.0012 (0.0103)	-0.0062 (0.0064)	0.0049 (0.0041)	3373	431

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Clustered standard errors for the reported AI coefficients are shown in parentheses. Reported coefficients are for AI disclosure variables only. H1 is evaluated as a section-specific association test using StrategyAI, RiskAI, and AccountingNoteAI jointly. Coefficients should therefore be interpreted by disclosure domain rather than as a single pooled AI-disclosure effect. AI disclosure variables are standardized before estimation; dependent variables remain in their reported units. Controls are size, leverage, cash ratio, and sales growth measured in the disclosure fiscal year. Firm fixed effects and disclosure-fiscal-year fixed effects are included but omitted for compactness. Event-window N varies because return-window availability differs across filing windows. Raw cumulative return is reported separately from market-adjusted CAR. Market-adjusted CAR subtracts the CRSP value-weighted market return from firm daily returns before cumulation. CAR refers only to the cumulative abnormal return specification. Sentence-intensity H1 estimates are reported in Appendix A.1, Table 10.



Note: Word-intensity market-adjusted strategy coefficients from Table 6. Figure plots the StrategyAI word-intensity coefficients from the market-adjusted CAR specifications only in Table 6. This figure illustrates the StrategyAI component only. H1 is evaluated using StrategyAI, RiskAI, and AccountingNoteAI jointly. RiskAI and AccountingNoteAI coefficients are reported in Table 6. Error bars report approximate 95 percent confidence intervals based on cleaned firm-panel identifier clustered standard errors. Stars denote p<0.10/0.05/0.01. AI disclosure variables are standardized before estimation. The dependent variable remains in return units.

Figure 3. StrategyAI component of H1 market-adjusted CAR results

4.3. H2 Future Operating-Performance Results

H2 receives its strongest support through the StrategyAI coefficients. The operating-performance evidence is concentrated in strategy-oriented AI disclosure rather than evenly distributed across all AI disclosure domains. The evidence is strongest for future profitability margins, while RiskAI and AccountingNoteAI do not show stable incremental associations.

In the word-intensity specification, a one-standard-deviation increase in StrategyAI is associated with a 0.77 percentage point increase in next-year operating income to revenue and a 1.04 percentage point increase in next-year net income to revenue. The same increase is associated with a 0.41 percentage point decrease in SG&A to revenue and a 0.64 percentage point decrease in other operating expenses to revenue. In the sentence-intensity specification, the operating income relation remains positive at 0.62 percentage points, while the net income and other operating expense results weaken.

The H2 evidence is concentrated in StrategyAI and is strongest for margin outcomes. It does not show that all AI disclosure domains predict future performance. RiskAI and AccountingNoteAI remain important in the specification because they show that the StrategyAI coefficient is not simply capturing a general tendency to mention AI anywhere in the 10-K filing.

The result is also economically interpretable because the dependent variables are future accounting ratios rather than contemporaneous disclosure or market-perception measures. Operating income to revenue and net income to revenue capture profitability outcomes, while selling, general, and administrative expense to revenue and other operating expenses to revenue capture cost-intensity outcomes. The evidence is strongest for profitability margins, suggesting that strategy-oriented AI language is more clearly associated with future margin outcomes than with a simple cost-reduction channel.

Control coefficients are reported in Appendix A.3, Tables 13 and 14. The StrategyAI coefficient remains positive for the main margin outcomes after conditioning on size, leverage, cash ratio, sales growth, firm fixed effects, and year fixed effects.

The H2 interpretation remains associational. The evidence does not show that AI disclosure causes operating performance or that the observed firms implemented AI internally. The model does not observe project-level deployment, implementation quality, managerial ability, product market positioning, employee level AI use, or prior investor information. The result instead shows that, within a model that jointly includes StrategyAI, RiskAI, and AccountingNoteAI, the future operating-performance association is concentrated in strategy-oriented AI language.

The H2 result is interpreted within the thesis's disclosure-informativeness design rather than as an internal AI use effect. The result is narrow because it focuses on section-specific 10-K disclosure domains and one-year-ahead accounting ratios. It therefore supports an operating-content interpretation only within the sample, measurement, and fixed-effect structure used here.

The H2 pattern should also be read with the lagged-performance boundary from Section 3.4.2 in mind. The estimates show a conditional association between strategy-oriented AI disclosure and future operating margins, not a causal or fully persistence-adjusted performance effect.

The distinction between profitability and expense outcomes is useful for interpreting the result. Operating income and net income capture broad profitability, while SG&A and other operating expenses capture cost intensity. The evidence is stronger for future margin outcomes than for a simple cost-reduction channel, suggesting that StrategyAI language is more clearly associated with later profitability than with isolated expense reductions.

The H2 evidence is strongest in the word specification, suggesting that repeated AI mentions within strategy sections carry more information than a sentence-count measure. The positive operating-income result in the sentence specification strengthens the finding relative to a single-measure result, while the weaker sentence-based net income and other operating expense coefficients limit the strength of the conclusion.

These magnitudes are associations measured in standard deviations of disclosure intensity, not counts of AI projects. A one-standard-deviation increase in StrategyAI should not be

interpreted as one additional AI project, one AI investment, or one internal AI-use event. The result shows that higher StrategyAI, conditional on firm fixed effects, year fixed effects, and disclosure-year controls, is associated with stronger subsequent operating margins.

The concentration of the H2 result in StrategyAI supports the section-level design and reduces the likelihood that the finding is merely a generic AI-mention effect. RiskAI and AccountingNoteAI remain useful controls because they show that the strategy coefficient is not simply a document wide tendency to mention AI.

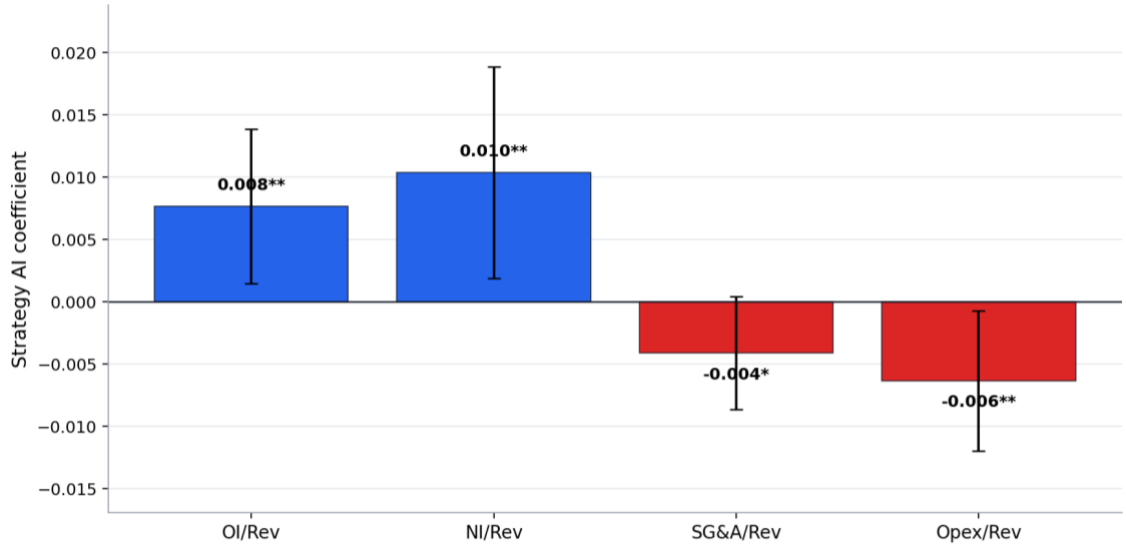
Appendix A.4 addresses this persistence boundary directly by re-estimating the H2 models after adding current and lagged values of the same operating outcome. The StrategyAI pattern remains directionally consistent, especially in the word-intensity specification, while RiskAI and AccountingNoteAI do not show similarly stable patterns. The robustness evidence is therefore supportive of the H2 interpretation, but it remains diagnostic rather than causal.

Table 7 reports the H2 operating-performance regressions for all three disclosure domains. Figure 4 summarizes the main word-intensity StrategyAI coefficients because the meaningful H2 evidence is concentrated in StrategyAI. RiskAI and AccountingNoteAI coefficients remain in Table 7, and the sentence-intensity estimates are reported in Table 7 to preserve the parallel-measurement design.

Table 7. H2 future operating-performance regressions

Spec	Outcome t+1	Strategy AI	Risk AI	AccountingNoteAI	N	Clusters
Word	Operating income / revenue	0.0077** (0.0032)	-0.0005 (0.0019)	0.0036 (0.0032)	3497	439
Word	Net income / revenue	0.0104** (0.0043)	-0.0028 (0.0025)	0.0023 (0.0029)	3497	439
Word	SG&A / revenue	-0.0041* (0.0023)	-0.0013 (0.0009)	-0.0004 (0.0009)	3406	429
Word	Other operating expenses / revenue	-0.0064** (0.0029)	0.0004 (0.0014)	-0.0010 (0.0018)	3496	439
Sentence	Operating income / revenue	0.0062* (0.0037)	0.0006 (0.0006)	0.0026 (0.0039)	3497	439
Sentence	Net income / revenue	0.0062 (0.0048)	0.0007 (0.0008)	0.0026 (0.0032)	3497	439
Sentence	SG&A / revenue	-0.0042* (0.0023)	0.0002 (0.0004)	-0.0001 (0.0006)	3406	429
Sentence	Other operating expenses / revenue	-0.0051 (0.0033)	0.0006 (0.0004)	-0.0014 (0.0021)	3496	439

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Clustered standard errors for the reported AI coefficients are shown in parentheses. Reported coefficients are for AI disclosure variables only. H2 is evaluated using all three section-specific AI disclosure domains. The StrategyAI coefficient is the expected operating-content coefficient, while RiskAI and AccountingNoteAI test whether risk-factor and accounting-note AI language contain incremental operating-performance associations. AI disclosure variables are standardized before estimation; dependent variables remain in their reported units. Controls are size, leverage, cash ratio, and sales growth measured in fiscal year t. Firm fixed effects and year fixed effects are included but omitted for compactness. Corresponding H2 coefficient-with-controls tables are reported in Appendix A.3, Tables 13 and 14. Standard errors are clustered at the cleaned firm-panel identifier level. N varies by outcome because t+1 outcome availability and SG&A coverage differ across firm-years.



Note: Word-intensity strategy coefficients from Table 7. Figure plots the word-intensity StrategyAI coefficients from the H2 operating-performance regressions in Table 7. This figure illustrates the StrategyAI component only because the H2 evidence is concentrated in StrategyAI. H2 is evaluated using StrategyAI, RiskAI, and AccountingNoteAI jointly. Error bars report approximate 95 percent confidence intervals based on cleaned firm-panel identifier clustered standard errors. Stars denote p<0.10/0.05/0.01. Positive coefficients for profitability outcomes and negative coefficients for expense-intensity outcomes are interpreted as favorable operating-performance associations. RiskAI and AccountingNoteAI coefficients, along with sentence-intensity estimates, are shown in Table 7.

Figure 4. StrategyAI component of H2 future operating-outcome results

4.4. H3 Strict R&D-Backing Results

H3 is not supported. The strict observed R&D interaction does not provide positive evidence that R&D backing strengthens the StrategyAI-performance relation.

In the strict observed R&D subsample, the StrategyAI main effects retain the directional pattern observed in H2, although the word-intensity NI/revenue coefficient is notably larger than the corresponding H2 estimate, reflecting the smaller and more R&D-intensive subsample.

The only statistically significant interaction appears in the word-intensity net-income specification, and its sign is negative. Because this negative interaction is not statistically significant in the sentence-intensity specification, it is interpreted as evidence against positive R&D-backed moderation rather than as robust evidence of a negative moderation mechanism.

The StrategyAI main effects in Table 8 are not directly comparable to the broader H2 coefficients in Table 7. H3 is estimated in a smaller strict observed R&D sample, the variables are standardized within the strict observed R&D coverage frame, and the model includes both strict weighted RDStockTA and the StrategyAI interaction term. The significant StrategyAI main effects show that the H2-style StrategyAI pattern remains visible in the observed R&D subsample. However, they do not support the moderation hypothesis, because H3 is tested through the StrategyAI × strict weighted RDStockTA interaction rather than through the StrategyAI main effect alone.

The strict R&D design also reduces firm clusters from 439 in the broader H2 frame to 204 in Table 8. The interaction estimates should therefore be interpreted as lower-power tests of a narrower mechanism. RiskAI and AccountingNoteAI are not interpreted as H3 moderation channels because they enter only as non-interacted section-specific disclosure controls.

The H3 result narrows the thesis interpretation. Observed R&D capacity is conceptually relevant, but the strict interaction test does not show that R&D stock reliably strengthens the strategy-disclosure relation. The conservative conclusion is that the StrategyAI remains directionally consistent with the H2 operating-margin evidence in the strict observed R&D subsample, while the specific strict weighted RDStockTA moderation mechanism is not supported. R&D is therefore interpreted only within H3 rather than blended into the broad H2 margin tests.

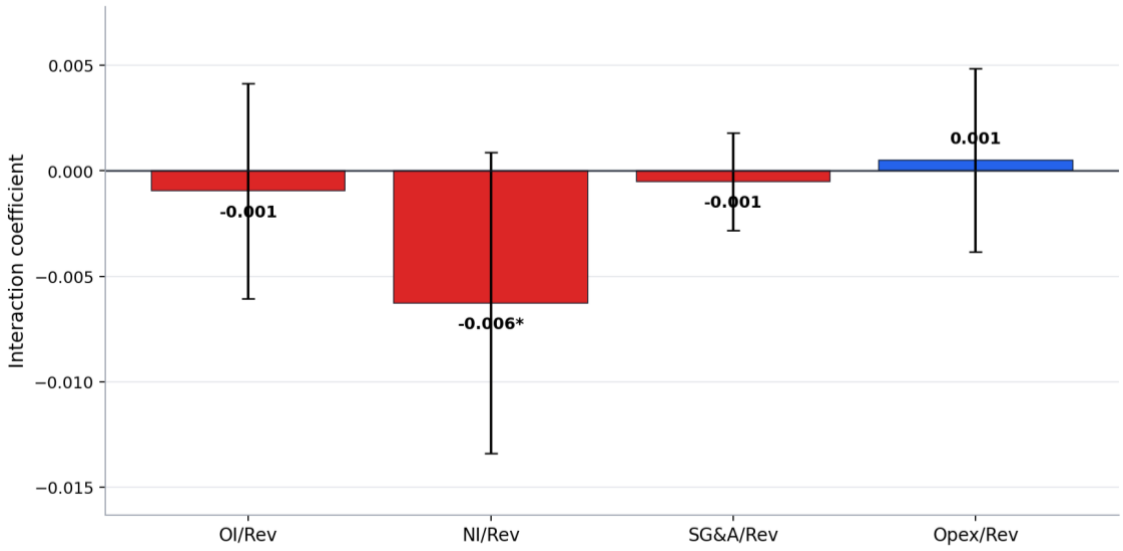
The H3 result should not be interpreted as evidence that investment backing is irrelevant. Strict weighted RDStockTA is not an AI-specific investment measure. It is a conservative observed R&D proxy that requires current R&D, four observed lagged R&D inputs, and positive disclosure-year assets. The test does not observe AI-specific software, data, cloud infrastructure, vendor systems, process redesign, or AI-related human capital. H3 therefore shows only that the specific broad R&D-stock proxy used here does not positively moderate the StrategyAI-performance association in the strict observed R&D subsample.

Table 8 reports the H3 strict R&D-backing regressions. Figure 5 plots the StrategyAI × strict weighted RDStockTA interaction coefficients used to interpret the moderation evidence.

Table 8. H3 strict R&D-backing interaction regressions

Spec	Outcome t+1	Strategy AI	Strict weighted RDStockTA	Strategy AI × strict weighted RDStockTA	N	Clusters
Word	Operating income / revenue	0.0083* (0.0043)	0.0152 (0.0131)	-0.0010 (0.0026)	1362	204
Word	Net income / revenue	0.0202*** (0.0066)	0.0278 (0.0180)	-0.0063* (0.0036)	1362	204
Word	SG&A / revenue	-0.0043* (0.0023)	-0.0063 (0.0070)	-0.0005 (0.0012)	1362	204
Word	Other operating expenses / revenue	-0.0077** (0.0035)	-0.0121 (0.0115)	0.0005 (0.0022)	1362	204
Sentence	Operating income / revenue	0.0079* (0.0046)	0.0165 (0.0132)	-0.0004 (0.0027)	1362	204
Sentence	Net income / revenue	0.0147** (0.0067)	0.0278 (0.0181)	-0.0046 (0.0035)	1362	204
Sentence	SG&A / revenue	-0.0050** (0.0021)	-0.0075 (0.0070)	-0.0007 (0.0013)	1362	204
Sentence	Other operating expenses / revenue	-0.0073** (0.0036)	-0.0131 (0.0116)	-0.0001 (0.0022)	1362	204

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Clustered standard errors for the reported StrategyAI, strict weighted RDStockTA, and interaction coefficients are shown in parentheses. StrategyAI, RiskAI, Accounting-note AI, and strict weighted RDStockTA are standardized within the strict observed R&D coverage frame before regression-specific outcome and control filters are applied. The interaction term is constructed by multiplying standardized StrategyAI by standardized strict weighted RDStockTA. Table 8 coefficients are estimated in the smaller strict observed R&D subsample and are not directly comparable to Table 7 H2 coefficients. RiskAI and Accounting-note AI are included as non-interacted disclosure controls and are reported in Appendix A.5, Tables 17 and 18. Dependent variables remain in their reported units. Controls are size, leverage, cash ratio, and sales growth measured in fiscal year t. Firm fixed effects and year fixed effects are included but omitted for compactness. Corresponding H3 coefficient-with-controls tables are reported in Appendix A.5, Tables 17 and 18. Standard errors are clustered at the cleaned firm-panel identifier level. N is smaller than the strict weighted RDStockTA coverage count as regression observations also require non-missing outcomes, controls, and fixed-effect identifiers.



Note: Bars show coefficients. Figure plots the H3 StrategyAI × strict weighted RDStockTA interaction coefficients across the reported operating outcomes in Table 8. Error bars report approximate 95 percent confidence intervals based on cleaned firm-panel identifier clustered standard errors. *, **, *** denote p<0.10, p<0.05, p<0.01. StrategyAI and strict weighted RDStockTA are standardized within the strict observed R&D coverage frame before regression-specific outcome and control filters are applied and before interaction construction. Dependent variables remain in their reported units.

Figure 5. H3 interaction coefficients across operating outcomes

4.5. Result Summary

The results support a conditional operating-content interpretation rather than an immediate market-signal interpretation. H1 provides mixed market-reaction evidence across disclosure domains. StrategyAI is negatively associated with short-window market-adjusted CARs, RiskAI is positively associated with the [0,+5] market-adjusted CAR, and AccountingNoteAI does not show a stable short-window association. H1 therefore provides partial evidence of section-specific filing-window associations, but it does not support the interpretation that AI disclosure is uniformly rewarded as favorable filing-date news.

H2 provides the clearest empirical support in the thesis, but the claim is narrow. The future operating performance evidence is concentrated in StrategyAI, especially for one-year-ahead operating income to revenue. The word-intensity specification also shows positive net-income evidence and negative expense-intensity evidence, while the sentence-intensity specification provides more limited support. Because financial firms, insurers, and real-estate firms remain in the cross-sector panel, the H2 results are interpreted as large-cap cross-sector operating-ratio associations rather than sector-specific production-function estimates. Because the tests span multiple outcomes and word/sentence specifications, interpretation emphasizes directional stability and economic coherence rather than isolated statistical significance.

H3 is not supported. Strict weighted RDStockTA does not positively moderate the StrategyAI-performance relation. The only statistically significant interaction is negative in the word-based net-income specification and does not replicate as statistically significant in the sentence-based specification. The strict R&D result is therefore interpreted as a boundary on the H2 interpretation rather than as evidence that observed R&D backing explains the StrategyAI operating-performance association.

Across the three tests, AI disclosure has mixed filing-window market associations by section, and future operating-performance content is concentrated in strategy-oriented AI disclosure. The evidence does not support a consistently positive filing-date reaction or a positive strict observed R&D moderation mechanism.

Table 9. Hypothesis status summary

Hypothesis	Main test	Evidence	Conclusion
H1	StrategyAI, RiskAI, and AccountingNoteAI in filing-window returns	StrategyAI short-window MCAR is negative; RiskAI [0,+5] is positive; AccountingNoteAI is not stable.	Partial evidence consistent with section-specific market association, not consistently positive AI-disclosure news
H2	StrategyAI, RiskAI, and AccountingNoteAI in t+1 operating outcomes	Evidence is concentrated in StrategyAI: OI/revenue is positive in word and sentence, NI/revenue is positive in word, and expense evidence is less stable. RiskAI and AccountingNoteAI are not stable.	Supported for section-specific, margin-centered operating content concentrated in StrategyAI
H3	StrategyAI × strict weighted RDStockTA, with RiskAI and AccountingNoteAI as controls	Interaction is not positively significant; word NI/revenue interaction is negative and not sentence-stable.	Not supported

The appendix reports supplementary event-study estimates, coefficient with controls tables, R&D robustness tests, ROIC evidence, reproducibility details, and the AI dictionary validation protocol.

5. Discussion

5.1. Signal, Shield, and Timing

The contrast between the market reaction evidence and the operating performance evidence is economically plausible. Event-study methodology treats abnormal returns as price reactions around a dated information event (Brown & Warner, 1985; MacKinlay, 1997). The H1 CAR estimates should therefore be interpreted as incremental filing-window evidence rather than as a test of the full economic importance of AI activity. The absence of consistently positive filing-window CARs, especially for StrategyAI, therefore does not mean that AI disclosure is meaningless. Instead, it indicates that the 10-K filing does not provide uniformly positive incremental market news across AI disclosure domains.

This timing distinction is important as investors may learn about major AI initiatives before the annual report. Relevant information may already have appeared through earlier earnings announcements, associated conference calls, investor presentations, product announcements, news coverage, or previous filings. Disclosure channel evidence shows that conference calls can transmit and accelerate investors' processing of earnings-related information, while the business press can package and disseminate firm information outside the annual report filing event (Frankel et al., 1999; Kimbrough, 2005; Bushee et al., 2010). The 10-K may therefore formalize, confirm, or legally document information that market participants have already partly incorporated into prices.

The operating-performance evidence provides a different type of informativeness test. While H1 asks whether AI disclosure is associated with filing-window market reactions, H2 asks whether the same disclosure is associated with future accounting outcomes. The H2 results suggest that StrategyAI can contain operating content even when the price response at filing is not favorable. This distinction helps reconcile the main findings: the market reaction evidence does not support the strongest immediate price signal interpretation, while the operating-performance evidence supports a more modest disclosure-informativeness interpretation.

The H3 result adds a further constraint. The strict observed R&D moderation test does not show that the StrategyAI-performance relation is positively amplified by strict weighted RDStockTA. The operating-performance evidence therefore cannot be attributed mechanically to observed R&D-backed innovation. StrategyAI may instead capture managerial attention, implementation intensity, product-market positioning, or organizational change that is not fully reflected in reported R&D expenses.

This interpretation is consistent with disclosure theory. Firms disclose under information asymmetry, proprietary costs, litigation pressure and managerial incentives (Verrecchia, 1983; Healy & Palepu, 2001; Beyer et al., 2010). AI language in the 10-K can therefore serve several purposes at once. It may communicate substantive business activity, frame

uncertainty, document risk exposure, or manage investor expectations. The signal and shield interpretations are therefore not mutually exclusive. A firm can disclose meaningful AI-related activity while also using broad language to manage uncertainty around implementation, costs, and outcomes.

The main implication is that AI disclosure should not be interpreted as automatically informative or automatically promotional. Its informativeness depends on timing, section location, outcome choice, and the observable support mechanism being tested. In this thesis, the strongest evidence appears in future operating margins rather than in immediate filing-window returns, while the strict R&D moderation test is not supported. Annual report AI language can carry operating information, but investors do not uniformly reward it as favorable filing-date news, and the result is not clearly explained by observed R&D backing.

5.2. Why Section Context Matters

The findings reinforce why section context matters in textual disclosure measurement. StrategyAI, RiskAI, and AccountingNoteAI are not interchangeable because they map onto different disclosure functions: business and MD&A language is closer to operating initiatives, risk-factor language is closer to uncertainty or exposure, and footnote language is closer to accounting detail. Keeping these domains separate makes the evidence easier to interpret and reduces the risk that the results are driven only by filing length, boilerplate, or a general tendency to mention AI anywhere in the annual report.

5.3. Interpreting R&D Backing

The R&D results are best read as a boundary on the operating-performance finding rather than as evidence against it. RDStockTA is designed as an observable innovation-capacity moderator, but it is not a direct measure of AI-specific investment. The H3 results do not show that strict weighted RDStockTA positively moderates the StrategyAI-performance relation. The thesis therefore cannot claim that observed R&D backing explains the H2 operating-performance association.

This result is important because R&D is only one possible channel through which firms may build AI capability. AI-related activity may also be embedded in software subscriptions, cloud infrastructure, vendor systems, data engineering, acquired platforms, proprietary data, process redesign, consulting projects, or employee training. These activities may not appear as separately reported R&D in financial statements. The strict R&D proxy is therefore observable and conservative, but incomplete.

The H3 design deliberately reflects this limitation. Missing R&D is not recoded as zero because missing R&D does not necessarily imply the absence of innovation activity (Koh & Reeb, 2015). Strict weighted RDStockTA also requires observed current R&D, four observed lagged R&D values, and disclosure-year total assets. This strengthens the

observed-data interpretation of the moderator, but it also narrows the sample and reduces statistical power.

The H3 evidence therefore has a constraining role. It prevents the thesis from attributing the StrategyAI operating-performance result mechanically to observed R&D-backed innovation. StrategyAI may capture managerial attention, operating initiatives, product-market positioning, implementation intensity, or organizational change. These mechanisms may be related to innovation activity, but they are not necessarily captured by reported R&D.

The appropriate conclusion is therefore narrow. The R&D interaction does not show that R&D backing strengthens the StrategyAI-performance relationship in the expected positive direction. The H2 operating-performance evidence remains the central result, while H3 shows that this evidence should not be interpreted as being clearly amplified by the specific observed R&D stock proxy used in the thesis. In this sense, R&D is a boundary condition rather than the main result: H3 tests whether one observable investment proxy strengthens the StrategyAI-performance relation, not whether AI capability of any kind underlies the disclosure.

5.4. Controls, Fixed Effects, and Credibility

The credibility of the empirical design comes from the alignment between the research question, the measurement structure, and the interpretation of the estimates. The regressions include controls for size, leverage, cash ratio, and sales growth to adjust for time-varying differences in firm scale, financing structure, liquidity, and operating momentum. Firm fixed effects absorb persistent firm characteristics, including business model, baseline disclosure culture, industry position, and other time-invariant differences. Year fixed effects absorb common macroeconomic, reporting-period, and AI-salience shocks.

The AI coefficients are therefore interpreted as conditional within-firm, year-adjusted associations rather than as simple cross-sectional differences between firms. The use of standard errors clustered at the cleaned firm-panel identifier level allows residual dependence within firms over time, consistent with finance and accounting panel-inference concerns (Petersen, 2009; Gow et al., 2010).

This structure does not make the design causal. AI disclosure may still be correlated with unobserved implementation quality, managerial capability, investor attention, competitive position, or earlier public information. The contribution is therefore narrower. The thesis tests whether section-specific AI disclosure remains associated with market reactions and future operating outcomes after observable controls and fixed effects. This is a disclosure informativeness test, not an experimental or causal design.

The interpretation is therefore aligned with the research design. The results are not presented as causal effects, and the method does not claim to identify randomized AI

implementation. Instead, the thesis asks whether section-specific 10-K AI disclosure is informative under a disciplined observational design. This level of inference is appropriate because the researcher observes disclosure, returns, and accounting outcomes, but not randomized AI adoption or project-level implementation.

Appendices A.2, A.3, and A.5 provide the coefficient-with-controls evidence, while Appendix A.4 reports the H2 persistence-control robustness diagnostics. Appendix B.1-B.4 record the text and data construction logic, and Appendix C.1-C.4 report the dictionary and validation protocol. The full coefficient tables report controls, standard errors, adjusted R^2, fixed effects, and cluster counts. The reproducibility appendix records the text download, section recovery, dictionary merge, and CRSP/Compustat Merged event-link logic. The dictionary appendix reports the AI terms, exclusion rules, ratio construction, and manual validation protocol. Together, these materials show how the thesis moves from raw filing text to empirical variables and regression results.

5.5. Limitations and Robustness Boundary

Because the firm universe is based on a large-cap constituent snapshot rather than historical index membership, the results should be interpreted as evidence from a retrospective large-cap issuer panel rather than as evidence for all public firms.

Second, the H1 market reaction design uses daily filing-window returns, rather than intraday returns around SEC filing acceptance timestamps. As a result, the H1 estimates should be interpreted as filing-date window associations rather than precise intraday announcement effects. Longer post-filing windows are especially exposed to non-filing information events, which is why the [0,+60] and [0,+126] windows are treated as diagnostic rather than as clean filing-news tests.

Third, the text measure captures explicit AI-related disclosure rather than latent internal AI capability. Firms may use AI through vendors, cloud services, acquired software, process redesign, or other channels without describing those activities in the 10-K, so the findings should be interpreted as evidence about disclosed AI language rather than comprehensive AI adoption.

Fourth, strict R&D treatment improves interpretation but reduces the H3 sample size and statistical power. Strict weighted RDStockTA requires observed current R&D, four observed lagged R&D inputs, and positive disclosure-year total assets. This conservative treatment avoids treating missing R&D as zero, but it also excludes firms that may have AI-related investment through software, vendors, cloud services, acquisitions, or process redesign. H3 therefore tests one observed investment-backing proxy, not all possible forms of AI capability.

Fifth, financial firms, insurers, and real-estate firms remain in the main cross-sector sample. Revenue-scaled operating ratios may not have identical economic meaning

across these sectors. The H2 and H3 findings should therefore be interpreted as large-cap cross-sector associations rather than sector-specific production-function estimates.

The interpretation also faces a multiple-testing and coefficient-pattern boundary. H1 is estimated across several return windows, while H2 and H3 are estimated across several operating outcomes. No formal family-wise multiple-testing correction is applied. The thesis therefore emphasizes direction, economic interpretation, and stability across word and sentence specifications rather than isolated significance. In addition, the H2 section-specific conclusion is based on coefficient patterns rather than formal tests of equality across StrategyAI, RiskAI, and AccountingNoteAI coefficients.

Inference is clustered at the firm level; because the panel covers only eight fiscal years, two-way clustering is not used as the main specification, so cross-sectional dependence remains a robustness boundary.

These limitations are addressed by reporting model-level observation counts, separate word and sentence specifications, and the supporting appendix diagnostics. The thesis reports observation counts by model, distinguishes word and sentence specifications, keeps R&D backing separate from broad H2 outcomes, and reports supplementary ROIC evidence in Appendix A.7 and reproducibility/dictionary diagnostics in Appendices B and C. These choices keep the conclusion conservative. StrategyAI disclosure is conditionally informative about future operating performance, but the evidence does not justify a broad positive market-reaction or strong R&D-amplification claim.

Finally, the accounting-ratio results may be affected by extreme observations. Fiscal-year-specific winsorization reduces the influence of extreme ratio values, but it does not replace robust-regression, no-winsorization, or influence-diagnostic checks. The reported estimates should therefore be read with this winsorization choice in mind.

Future research could extend the design in three directions. First, it could combine 10-K text with earnings-call transcripts, press releases, and product announcements to observe the full disclosure sequence. Second, future studies could use intraday return data around SEC filing acceptance timestamps to sharpen the H1 event window. Third, it could use project-level data, patent data, job postings, cloud-service expenditures, or software capitalization disclosures to measure AI capability more directly than observed R&D permits.

A further extension would be to examine how investors respond to the specificity of AI disclosure. The current dictionary measures explicit AI intensity, but it does not separately score whether the firm describes a product, a process, a risk, a governance framework, a customer use case, or a financial investment. More granular classification could distinguish operationally specific AI disclosure from broad strategic language. That would help separate signal from shield more directly.

6. Conclusion

This thesis examines whether AI-related disclosure in 10-K filings from a retrospective early-2026 large-cap issuer snapshot is more consistent with an informative signal, a strategic shield, or a combination of both. The evidence is mixed. AI disclosure is not uniformly associated with favorable filing-window market reactions, but strategy-oriented AI disclosure is associated with future operating-performance outcomes.

The filing-window evidence does not suggest that the market simply rewards AI disclosure. Rather, it shows that AI language is not one uniform signal, and that market associations depend on the section of the 10-K in which the language appears.

The operating-performance evidence is the clearest support for disclosure-informativeness interpretation. StrategyAI is associated with higher one-year-ahead operating margins, especially operating income to revenue. The word-intensity specification also shows positive net-income evidence and negative expense-intensity evidence, although the sentence-intensity results are weaker. RiskAI and AccountingNoteAI do not show similarly stable operating-performance patterns. This pattern supports the section-specific design and suggests that AI language carries the most information when it appears in business and management discussion sections.

The strict observed R&D backing test does not support the expected moderation mechanism. Strict weighted RDStockTA does not amplify the StrategyAI-performance relation. The operating-performance result therefore cannot be attributed to observed R&D backing. AI-related capability may instead be built through other channels, such as software, cloud services, vendor systems, acquisitions, data infrastructure, process redesign, or human capital investments.

The strongest finding is that strategy-oriented AI disclosure is associated with higher one-year-ahead operating-performance outcomes in a way consistent with operating-content informativeness, but the evidence does not support a broadly positive filing-date market reaction or a robust observed R&D moderation mechanism.

7. References

Aboody, D., & Lev, B. (2000). Information asymmetry, R&D, and insider gains. *The Journal of Finance*, *55*(6), 2747-2766. https://doi.org/10.1111/0022-1082.00305

Akyol, A. C., & Shekhar, C. (2026). Information environment, proprietary costs and disclosure content in firm 10-K reports. *International Journal of Managerial Finance*, *22*(1), 149-175. https://doi.org/10.1108/ijmf-12-2024-0683

Ante, L., & Saggu, A. (2025). Quantifying a firm's AI engagement: Constructing objective, data-driven, AI stock indices using 10-K filings. *Technological Forecasting and Social Change*, *212*, 123965. https://doi.org/10.1016/j.techfore.2024.123965

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, *6*(2), 159-178. https://doi.org/10.2307/2490232

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of Accounting Research*, *6*, 67-92. https://doi.org/10.2307/2490070

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, *50*(2-3), 296-343. https://doi.org/10.1016/j.jacceco.2010.10.003

Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2-3), 329-357. https://doi.org/10.1016/j.jacceco.2017.03.002

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns. *Journal of Financial Economics*, *14*(1), 3-31. https://doi.org/10.1016/0304-405x(85)90042-x

Bushee, B. J., Core, J. E., Guay, W., & Hamm, S. J. (2010). The role of the business press as an information intermediary. *Journal of Accounting Research*, *48*(1), 1-19. https://doi.org/10.1111/j.1475-679x.2009.00357.x

Cao, S., Jiang, W., Yang, B., & Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of AI. *Review of Financial Studies*, *36*(9), 3603-3642. https://doi.org/10.1093/rfs/hhad021

Center for Research in Security Prices. (2026, May 8). *CRSP US Stock Databases - Center for Research in Security Prices*. https://www.crsp.org/research/crsp-us-stock-databases/

Chan, L. K. C., Lakonishok, J., & Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *The Journal of Finance*, *56*(6), 2431-2456. https://doi.org/10.1111/0022-1082.00411

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from latent Dirichlet allocation. *Journal of Accounting and Economics*, *64*(2-3), 221-245. https://doi.org/10.1016/j.jacceco.2017.07.002

Einhorn, E. (2005). The nature of the interaction between mandatory and voluntary disclosures. *Journal of Accounting Research*, *43*(4), 593-621. https://doi.org/10.1111/j.1475-679x.2005.00183.x

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*(1), 1-21. https://doi.org/10.2307/2525569

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/10.2307/2329112

Frankel, R., Johnson, M., & Skinner, D. J. (1999). An empirical examination of conference calls as a voluntary disclosure medium. *Journal of Accounting Research*, *37*(1), 133-150. https://doi.org/10.2307/2491400

Gow, I. D., Ormazabal, G., & Taylor, D. J. (2010). Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review*, *85*(2), 483-512. https://doi.org/10.2308/accr.2010.85.2.483

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1-3), 405-440. https://doi.org/10.1016/s0165-4101(01)00018-0

Kimbrough, M. D. (2005). The effect of conference calls on analyst and market underreaction to earnings announcements. *The Accounting Review*, *80*(1), 189-219. https://doi.org/10.2308/accr.2005.80.1.189

Koh, P., & Reeb, D. M. (2015). Missing R&D. *Journal of Accounting and Economics*, *60*(1), 73-94. https://doi.org/10.1016/j.jacceco.2015.03.004

Kothari, S. P., Laguerre, T. E., & Leone, A. J. (2002). Capitalization versus expensing: Evidence on the uncertainty of future earnings from capital expenditures versus R&D outlays. *Review of Accounting Studies*, *7*(4), 355-382. https://doi.org/10.1023/a:1020764227390

Leone, A. J., Minutti-Meza, M., & Wasley, C. E. (2019). Influential observations and inference in accounting research. *The Accounting Review*, *94*(6), 337-364. https://doi.org/10.2308/accr-52396

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. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2-3), 221-247. https://doi.org/10.1016/j.jacceco.2008.02.003

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

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. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, *69*(4), 1643-1671. https://doi.org/10.1111/jofi.12162

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1230. https://doi.org/10.1111/1475-679x.12123

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature, 35*(1), 13-39. http://www.jstor.org/stable/2729691

Merkley, K. J. (2014). Narrative disclosure and earnings performance: Evidence from R&D disclosures. *The Accounting Review*, *89*(2), 725-757. https://doi.org/10.2308/accr-50649

Opler, T., Pinkowitz, L., Stulz, R., & Williamson, R. (1999). The determinants and implications of corporate cash holdings. *Journal of Financial Economics*, *52*(1), 3-46. https://doi.org/10.1016/s0304-405x(99)00003-3

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, *22*(1), 435-480. https://doi.org/10.1093/rfs/hhn053

Rajan, R. G., & Zingales, L. (1995). What do we know about capital structure? Some evidence from international data. *The Journal of Finance*, *50*(5), 1421-1460. https://doi.org/10.1111/j.1540-6261.1995.tb05184.x

S&P Dow Jones Indices. (2026). S&P 500®. S&P Global. https://www.spglobal.com/spdji/en/indices/equity/sp-500/

S&P Global Market Intelligence. (2026). S&P Capital IQ Pro. https://www.spglobal.com/market-intelligence/en/solutions/products/sp-capital-iq-pro

U.S. Securities and Exchange Commission. (n.d.). Search filings. Retrieved May 17, 2026, from https://www.sec.gov/search-filings

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics, 5*, 179-194. https://doi.org/10.1016/0165-4101(83)90011-3

Wharton Research Data Services. (2026). Linking CRSP with Compustat. https://wrds-www.wharton.upenn.edu/pages/wrds-research/database-linking-matrix/linking-crsp-with-compustat/

8. Appendix

Appendix A. Detailed Empirical Evidence

Appendix A.1 reports supplementary H1 event-study estimates, Appendices A.2, A.3, and A.5 report the coefficient-with-controls tables, Appendix A.4 reports H2 persistence-control robustness diagnostics, Appendix A.6 reports the correlation matrix, and Appendix A.7 reports ROIC supplementary evidence.

Appendix A.1 H1 Event-Study Regressions

Table 10. Full H1 event-study regressions

Spec	Window	Dependent variable	Strategy AI	Risk AI	AccountingNoteAI	N	Clusters
Word	[-1,+1]	Cumulative raw return	-0.0032	0.0017	0.0003	3301	431
Word	[-1,+1]	Market-adjusted CAR	-0.0036*	0.0016	-0.0004	3301	431
Word	[0,+5]	Cumulative raw return	-0.0048	0.0032*	0.0027	3376	431
Word	[0,+5]	Market-adjusted CAR	-0.0072**	0.0033**	0.0019	3376	431
Word	[0,+20]	Cumulative raw return	-0.0033	-0.0005	0.0048	3376	431
Word	[0,+20]	Market-adjusted CAR	-0.0063	-0.0011	0.0022	3376	431
Word	[0,+60]	Cumulative raw return	-0.0016	0.0001	0.0006	3374	431
Word	[0,+60]	Market-adjusted CAR	-0.0024	-0.0008	-0.0006	3374	431
Word	[0,+126]	Cumulative raw return	0.0024	-0.0066	0.0052	3373	431
Word	[0,+126]	Market-adjusted CAR	0.0012	-0.0062	0.0049	3373	431
Sentence	[-1,+1]	Cumulative raw return	-0.0034*	0.0003	0.0015	3301	431
Sentence	[-1,+1]	Market-adjusted CAR	-0.0036**	0.0003	0.0011	3301	431

Spec	Window	Dependent variable	Strategy AI	Risk AI	AccountingNoteAI	N	Clusters
Sentence	[0,+5]	Cumulative raw return	-0.0041	0.0009	0.0022	3376	431
Sentence	[0,+5]	Market-adjusted CAR	-0.0063**	0.0008	0.0019	3376	431
Sentence	[0,+20]	Cumulative raw return	-0.0052	0.0009	0.0076*	3376	431
Sentence	[0,+20]	Market-adjusted CAR	-0.0075**	-0.0002	0.0059	3376	431
Sentence	[0,+60]	Cumulative raw return	-0.0070	0.0016	0.0060	3374	431
Sentence	[0,+60]	Market-adjusted CAR	-0.0075	0.0007	0.0046	3374	431
Sentence	[0,+126]	Cumulative raw return	-0.0066	-0.0019	0.0158**	3373	431
Sentence	[0,+126]	Market-adjusted CAR	-0.0071	-0.0017	0.0144*	3373	431

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes.

Appendix A.2 H1 Coefficients With Controls

Cells report coefficient with clustered standard error in parentheses. All specifications include firm and year fixed effects.

Table 11. H1 word-intensity coefficients with controls

Variable	CAR[-1,+1]	CAR[0,+5]	CAR[0,+20]	CAR[0,+60]	CAR[0,+126]
Strategy AI	-0.0036* (0.0018)	-0.0072** (0.0031)	-0.0063 (0.0042)	-0.0024 (0.0065)	0.0012 (0.0103)
Risk AI	0.0016 (0.0013)	0.0033** (0.0016)	-0.0011 (0.0023)	-0.0008 (0.0036)	-0.0062 (0.0064)
AccountingNoteAI	-0.0004 (0.0019)	0.0019 (0.0028)	0.0022 (0.0038)	-0.0006 (0.0037)	0.0049 (0.0041)
Size	-0.0053 (0.0034)	-0.0061 (0.0070)	-0.0153 (0.0116)	-0.0533*** (0.0145)	-0.1215*** (0.0209)
Leverage	-0.0041 (0.0125)	0.0189 (0.0181)	0.0353 (0.0294)	0.0237 (0.0409)	0.0708 (0.0557)

Variable	CAR[-1,+1]	CAR[0,+5]	CAR[0,+20]	CAR[0,+60]	CAR[0,+126]
Cash ratio	-0.0331 (0.0220)	-0.0221 (0.0232)	0.0434 (0.0401)	-0.0061 (0.0625)	-0.0424 (0.0887)
Sales growth	0.0088** (0.0042)	-0.0046 (0.0066)	0.0015 (0.0122)	0.0206 (0.0162)	0.0354 (0.0257)
Observations	3301	3376	3376	3374	3373
Adjusted R2	0.033	0.022	0.063	0.079	0.115

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes.

Table 12. H1 sentence-intensity coefficients with controls

Variable	CAR[-1,+1]	CAR[0,+5]	CAR[0,+20]	CAR[0,+60]	CAR[0,+126]
Strategy AI	-0.0036** (0.0017)	-0.0063** (0.0027)	-0.0075** (0.0038)	-0.0075 (0.0059)	-0.0071 (0.0096)
Risk AI	0.0003 (0.0004)	0.0008 (0.0008)	-0.0002 (0.0007)	0.0007 (0.0010)	-0.0017 (0.0020)
AccountingNoteAI	0.0011 (0.0020)	0.0019 (0.0020)	0.0059 (0.0044)	0.0046 (0.0058)	0.0144* (0.0084)
Size	-0.0048 (0.0034)	-0.0055 (0.0071)	-0.0147 (0.0116)	-0.0517*** (0.0147)	-0.1197*** (0.0208)
Leverage	-0.0041 (0.0128)	0.0196 (0.0182)	0.0362 (0.0293)	0.0228 (0.0409)	0.0696 (0.0549)
Cash ratio	-0.0325 (0.0219)	-0.0210 (0.0232)	0.0423 (0.0402)	-0.0092 (0.0624)	-0.0473 (0.0890)
Sales growth	0.0089** (0.0042)	-0.0042 (0.0066)	0.0010 (0.0121)	0.0200 (0.0160)	0.0334 (0.0252)
Observations	3301	3376	3376	3374	3373
Adjusted R2	0.033	0.021	0.064	0.080	0.116

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes.

Appendix A.3 H2 Coefficients With Controls

Controls are measured in year t. Dependent variables are measured in year t+1. Standard errors are clustered by firm.

Table 13. H2 word-intensity coefficients with controls

Variable	OI/revenue	NI/revenue	SG&A/revenue	Other opex/revenue
Strategy AI	0.0077** (0.0032)	0.0104** (0.0043)	-0.0041* (0.0023)	-0.0064** (0.0029)
Risk AI	-0.0005 (0.0019)	-0.0028 (0.0025)	-0.0013 (0.0009)	0.0004 (0.0014)
AccountingNoteAI	0.0036 (0.0032)	0.0023 (0.0029)	-0.0004 (0.0009)	-0.0010 (0.0018)
Size	-0.0258*** (0.0099)	-0.0429*** (0.0132)	-0.0118 (0.0145)	0.0020 (0.0186)
Leverage	0.0351 (0.0322)	0.0468 (0.0335)	-0.0022 (0.0161)	0.0011 (0.0234)
Cash ratio	0.0478 (0.0356)	0.0494 (0.0452)	0.0194 (0.0206)	-0.0089 (0.0321)
Sales growth	0.0572*** (0.0171)	0.0583*** (0.0167)	-0.0247*** (0.0094)	-0.0455*** (0.0124)
Observations	3497	3497	3406	3496
Adjusted R2	0.746	0.510	0.955	0.914

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes.

Table 14. H2 sentence-intensity coefficients with controls

Variable	OI/revenue	NI/revenue	SG&A/revenue	Other opex/revenue
Strategy AI	0.0062* (0.0037)	0.0062 (0.0048)	-0.0042* (0.0023)	-0.0051 (0.0033)
Risk AI	0.0006 (0.0006)	0.0007 (0.0008)	0.0002 (0.0004)	0.0006 (0.0004)
AccountingNoteAI	0.0026 (0.0039)	0.0026 (0.0032)	-0.0001 (0.0006)	-0.0014 (0.0021)
Size	-0.0259*** (0.0099)	-0.0428*** (0.0133)	-0.0117 (0.0146)	0.0021 (0.0187)
Leverage	0.0344 (0.0322)	0.0451 (0.0334)	-0.0020 (0.0162)	0.0018 (0.0234)
Cash ratio	0.0483 (0.0356)	0.0478 (0.0455)	0.0183 (0.0206)	-0.0091 (0.0322)
Sales growth	0.0569*** (0.0170)	0.0576*** (0.0167)	-0.0249*** (0.0094)	-0.0454*** (0.0124)
Observations	3497	3497	3406	3496
Adjusted R2	0.745	0.510	0.955	0.914

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes.

Appendix A.4 H2 Persistence-Control Robustness

This appendix reports persistence-control diagnostics for H2. The baseline H2 specification is re-estimated after adding the current value of the same operating ratio, its one-year lagged value, and both controls together. StrategyAI, RiskAI, and AccountingNoteAI are included jointly in all specifications, so the tests examine whether the StrategyAI pattern remains after accounting for persistent operating performance.

Table 15 reports the word-intensity version. StrategyAI remains directionally consistent across all four outcomes, with the strongest support for operating income/revenue, net income/revenue, and other operating expenses/revenue. RiskAI and AccountingNoteAI do not show similarly stable patterns.

Table 15. H2 persistence-control robustness, word-intensity specification

Outcome t+1	Specification	StrategyAI	RiskAI	AccountingNoteAI	N	Clusters
Operating income / revenue	Baseline	0.0077** (0.0032)	-0.0005 (0.0019)	0.0036 (0.0032)	3497	439
Operating income / revenue	Current Y(t)	0.0064** (0.0025)	-0.0003 (0.0016)	0.0029 (0.0026)	3497	439
Operating income / revenue	Lagged Y(t-1)	0.0077** (0.0032)	-0.0005 (0.0020)	0.0036 (0.0032)	3497	439
Operating income / revenue	Current Y(t) + lagged Y(t-1)	0.0066** (0.0026)	-0.0006 (0.0017)	0.0030 (0.0026)	3497	439
Net income / revenue	Baseline	0.0104** (0.0043)	-0.0028 (0.0025)	0.0023 (0.0029)	3497	439
Net income / revenue	Current Y(t)	0.0094** (0.0042)	-0.0026 (0.0024)	0.0024 (0.0028)	3497	439
Net income / revenue	Lagged Y(t-1)	0.0110** (0.0045)	-0.0033 (0.0026)	0.0024 (0.0030)	3497	439
Net income / revenue	Current Y(t) + lagged Y(t-1)	0.0100** (0.0043)	-0.0031 (0.0024)	0.0025 (0.0028)	3497	439
SG&A / revenue	Baseline	-0.0041* (0.0023)	-0.0013 (0.0009)	-0.0004 (0.0009)	3406	429
SG&A / revenue	Current Y(t)	-0.0020 (0.0015)	-0.0009 (0.0005)	-0.0007 (0.0006)	3401	429
SG&A / revenue	Lagged Y(t-1)	-0.0034* (0.0021)	-0.0013* (0.0008)	-0.0006 (0.0008)	3394	429
SG&A / revenue	Current Y(t) + lagged Y(t-1)	-0.0022 (0.0015)	-0.0008 (0.0005)	-0.0006 (0.0006)	3393	429

Outcome t+1	Specification	StrategyAI	RiskAI	AccountingNoteAI	N	Clusters
Other operating expenses / revenue	Baseline	-0.0064** (0.0029)	0.0004 (0.0014)	-0.0010 (0.0018)	3496	439
Other operating expenses / revenue	Current Y(t)	-0.0049** (0.0022)	0.0002 (0.0011)	-0.0010 (0.0012)	3495	439
Other operating expenses / revenue	Lagged Y(t-1)	-0.0062** (0.0028)	0.0003 (0.0014)	-0.0010 (0.0017)	3494	439
Other operating expenses / revenue	Current Y(t) + lagged Y(t-1)	-0.0050** (0.0023)	0.0004 (0.0011)	-0.0009 (0.0013)	3494	439

Note: Cells report coefficients with firm-clustered standard errors in parentheses. All specifications include firm fixed effects, year fixed effects, and the baseline controls: log assets, leverage, cash ratio, and sales growth. Current Y(t) denotes the current value of the same operating ratio used as the t+1 dependent variable. Lagged Y(t-1) denotes that operating ratio lagged one fiscal year. StrategyAI, RiskAI, and AccountingNoteAI are included jointly in every model. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table 16 reports the sentence-intensity version. StrategyAI signs remain directionally consistent, but statistical significance is weaker than in the word-intensity specification. This supports the interpretation that the H2 evidence is concentrated in StrategyAI, while remaining diagnostic rather than causal.

Table 16. H2 persistence-control robustness, sentence-intensity specification

Outcome t+1	Specification	StrategyAI	RiskAI	AccountingNoteAI	N	Clusters
Operating income / revenue	Baseline	0.0062* (0.0037)	0.0006 (0.0006)	0.0026 (0.0039)	3497	439
Operating income / revenue	Current Y(t)	0.0056** (0.0028)	0.0005 (0.0005)	0.0018 (0.0030)	3497	439
Operating income / revenue	Lagged Y(t-1)	0.0062* (0.0037)	0.0006 (0.0006)	0.0026 (0.0039)	3497	439
Operating income / revenue	Current Y(t) + lagged Y(t-1)	0.0055* (0.0029)	0.0005 (0.0005)	0.0020 (0.0031)	3497	439
Net income / revenue	Baseline	0.0062 (0.0048)	0.0007 (0.0008)	0.0026 (0.0032)	3497	439
Net income / revenue	Current Y(t)	0.0057 (0.0046)	0.0008 (0.0008)	0.0025 (0.0031)	3497	439

Outcome t+1	Specification	StrategyAI	RiskAI	AccountingNoteAI	N	Clusters
Net income / revenue	Lagged Y(t-1)	0.0064 (0.0049)	0.0006 (0.0010)	0.0029 (0.0033)	3497	439
Net income / revenue	Current Y(t) + lagged Y(t-1)	0.0058 (0.0047)	0.0006 (0.0009)	0.0029 (0.0031)	3497	439
SG&A / revenue	Baseline	-0.0042* (0.0023)	0.0002 (0.0004)	-0.0001 (0.0006)	3406	429
SG&A / revenue	Current Y(t)	-0.0022 (0.0015)	0.0001 (0.0002)	-0.0003 (0.0005)	3401	429
SG&A / revenue	Lagged Y(t-1)	-0.0036* (0.0021)	0.0001 (0.0003)	-0.0002 (0.0006)	3394	429
SG&A / revenue	Current Y(t) + lagged Y(t-1)	-0.0023 (0.0016)	0.0001 (0.0002)	-0.0003 (0.0005)	3393	429
Other operating expenses / revenue	Baseline	-0.0051 (0.0033)	0.0006 (0.0004)	-0.0014 (0.0021)	3496	439
Other operating expenses / revenue	Current Y(t)	-0.0042* (0.0024)	0.0003 (0.0003)	-0.0012 (0.0016)	3495	439
Other operating expenses / revenue	Lagged Y(t-1)	-0.0051 (0.0032)	0.0005 (0.0004)	-0.0014 (0.0021)	3494	439
Other operating expenses / revenue	Current Y(t) + lagged Y(t-1)	-0.0042* (0.0025)	0.0004 (0.0003)	-0.0011 (0.0016)	3494	439

Note: Cells report coefficients with firm-clustered standard errors in parentheses. All specifications include firm fixed effects, year fixed effects, and the baseline controls: log assets, leverage, cash ratio, and sales growth. Current Y(t) denotes the current value of the same operating ratio used as the t+1 dependent variable. Lagged Y(t-1) denotes that operating ratio lagged one fiscal year. StrategyAI, RiskAI, and AccountingNoteAI are included jointly in every model. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Appendix A.5 H3 Strict R&D Coefficients With Controls

The table uses strict observed R&D firm-years and strict weighted RDStockTA. Missing R&D is not coded as zero.

Table 17. H3 word-intensity coefficients with strict R&D controls

Variable	OI/revenue	NI/revenue	SG&A/revenue	Other opex/revenue
Strategy AI	0.0083* (0.0043)	0.0202*** (0.0066)	-0.0043* (0.0023)	-0.0077** (0.0035)
Risk AI	0.0023 (0.0028)	-0.0019 (0.0039)	-0.0026** (0.0013)	-0.0012 (0.0018)
AccountingNoteAI	0.0028 (0.0030)	0.0004 (0.0032)	-0.0005 (0.0008)	-0.0012 (0.0017)
Strict weighted RDStockTA	0.0152 (0.0131)	0.0278 (0.0180)	-0.0063 (0.0070)	-0.0121 (0.0115)
Strategy AI x strict weighted RDStockTA	-0.0010 (0.0026)	-0.0063* (0.0036)	-0.0005 (0.0012)	0.0005 (0.0022)
Size	-0.0108 (0.0179)	-0.0429 (0.0288)	-0.0029 (0.0111)	0.0077 (0.0144)
Leverage	-0.0232 (0.0452)	-0.0099 (0.0562)	-0.0010 (0.0218)	0.0254 (0.0321)
Cash ratio	0.0080 (0.0451)	0.0390 (0.0644)	0.0061 (0.0281)	-0.0267 (0.0391)
Sales growth	0.0917*** (0.0164)	0.0708*** (0.0191)	-0.0290*** (0.0084)	-0.0597*** (0.0118)
Observations	1362	1362	1362	1362
Adjusted R2	0.798	0.521	0.955	0.950

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes. N is smaller than the strict weighted RDStockTA coverage count as regression observations also require non-missing outcomes, controls, and fixed-effect identifiers.

Table 18. H3 sentence-intensity coefficients with strict R&D controls

Variable	OI/revenue	NI/revenue	SG&A/revenue	Other opex/revenue
Strategy AI	0.0079* (0.0046)	0.0147** (0.0067)	-0.0050** (0.0021)	-0.0073** (0.0036)
Risk AI	0.0011* (0.0006)	0.0015** (0.0006)	-0.0000 (0.0004)	0.0001 (0.0004)
AccountingNoteAI	0.0020 (0.0033)	0.0009 (0.0031)	-0.0005 (0.0006)	-0.0018 (0.0019)
Strict weighted RDStockTA	0.0165 (0.0132)	0.0278 (0.0181)	-0.0075 (0.0070)	-0.0131 (0.0116)

Variable	OI/revenue	NI/revenue	SG&A/revenue	Other opex/revenue
Strategy AI x strict weighted RDStockTA	-0.0004 (0.0027)	-0.0046 (0.0035)	-0.0007 (0.0013)	-0.0001 (0.0022)
Size	-0.0101 (0.0184)	-0.0428 (0.0294)	-0.0035 (0.0113)	0.0076 (0.0143)
Leverage	-0.0222 (0.0462)	-0.0148 (0.0555)	-0.0026 (0.0226)	0.0249 (0.0327)
Cash ratio	0.0136 (0.0457)	0.0378 (0.0651)	0.0010 (0.0285)	-0.0300 (0.0394)
Sales growth	0.0919*** (0.0162)	0.0709*** (0.0191)	-0.0294*** (0.0085)	-0.0597*** (0.0119)
Observations	1362	1362	1362	1362
Adjusted R2	0.797	0.520	0.955	0.950

Note: *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered as indicated in the model notes. N is smaller than the strict weighted RDStockTA coverage count as regression observations also require non-missing outcomes, controls, and fixed-effect identifiers.

Appendix A.6 Correlation matrix

Table 19. Correlation matrix. (Values in %)

	Strat-W	Risk-W	Note-W	Strat-S	Risk-S	Note-S	OI/Rev	NI/Rev	SG&A/Rev	OOpEx/Rev	Size	Lev.	Cash	SalesGr.	RDStockTA
Strat-W	100.0														
Risk-W	43.0	100.0													
Note-W	70.7	22.5	100.0												
Strat-S	97.7	42.2	70.1	100.0											
Risk-S	25.7	51.2	13.9	26.4	100.0										
Note-S	67.2	18.4	92.5	66.2	11.5	100.0									
OI/Rev	5.1	7.5	4.5	4.3	4.0	4.7	100.0								
NI/Rev	9.6	8.4	6.0	8.5	4.5	5.8	80.1	100.0							
SG&A/Rev	4.2	10.7	1.1	3.3	4.8	-0.1	11.6	10.0	100.0						
OOpEx/Rev	10.4	9.2	4.4	9.4	3.8	2.7	10.7	6.5	77.0	100.0					
Size	4.3	8.6	6.3	6.0	5.3	5.1	8.8	3.2	10.4	11.7	100.0				
Lev.	-8.2	-4.4	-1.8	-8.7	-2.5	-1.2	3.4	-10.6	-15.7	-2.7	-12.0	100.0			
Cash	10.2	3.0	5.0	8.9	0.7	4.0	0.5	5.5	19.9	12.9	-31.0	-14.1	100.0		
SalesGr.	5.8	-2.7	5.3	5.4	-1.9	5.6	15.0	17.1	-1.7	-0.2	-6.5	-8.5	7.1	100.0	
RDStockTA	29.8	11.4	13.9	26.7	3.7	11.7	14.6	19.1	17.5	42.8	-16.3	-18.3	47.2	12.2	100.0

Note: This table reports pairwise Pearson correlations among word-intensity AI measures (W), sentence-intensity AI measures (S), main operating outcomes, controls, and strict observed R&D stock scaled by assets. Values are multiplied by 100. Pairwise observations vary by variable availability, especially for RDStockTA. Variable labels: Strat-W = StrategyAI word-intensity ratio; Risk-W = RiskAI word-intensity ratio; Note-W = AccountingNoteAI word-intensity ratio; Strat-S = StrategyAI sentence-intensity ratio; Risk-S = RiskAI sentence-intensity ratio; Note-S = AccountingNoteAI sentence-intensity ratio; OI/Rev = operating income divided by revenue; NI/Rev = net income divided by revenue; SG&A/Rev = selling, general, and administrative expenses divided by revenue; OOpEx/Rev = other operating expenses divided by revenue; Size = log total assets; Lev. = total debt divided by total assets; Cash = cash and equivalents divided by total assets; SalesGr. = sales growth; RDStockTA = strict observed weighted R&D stock divided by total assets.

Appendix A.7 ROIC Supplementary Evidence

ROIC is retained as supplementary evidence only. It is not used in the main H2/H3 interpretation.

Table 20. ROIC supplementary evidence

Spec	Outcome	Strategy AI	N	Clusters	Scope
Word	ROIC t+1	-0.0197 (p=0.351)	3166	435	Appendix only
Sentence	ROIC t+1	-0.0269 (p=0.512)	3166	435	Appendix only

Appendix B. Technical Reproducibility Notes

Appendix B records the main reproducibility logic used to construct the SEC EDGAR text corpus, merge the accounting panel, estimate the CRSP event-study results, and build the thesis evidence. Capital IQ Pro data were downloaded manually from the browser interface. The underlying Capital IQ Pro accounting downloads were completed during March 1-3, 2026, and the cleaned accounting panel was regenerated after those files were available. Fiscal-year 2025 observations are used only to construct t+1 outcomes for fiscal-year 2024 disclosure observations where available. The reproducible local steps begin after the downloaded files are available on disk. Paths below are repository-relative and avoid machine-specific user folders.

Appendix B.1 EDGAR 10-K Full-Text Download Logic

The 10-K download stage uses the local EDGARtools downloader. It discovers the ticker universe from the existing section corpus, queries SEC EDGAR filings through edgartools, selects annual 10-K filings by fiscal year using period_of_report, and writes full filing text plus manifest/error logs. The local project folder for the EDGAR text corpus is named EDGAR in the repository.

An independent manifest audit found one study-window observation labeled 10-K/A: DOV fiscal year 2017. That row entered the pre-exclusion local manifest before the final filing-type audit. The main H1, H2, and H3 estimation samples exclude it, and the code excerpt below shows the stricter original-10-K selection rule used for reproducible reruns.

Core selection and write logic:

```
def select_annual_filings(filings, *, start_year, end_year):
    selected = {}
    for filing in filings:
        if str(getattr(filing, “form”, ““)).upper() != “10-K”:
            continue
        period_end = parse_iso_date(getattr(filing, “period_of_report”, None))
        filing_dt = parse_iso_date(getattr(filing, “filing_date”, None))
        if not period_end or not filing_dt:
            continue
        fiscal_year = period_end.year
        if fiscal_year < start_year or fiscal_year > end_year:
            continue
        current = selected.get(fiscal_year)
        if current is None or filing_dt > parse_iso_date(getattr(current, “filing_date”, None)):
            selected[fiscal_year] = filing
    return selected


def write_filing_text(output_file, filing):
```

```
    output_file.parent.mkdir(parents=True, exist_ok=True)
    text = filing.text()
    output_file.write_text(text, encoding="utf-8")
    return len(text)
```

Repository-relative run command:

```
export EDGAR_IDENTITY="<name email for SEC fair-access identity>"
./.venv/bin/python -u -m EDGARtools.downloader \
  --output-root EDGARtools \
  --universe-root EDGAR/footnote_sentence \
  --start-year 2015 \
  --end-year 2024 \
  --sleep-seconds 0
```

The downloader writes EDGARtools/manifest.csv, EDGARtools/errors.csv, and filing text under EDGARtools/filing_text/<TICKER>/<year>__<cik>__<period_end>.txt.

Appendix B.2 Section Recovery and Sentence Formatting Logic

When a firm-year filing required recovery from full filing text, the local backfill script rebuilt Item 1, Item 1A, MD&A, and footnote sentence files using the same sentence formatting routine used elsewhere in the EDGAR corpus.

Core extraction logic:

```
def extract_sections_from_filing_text(text):
    item1_start = first_heading_start(text, ITEM_1_RE)
    item1a_start = first_heading_start(text, ITEM_1A_RE, start=item1_start or 0)
    item1a_end = first_heading_start(text, ITEM_1A_END_RE, start=item1a_start or 0)
    item7_start = first_heading_start(text, ITEM_7_RE, start=item1a_start or 0)
    item8_start = first_heading_start(text, ITEM_8_RE, start=item7_start or 0)
    item9_start = first_heading_start(text, ITEM_9_RE, start=item8_start or 0)
    notes_start = first_heading_start(text, NOTES_RE, start=item8_start or 0)
    return {
        "1": drop_first_heading_line(slice_section(text, item1_start, item1a_start)),
        "1a": drop_first_heading_line(slice_section(text, item1a_start, item1a_end)),
        "md&a": normalize_mda_heading(slice_section(text, item7_start, item8_start)),
        "footnote": slice_section(text, notes_start, item9_start),
    }


def write_section_outputs(ticker, filing_path, sections):
    for section_name, content in sections.items():
        if not content.strip():
            continue
        raw_path = EDGAR_ROOT / section_name / ticker / filing_path.name
        sentence_path = EDGAR_ROOT / f"{section_name}_sentence" / ticker / filing_path.name
        raw_path.parent.mkdir(parents=True, exist_ok=True)
        sentence_path.parent.mkdir(parents=True, exist_ok=True)
        raw_path.write_text(content.strip() + "\n", encoding="utf-8")
        sentence_path.write_text(
            format_item1_text(content, section=section_name),
            encoding="utf-8",
        )
```

Repository-relative recovery and validation commands:

```
./.venv/bin/python EDGAR/qa/backfill_sections_from_filing_text.py <TICKER> \
  --source-root EDGARtools/parallel_3/filing_text


./.venv/bin/python -m unittest EDGAR/tests/test_backfill_sections_from_filing_text.py
./.venv/bin/python -m unittest EDGAR/tests/test_item1_sentence_paragraphs.py
```

Appendix B.3 AI Dictionary Counts and Panel Merge Logic

The AI disclosure measures are produced from section-level sentence corpora and merged into the cleaned firm-year accounting panel. The default empirical dictionary uses the conservative include bucket and applies exclusion/noise patterns before counting matched words and matched sentences.

Core count logic:

```
DEFAULT_INCLUDE_BUCKETS = (“default”,)
DEFAULT_EXCLUDE_BUCKETS = (“exclude”,)

def load_dictionary_patterns(dictionary_path):
    include_patterns = []
    exclude_patterns = []
    with dictionary_path.open(“r”, encoding=“utf-8-sig”, newline=““) as handle:
        for line in handle:
            row = parse_dictionary_row(line)
            if len(row) < 4 or not row[3].strip():
                continue
            bucket = row[0].strip().lower()
            pattern = re.compile(row[3].strip(), re.IGNORECASE)
            if bucket in DEFAULT_INCLUDE_BUCKETS:
                include_patterns.append(pattern)
            elif bucket in DEFAULT_EXCLUDE_BUCKETS:
                exclude_patterns.append(pattern)
    return include_patterns, exclude_patterns

def compute_file_counts(text_path, section_name, ticker, include_pattern, exclude_pattern):
    year, cik, filing_date = parse_file_metadata(text_path)
    content = text_path.read_text(encoding=“utf-8”, errors=“replace”)
    matched_words = total_words = matched_sentences = 0
    sentences = [line.strip() for line in content.splitlines() if line.strip()]
    for sentence in sentences:
        total_words += count_words(sentence)
        if exclude_pattern.search(sentence):
            continue
        if not include_pattern.search(sentence):
            continue
        sentence_matches = count_matched_words(sentence, include_pattern)
        matched_words += sentence_matches
        matched_sentences += int(sentence_matches > 0)
    return {
```

```
        "section": section_name,
        "ticker": ticker,
        "year": year,
        "matched_word_count": matched_words,
        "total_word_count": total_words,
        "matched_sentence_count": matched_sentences,
        "total_sentence_count": len(sentences),
    }
```

Repository-relative commands:

```
./.venv/bin/python EDGAR/qa/build_ai_dictionary_final_2026-03-16.py
./.venv/bin/python EDGAR/qa/aggregate_ai_dictionary_counts.py --workers 16
./.venv/bin/python EDGAR/qa/merge_ai_dictionary_into_panel.py
```

The main merged files are data/panel_all_financials_cleaned_with_ai_dictionary.csv and data/panel_all_financials_cleaned_with_ai_dictionary_analysis_ready.csv.

Appendix B.4 CRSP and CCM Event-Study Link Logic

The audited H1 implementation maps each SEC filing event from archive CIK into the local Compustat/CCM-linked identifier table, then uses WRDS-hosted CCM link records to identify the associated CRSP PERMNO/LPERMNO and pull CRSP daily returns. The script records ticker sanity checks, CUSIP checks, event-date gaps, and failure reasons.

Core CIK, link, and return-window logic:

```
def cik_from_sec_url(value):
    text = str(value or "")
    if "/data/" not in text:
        return pd.NA
    return clean_int(text.split("/data/", 1)[1].split("/", 1)[0])


def first_link_for_event(group, event_date, event_ticker_key):
    dates = group["date"].to_numpy()
    idx = int(np.searchsorted(dates, np.datetime64(pd.Timestamp(event_date).normalize()), side="left"))
    if idx >= len(group):
        return None
    same_day = group[group["date"].eq(group.iloc[idx]["date"])].copy()
    same_day["ticker_match_rank"] = np.where(
        same_day["ccm_ticker_key"].eq(event_ticker_key),
        0,
        1,
    )
    return same_day.sort_values(
        ["ticker_match_rank", "linkprim_rank", "active_rank", "common_rank", "permno"]
    ).iloc[0]


def market_adjusted_return(stock_ret, market_ret):
    return stock_ret - market_ret
```

Repository-relative command and main outputs:

```
./.venv/bin/python thesis-work/scripts/run_h1_crsp_ccm_audited_event_study.py


thesis-work/h1_event_study/h1_crsp_ccm_audited_car_windows.csv
thesis-work/h1_event_study/h1_crsp_ccm_audited_link_audit.csv
thesis-work/h1_event_study/h1_crsp_ccm_audited_failures.csv
thesis-work/h1_event_study/h1_crsp_ccm_audited_car_regressions.csv
thesis-work/h1_event_study/h1_crsp_ccm_audited_event_study_summary.csv
```

The CRSP daily extract contains delisting metadata columns, including DlyDelFlg, DelActionType, DelStatusType, DelReasonType, and DelPaymentType, but not a separate DLRET return column. Given the sample is built from a retrospective early 2026 large-cap constituent universe rather than a historical constituent universe, the main limitation is survivorship and sample-definition scope. DLRET can be added if the design is extended beyond the current constituent universe.

Appendix C. AI Dictionary Scope and Validation Protocol

Appendix C documents the dictionary scope, exclusion rules, ratio construction, and validation protocol used to construct the AI disclosure intensity measures. The purpose is to make clear that the thesis measures explicit AI-related language, not latent AI capability or disclosure credibility.

Appendix C.1 Dictionary Scope

The default dictionary is phrase-based and uses case-insensitive regular expressions. It includes 27 narrow AI, AI-adjacent, model-family, applied-business, and governance/risk patterns. The final empirical analysis uses these default include patterns together with five exclusion/noise patterns. Eleven extended automation terms are retained separately and are not counted in the main measure as they introduce more false-positive risk.

The term list was derived from a corpus-level candidate-discovery workflow. The development scan covered 17,560 section-formatted 10-K files and used a broad AI-context detector to surface 6,831 context files and 33,700 candidate context sentences. The resulting high-frequency terms and n-grams were reviewed manually, then separated into default include terms, extended include terms, dropped fragments, and exclusion/noise patterns. The main empirical analysis uses the 27 default include patterns and five exclusion patterns only. Extended automation terms remain outside the main measure as they are more likely to capture generic automation rather than AI disclosure.

Table 21 reports the conservative default AI dictionary terms used in the main measure.

Table 21. Default AI dictionary terms

Group	Term	Matching rule
Core shortform	AI	\bAI\b
Core technology	artificial intelligence	\bartificial intelligence\b
Core technology	generative AI	generative ai or generative artificial intelligence
Core shortform	GenAI	GenAI or Gen AI
Core technology	machine learning	\bmachine learning\b
Core shortform	ML	ML-powered, ML-driven, ML-enabled, ML-based, and equivalent business-context phrases
Core shortform	AI/ML	AI/ML, AI-ML, AI and ML
Core technology	deep learning	\bdeep learning\b
Core technology	neural network	neural network(s)
Core technology	computer vision	\bcomputer vision\b

Group	Term	Matching rule
Core technology	natural language processing	\bnatural language processing\b
Core shortform	NLP	\bNLP\b
Model family	large language model	large language model(s) and language model(s)
Model family	LLM	LLM(s)
Model family	foundation model	foundation model(s)
Model family	multimodal model	multimodal model(s)
Application form	AIOps	AIOps
Application form	chatbot / virtual assistant	chatbot(s) and virtual assistant(s)
Analytics family	predictive analytics	predictive analytics
Applied business AI	intelligent automation	intelligent automation
Applied business AI	AI-powered / AI-driven / AI-enabled	ai-powered, ai-driven, ai-enabled, ai-based, and ai-generated
Applied business AI	machine-learning methods	machine learning algorithms, machine learning techniques, and machine learning-based
Applied business AI	automated decision making	automated decision and automated decision making
Governance/risk	responsible AI	responsible ai
Governance/risk	AI governance / regulation	governing AI, regulating AI, AI regulation(s), and AI law(s)
Governance/risk	artificial intelligence act	Artificial Intelligence Act, EU AI Act, and related variants
Business context	AI strategy / initiative / workload	AI strategy, AI initiative(s), AI feature(s), AI training, and AI workload(s)

Appendix C.2 Exclusion Rules

The dictionary carries exclusion patterns for phrases that frequently look technological but are not AI disclosure in the intended sense. These include industrial automation, lab automation, automated clearing house or automated teller language, automated valuation or controls, and transformer references related to power equipment. These exclusions reduce false positives in Utilities, Industrials, Financials, and Health Care filings.

Appendix C.3 Matching Unit and Ratio Construction

The matching unit is the sentence. A sentence is classified as AI-related when at least one default dictionary expression appears in the sentence after section text has been converted to the local sentence-oriented EDGAR format. Matched-word counts are also calculated by overlapping dictionary-match spans with alphabetic word tokens. If multiple dictionary expressions overlap the same word token, the token is counted once.

The main section-level variables are:

$$AIWordRatio_{i,t,s} = \frac{MatchedAIWords_{i,t,s}}{TotalWords_{i,t,s}} \tag{9}$$

$$AISentenceRatio_{i,t,s} = \frac{MatchedAISentences_{i,t,s}}{TotalSentences_{i,t,s}} \tag{10}$$

Here, s denotes the 10-K section, i denotes the firm, and t denotes the fiscal year. The regression variables combine Item 1 and MD&A into the strategy block, retain Item 1A as the risk block, and retain footnotes as the accounting-note block.

Appendix C.4 Manual Validation Protocol and Results

The validation protocol addresses the risk that generic dictionary matching can misclassify financial text, which is why the AI dictionary is checked against the thesis corpus rather than accepted as a generic keyword list (Loughran & McDonald, 2011). Table 22 reports the audit checks used for the manual validation protocol.

Table 22. Manual validation protocol

Audit item	Target sample	Coding output
Matched AI sentences	200 random matched sentences	True AI disclosure, false positive, ambiguous
Technology-related unmatched sentences	50 random unmatched technology sentences	False negative, true negative, ambiguous
Section-level review	Matched sentences by Item 1, Item 1A, MD&A, and footnotes	Section-specific precision concerns
Exclusion review	50 sentences filtered by exclusion patterns	Correct exclusion or over-exclusion

The implemented manual review uses a fixed-seed sample of 300 sentence-level observations. The sample contains 200 sentences counted by the default dictionary after exclusion rules, 50 sentences filtered by exclusion rules, and 50 technology-context sentences not counted by the main dictionary. Each sentence is coded as clear AI disclosure, false positive, or ambiguous. Ambiguous observations are retained separately rather than forced into the clear-AI category.

Table 23 reports the manual validation results and the conservative precision floor.

Table 23. Manual validation results

Sample bucket	N	Clear AI	False positive / true negative	Ambiguous	Interpretation
Matched default dictionary sentences	200	192	3	5	Conservative precision floor of 96.0 percent when ambiguous cases are not counted as clear AI.
Filtered by exclusion rules	50	0	47	3	Exclusion rules mainly filter non-AI automation, banking automation, transformer, and lab automation noise.
Unmatched technology-context sentences	50	0	49	1	Small negative-control check. No clear missed AI sentence appears in this sample.

The clear false positives in the matched-default sample involve an entity name containing AI and two intelligent-automation sentences that do not explicitly mention AI or machine learning. The ambiguous cases mainly involve predictive analytics,

machine vision, robotics, advanced analytics, or noisy extracted table text. Treating all ambiguous matched-default observations as non-clear AI gives the 192/200, or 96.0 percent, conservative precision floor.

This validation exercise supports the view that the implemented dictionary has high precision for explicit AI and AI-adjacent disclosure language. It does not estimate formal recall, as the unmatched technology-context check is a small diagnostic sample rather than a complete search for all omitted AI sentences. The empirical results are therefore interpreted as associations involving high-precision, rule-based AI disclosure intensity rather than as direct evidence of AI capability, exhaustive AI coverage, or disclosure credibility.

Appendix D. AI Usage Report

Generative AI tools were used in a controlled and limited manner during the development of this thesis. We used OpenAI's ChatGPT GPT-5.4 and GPT-5.5, Anthropic's Claude Opus 4.6, and Google Gemini 3.1. These tools were used as coding, editing, and consistency-support tools, and all outputs were critically reviewed by us before utilization.

First, AI was used as coding assistance. ChatGPT supported the development, checking, and debugging of code related to the text-processing pipeline, AI dictionary counts, data merging, event-window construction, and regression outputs. All code suggestions were reviewed and tested by the authors before use.

Second, AI was used as editing support. ChatGPT, Claude, and Gemini helped improve phrasing, grammar, sentence structure, paragraph flow, and terminology consistency. This contributed to the quality of the thesis by making technical explanations clearer and improving consistency across sections.

Third, AI was used for consistency checks. The tools helped identify repeated wording, unclear transitions, overly strong claims, and places where the interpretation needed to remain aligned with the non-causal research design. AI was also used for limited checks of citation wording and reference formatting. Final citation choices, source interpretation, empirical claims, and conclusions were reviewed and decided by us.

The main risks were inaccurate outputs, unsupported suggestions, distortion of analytical meaning, and overstatement of findings. These risks were reduced by using AI only as a support function, manually reviewing all suggestions, testing code before use, and rejecting any wording that implied causality, overstated the evidence, or treated AI disclosure as direct evidence of internal AI use.

Using AI tools showed that they can improve efficiency, clarity, and consistency, especially for editing, debugging, and checking long technical sections. However, the process also showed that AI outputs require careful judgment and cannot replace the authors' own analysis, interpretation, or academic responsibility. All substantive reasoning, empirical work, methodological choices, and conclusions remain ours.