When AI-Washing Meets Initial Public Offerings

Rui Dai, Gangxing (Gary) Guo*, and Bohui Zhang

Current version: July 2026

* Rui Dai is from Wharton Research Data Services (WRDS) and CUHK-Shenzhen, The Wharton School, University of Pennsylvania, 3819 Chestnut Street Philadelphia PA, 19104, USA, rdai@wharton.upenn.edu; Gangxing (Gary) Guo is from School of Management and Economics, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), 2001 Longxiang Road, Longgang District, Shenzhen, China, 518172, gangxingguo@link.cuhk.edu.cn; and Bohui Zhang is from Shenzhen Finance Institute and School of Management and Economics, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), 2001 Longxiang Road, Longgang District, Shenzhen, China, 518172, bohuizhang@cuhk.edu.cn. We have received helpful comments and suggestions from Arnie Cowan (discussant), Emilio Bisetti (discussant), Susan Shu, and seminar and conference participants at the 2025 CUHK-Shenzhen FIN Ph.D. Student Lunch Seminar, the 22nd Chinese Finance Annual Meeting, the 2025 China Tech-Fin Research Conference, JIBS-Tsinghua Paper Development Workshop, the 2026 China Fintech Research Conference, the 2026 Financial Management Association European Conference, the 5th Annual Hong Kong Conference on FinTech and AI in Finance, the 2026 Asia Meeting of the Econometrics Society – China and the 2026 China International Conference in Finance. This paper won the best paper award of the 22nd Chinese Finance Annual Meeting.

When AI-Washing Meets Initial Public Offerings

Abstract

Valuing technological intangibles is central to IPO pricing, yet the literature lacks reliable measures of firms' technological footprints before they go public. We address this challenge in the context of artificial intelligence (AI), a general-purpose technology with substantial value-creation potential but difficult-to-verify firm-level inputs. Using granular data on AI-skilled employees and AI patent filings to benchmark AI-related language in S-1 registration statements, we develop a novel framework that classifies U.S. IPOs from 2010 to 2024 as AI-washing, AI-investment, or non-AI firms. We validate this classification using independent evidence from firms' post-IPO AI activities and pre-filing regulatory scrutiny. AI-washing firms receive more favorable initial market reactions but significantly underperform after listing, whereas firms with substantive AI inputs earn strong long-run returns. Cross-sectional evidence supports three complementary channels. The *value premium channel* shows that AI-washing firms benefit more from AI-related valuation premiums; the *verification channel* shows that unsupported AI-related claims are harder to detect when underlying AI capabilities are more difficult to verify; and the *litigation risk channel* shows that legal exposure appears to discipline such claims after listing, with suggestive evidence that higher-risk AI-washing firms subsequently scale back their AI-related disclosures. Our findings show how the difficulty of verifying emerging technologies enables opportunistic disclosure to distort IPO valuation and provide a new framework for assessing whether pre-offering technological narratives are supported by substantive investment.

Keywords: AI-washing; Initial public offerings; Prospectus disclosure.
JEL code: G14, G24, O32, M41

1 Introduction

The emergence of artificial intelligence (AI) in the U.S. IPO market creates a novel challenge regarding how to value technological intangibles at issuance. AI-related disclosure, already increasing before the release of ChatGPT in late 2022, has since become pervasive. By April 2023, 54% of U.S. IPOs completed year to date referenced AI in their prospectuses, following six consecutive years of increases.[1] More recently, AI was the main component to the valuation narrative of SpaceX's June 2026 offering, which raised $85.7 billion in gross proceeds, making it the largest IPO on record.[2] The prospective AI IPO pipeline is similarly large. Anthropic and OpenAI subsequently entered the IPO pipeline through confidential S-1 submissions, with combined private-market valuations of approximately $1.82 trillion, or 21% of the aggregate value of global unicorns.[3] These developments indicate that AI has become increasingly salient in both the content and scale of the IPO market.

However, this growing importance gives rise to a more difficult valuation problem. Because AI can materially affect productivity, product innovation, and scalability, it can generate large differences in firm value and growth potential, leading investors to respond strongly even to modest claims of AI capability. Yet genuine AI capability rests on technically complex algorithms, proprietary data, and specialized human capital that outsiders cannot readily assess without technical expertise, particularly when IPO firms lack established public track records. They create incentives for issuers to overstate their AI capabilities to attract investors, generating a mismatch between AI disclosures ("AI-talk") and AI inputs ("AI-walk"), a phenomenon known as *AI-*

[1] Renaissance Capital, "AI Is the IPO Market's Next Buzzword," April 26, 2023.
[2] Space Exploration Technologies Corp., Form 424B4, June 12, 2026; SpaceX, "Announcement of Closing of Initial Public Offering," June 15, 2026; EY, "Global IPO Proceeds Triple in First Half of 2026," July 16, 2026.
[3] Anthropic, "Anthropic Confidentially Submits Draft S-1 to the SEC," June 1, 2026, and "Anthropic Raises $65B in Series H Funding," May 28, 2026; OpenAI, "Confidential Submission of Draft S-1 to the SEC," June 8, 2026, and "OpenAI Raises $122 Billion to Accelerate the Next Phase of AI," March 31, 2026; Russell Investments, "Four Watchpoints for 2026's Potential Mega IPO Class," May 26, 2026. The 21% figure is the authors' calculation based on the firms' respective post-money valuations of $965 billion and $852 billion and the $8.6 trillion aggregate valuation of global unicorns.

washing.[4] These incentives are especially strong in IPOs, where issuers lack public track records and valuations depend heavily on uncertain growth prospects. This setting raises three questions. How prevalent is AI-washing in IPO prospectuses? Can investors identify it ex ante? How does it affect initial IPO valuations and long-run stock performance?

In this paper, we develop a measurement framework that identifies this disclosure–input mismatch before firms go public. We measure AI-talk using dictionary-based textual analysis of S-1 registration statements. We measure AI-walk using two granular indicators of pre-IPO technological capital: AI-skilled employees identified from LinkedIn profiles and AI-related patent filings obtained from the U.S. Patent and Trademark Office. We classify an IPO as AI-washing when its AI disclosure exceeds a historical, industry-specific 95th-percentile threshold but the issuer has neither observable AI employees nor AI patent filings. IPOs with either type of observable input are classified as AI-investment IPOs, and the remaining firms are classified as non-AI IPOs. The stringent zero-input requirement makes the classification conservative and reduces the likelihood of labeling firms with observable AI activity as AI-washing.

Applying this framework to 1,671 U.S. IPOs from 2010 to 2024, we identify 107 AI-washing IPOs, representing 6.4% of the sample. Another 393 issuers, or 23.5%, have observable AI inputs and are classified as AI-investment IPOs. The remaining 1,171 issuers, or 70.1%, are classified as non-AI IPOs. AI-washing occurs throughout the sample period, reaches 14.3% of IPOs in 2023, and appears across industries rather than only in conventional technology sectors. The disclosure–input mismatch is therefore economically meaningful rather than anecdotal.

Our main results reveal a sharp reversal in how the market prices AI-washing firms. At issuance, AI-washing IPOs earn first-day returns that are 7.2 percentage points higher than those

[4] In March 2024, the U.S. Securities and Exchange Commission charged two investment advisers, Delphia (USA) Inc. and Global Predictions, Inc., for overstating their use of artificial intelligence in investor communications, alleging that their AI-related claims were not supported by actual implementation:
https://www.hklaw.com/en/insights/publications/2024/04/beyond-the-hype-the-secs-intensified-focus-on-ai-washing-practices
In January 2025, the SEC charged Presto Automation Inc., a restaurant-technology company that went public through a SPAC, for making materially misleading AI-related statements between 2021 and 2023:
https://www.sec.gov/enforcement-litigation/administrative-proceedings/33-11352-s

of otherwise comparable non-AI IPOs, whereas AI-investment IPOs do not earn significantly higher initial returns. This favorable initial response does not persist. Relative to non-AI IPOs, AI-washing IPOs underperform by 8.8 percentage points over two years and 13.9 percentage points over three years. By contrast, AI-investment IPOs outperform non-AI IPOs by 10.2 and 22.0 percentage points over the same horizons. The resulting performance gap between AI-washing and AI-investment IPOs is 19.0 percentage points after two years and 36.0 percentage points after three years, with both differences statistically significant at the 1% level. The combination of favorable first-day reactions and subsequent underperformance is consistent with a gradual correction of AI-related valuations as information about issuers' underlying technological capabilities becomes available.

Several additional tests support the interpretation of our classification. AI-washing firms discuss AI more frequently than non-AI firms immediately after listing, but this discussion declines markedly after the fourth post-IPO quarter. Their subsequent AI hiring and patenting activity also remains closer to that of non-AI firms than to that of AI-investment firms. Around the expiration of IPO lockups, AI-washing and AI-investment firms exhibit similar pre-expiration trends, but AI job postings at AI-washing firms decline significantly afterward. This pattern is less consistent with firms continuing to build AI capabilities after an unsuccessful initial attempt. In addition, greater SEC comment-letter activity before an issuer's filing is associated with a lower probability of AI-washing but is unrelated to the probability of observable AI investment. This asymmetric relation provides institutional validation that our measure captures aggressive disclosure rather than AI activity more generally.

The return evidence is robust to alternative empirical choices. The results persist under propensity-score matching and entropy balancing, alternative return benchmarks, different AI-disclosure thresholds, and a continuous residual-based measure of AI-washing. Calendar-time portfolios of AI-washing firms generate Fama–French five-factor alphas of −1.49% and −1.13% per month over 24- and 36-month holding windows, respectively. The findings also remain after

excluding issuers that disclose third-party AI tools or outsourcing, appear as AI customers in supply-chain data, or have identifiable AI collaboration partners. Thus, the results are unlikely to reflect standard risk exposures, observable firm differences, or unmeasured reliance on readily identifiable external AI capabilities.

To illuminate the underlying economic drivers, we explore three distinct channels through which AI-washing generates post-IPO underperformance. We investigate the *value premium channel*, the *verification channel*, and the *litigation risk channel*. Together, they capture why unsupported AI narratives may be initially priced, why they are difficult to discount at issuance, and why they become costly to maintain after listing.

First, the *value premium channel* posits that initial overvaluation is amplified in sectors where investors extrapolate AI innovation value and assign steeper valuation premiums to genuine technological growth opportunities. This mechanism builds on the foundational view that AI generates profound productivity shocks and innovation gains that are progressively capitalized into corporate valuation (Bresnahan and Trajtenberg 1995; Cockburn, Henderson, and Stern 2018; Babina et al. 2024). We measure this premium using the extent to which AI patent stocks are reflected in public-firm valuations within each Fama–French 12 industry (Hall, Jaffe, and Trajtenberg 2005; Kogan et al. 2017). Consistent with this channel, AI-washing firms receive higher price-to-sales multiples in industries where AI innovation commands a larger valuation premium. Their subsequent two- and three-year abnormal returns are also more negative in these industries.

Second, the *verification channel* focuses on the asymmetric information environment surrounding corporate innovation. Genuine AI capability is difficult for outside investors to assess because it depends on underlying inputs, such as data assets, technical talent, and algorithms, that are largely unobservable (Cockburn, Henderson, and Stern 2018; Abis and Veldkamp 2023). This verification friction is likely greater when the issuer's disclosed business model is semantically distant from established AI technologies, making it difficult for investors to evaluate true AI

capacities. To map this friction, We measure the difficulty of evaluating an issuer's AI claims using the semantic distance between the Business section of its S-1 filing and a broad corpus of AI patents. The long-run underperformance of AI-washing firms is significantly more pronounced when the issuer's business is farther from established AI technologies. Semantic distance alone does not predict long-run IPO returns. Its effect arises primarily in conjunction with AI-washing, consistent with verification difficulty allowing unsupported claims to receive more favorable initial assessments.

Third, the *litigation risk channel* operationalizes an ex-post disciplining mechanism. Exaggerated AI disclosures in S-1 prospectuses expose issuers to severe legal exposure once they go public. Under the Private Securities Litigation Reform Act of 1995, the statutory safe harbor protecting forward-looking statements does not apply to traditional IPOs (Johnson, Kasznik, and Nelson 2001). Consequently, if AI-washing firms present aspirational capabilities as current material facts, they lack legal immunity if subsequent performance falls short. Following Kim and Skinner (2012) and Donelson, Flam, and Yust (2022), we measure expected litigation risk based on post-listing firm characteristics one year after the listing. we find that AI-washing firms experience more negative second-year abnormal returns when legal exposure is higher. These firms also reduce AI-related language more sharply in their first Form 10-K filings and subsequent earnings calls relative to their S-1 disclosures. Together, these findings are consistent with litigation risk increasing the cost of maintaining unsupported AI narratives and contributing to their post-listing retrenchment.

Our paper contributes to several strands of literature. First, we contribute to the IPO literature on issuance valuation and post-listing long-term performance (Ritter 1991; Loughran and Ritter 1995; Ritter and Welch 2002). While previous literature extensively focuses on how information affects the pricing process (Teoh, Welch, and Wong 1998; Hanley and Hoberg 2010), technology presents a unique challenge at the IPO stage because it functions as a largely unobservable intangible asset. We extend this line of research by investigating how the market values, and

sometimes misvalues, these technological intangibles at issuance. By utilizing granular, micro-level datasets on human capital and patent filings, we provide a new framework to verify these unobservable assets and explain post-IPO returns when traditional metrics fail to capture technological realities.

Second, we contribute to the strategic disclosure literature. Prior work studies settings in which firms' public communications diverge from underlying actions, including greenwashing and diversity washing (Delmas and Burbano 2011; Duchin, Gao, and Xu 2025; Baker et al. 2024). Unlike the settings examined in prior studies, artificial intelligence is adopted across a wider range of industries and creates a more ambiguous form of information asymmetry that is difficult for investors to untangle. We add to this literature by operationalizing a systematic approach to measure the long-run market responses to such strategically overstated technological claims within the high-stakes IPO window.

Third, we add to the literature on market reactions to major technological waves. Prior studies show that investor enthusiasm for new technologies can affect prices during episodes such as the Dot-com bubble and blockchain mania (Ljungqvist and Wilhelm 2003; Ofek and Richardson 2003; Cheng et al. 2019). However, AI is fundamentally distinct because of its genuine, broad-based capacity for corporate value creation (Babina et al. 2024; Wu, Lou, and Hitt 2024). Crucially, our findings underscore that this intrinsic capacity for value creation does not shield the market from informational frictions; when technological investments remain difficult for investors to verify, overstated claims about a transformative technology can paradoxically amplify, rather than mitigate, market mispricing.

2 Data and Variable Construction

2.1 Data

We construct our sample from U.S. IPOs recorded in SDC Platinum between 2010 and 2024. We begin the sample in 2010 because the early 2010s marked the rise of the modern AI wave,

when advances in deep learning accelerated the diffusion of AI into firms' innovation processes, products, and business models. Following Lowry et al. (2017), we exclude IPOs with an offer price less than $5, American Depository Receipts (ADRs), real estate investment trusts (REITs), and closed-end funds, as well as firms without first-day trading records in CRSP. We then link the IPO sample to data on S-1 disclosures, AI-related human capital, AI patent filings, post-IPO stock returns, and firm fundamentals. Appendix A defines all variables, and Internet Appendix Table IA1 reports the detailed sample-selection procedure. The final sample used in the main analyses contains 1,671 IPOs.

2.2 Variables

2.2.1 AI employees, AI patents, and AI disclosures

Our classification of IPO firms relies on three dimensions of AI activity: AI-related disclosure in IPO prospectuses, AI-related human capital, and AI-related patent filings. The first captures the extent to which firms emphasize AI in IPO disclosures, whereas the latter two capture observable AI inputs.

We measure AI disclosure using firms' S-1 registration statements, in which issuers describe their business models, technologies, risk factors, and growth opportunities as part of the IPO process. We construct an AI keyword dictionary based on prior studies of AI technologies and innovation (Bloom et al. 2021; Giczy, Pairolero, and Toole 2022; Henderson 2018; Lou and Wu 2021; WIPO 2019; Wu, Lou, and Hitt 2024; Babina et al. 2024).[5] The dictionary contains 97 AI-related terms, 57 of which appear at least once in the IPO corpus, as reported in Appendix B. Internet Appendix Figure IA1 plots the frequency distribution of AI-related keywords. The distribution is highly concentrated: broad terms, including "machine learning", "artificial intelligence" and "predictive analytics", account for more than 70% of all AI-related mentions,

[5] We use regular expressions (*regex*) with pattern matching, for example, "machine learn" can match "machine learning"; "ada boost" can match "adaboost", "ada-boost", and "ada boosting". Therefore, our matched keyword pairs are far more than those keywords in the appendix list.

suggesting that issuers often rely on broad AI terminology when describing their AI activities. We define AI disclosure intensity (*AIWord*) as the total number of AI-related keyword mentions in a firm's S-1 prospectus.

We measure firms' substantive AI input using two proxies: AI employees and AI patents. We require both proxies to be observed before the IPO listing date to mitigate look-ahead bias. We capture AI employees (*AIEmployee*) as the number of AI employees employed by the firm as of the listing date. The data are obtained from the April 2025 snapshot of *Revelio Labs* on *WRDS*, a premier workforce data provider that compiles a large collection of *LinkedIn* profiles and job postings covering a substantial portion of the labor market. Importantly, the dataset records each worker's employment history, including the start and end dates of each position, as well as information on skills and educational background.

We classify an employee as an AI employee if the employee reports AI-related skills based on the skill list in Babina et al. (2024).[6] We exclude interns who have not been formally hired. To reduce noise from self-reported profiles, we impose two additional restrictions. First, the employee must hold at least one degree in a STEM field, including engineering, information technology, mathematics, statistics, or physics. Second, AI skills must account for a sufficiently large share of the employee's total reported skills. Internet Appendix Table IA2 provides summary statistics for the LinkedIn profiles and lists the AI-related skills used in the classification.

We also capture AI input using AI patent filings. We define AI patents (*AIPatent*) as the number of AI patent filings from three years before the listing date through the listing date. We use patent filing dates rather than patent grant dates to better capture the timing of firms' AI investment decisions (Luong et al. 2017). The patent data are obtained from the USPTO Artificial Intelligence Patent Dataset (AIPD) provided by the Office of the Chief Economist (OCE), which

[6] To better unify employees with different cultural backgrounds, we first translate all languages into English using *Google Translate* API and search for the keywords in English. To further mitigate the wrong interpretation of some skill words, we further exclude those skills with less than 4 characters.

covers all United States Patent and Trademark Office (USPTO) patent documents (i.e., patents and pre-grant publications, or PGPubs) (Giczy, Pairolero, and Toole 2022).[7]

First, we use the *PERMNO-patent_number* links provided by Kogan et al. (2017) and additional fuzzy name matching to match patent filings to firms, including both granted patents and pre-grant publications.[8] Second, we use the AIPD *predict93_any_ai* indicator, which equals one if a patent document is classified as AI-related under the 93% prediction-threshold criterion. Internet Appendix Table IA3 reports summary statistics of matched AI patents.

2.2.2 IPO returns

We measure IPO performance using both long-term buy-and-hold abnormal returns (*BHAR*) and initial returns. Specifically, *BHAR[0,1yr]*, *BHAR[0,2yr]*, and *BHAR[0,3yr]* are buy-and-hold abnormal returns from the close of the first trading day to one, two, and three years after the IPO, respectively. *BHAR[1,2yr]* is defined as the abnormal return from one year after the IPO to two years after the IPO for firms that remain listed at the end of the first year. All abnormal returns are benchmark-adjusted using the Fama–French 5×5 size and book-to-market portfolios (Barber and Lyon 1997; Qian, Shao, and Liao 2024; Liu and Wu 2021). We use value-weighted and equal-weighted market returns as alternative benchmarks in robustness tests. Initial returns (*InitialReturn*) are defined as the percentage change between the first-day closing price and the offer price.

2.2.3 Other variables

To control for other factors that may affect post-IPO stock performance, we include a set of standard control variables following the IPO literature (Liu and Wu 2021; Ritter 1991). At the IPO level, the control variables include initial returns (*InitialReturn*), high-reputation underwriters

[7] USPTO artificial intelligence patent dataset: https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset

[8] We match AI patent predictions to IPO firms using KPSS patent–PERMNO links for granted patents and manually verified assignee–PERMNO links for pre-grant applications matched through USPTO *pgpub_id*.

(*HighUW*), venture capital backing (*VC*), Big 4 auditors (*Big4*), carve-out IPOs (*Carve*), pure primary offerings (*Primary*), the ratio of secondary shares offered to total shares (*Second*), IPO overhang (*Overhang*), and the log of gross proceeds (*Proceeds*). When the initial return (*InitialReturn*) is the dependent variable, we replace it with price revision (*PrcRev*) as the control for investor demand during the book-building process.

At the firm level, we control for the log of firm age (*Age*), industry indicators (*HighTech*, *Biotech*, and *Internet*), the log of pre-IPO sales (*Sales*), the log of pre-IPO total assets (*Assets*), and a dummy of the positive earnings per share prior to the IPO (*PositiveEPS*). All continuous variables are winsorized at the 1st and 99th percentiles to mitigate the influence of outliers.

2.3 Summary statistics

Panel A of Table 1 reports the summary statistics for our main variables. Panel A1 reports IPO stock market performance, including buy-and-hold abnormal returns and initial returns. Post-IPO abnormal returns are negative on average: the mean buy-and-hold abnormal return is −11.0% over the first year, −13.7% over the first two years, and −17.2% over the first three years. The average abnormal return during the second post-IPO year is −9.4%. The average first-day return is 19.7%. These patterns are consistent with the well-documented long-run underperformance of IPO firms (Ritter 1991).

Panel A2 reports IPO AI characteristics, including AI keyword frequency, AI patents, and AI employees. On average, IPO firms mention approximately 2.8 AI keywords in their prospectuses, while the mean numbers of pre-IPO AI employees and AI patents are 0.5 and 1.5, respectively. This pattern suggests that firms emphasize AI in their IPO disclosures, while observable AI inputs, as measured by employees and patents, remain relatively limited.

Panel A3 reports summary statistics for other IPO and firm characteristics. On average, 51.3% of IPOs are underwritten by high-reputation underwriters, 59.1% receive venture capital backing, and 39.3% are audited by Big 4 accounting firms, indicating that a substantial fraction of IPO firms

is supported by reputable intermediaries. The mean gross proceeds are approximately $248.8 million, while the median firm age is about 10 years. The average pre-IPO total assets and sales are $1.4 billion and $652.5 million, respectively, although the distributions are highly skewed. In terms of industry composition, 55.0% of firms operate in high-tech industries, 32.3% in biotech sectors, and 9.5% are Internet-related firms. Overall, the sample consists primarily of young, growth-oriented firms.

Panel B reports pairwise correlations among the main variables. Panel B1 shows that buy-and-hold abnormal returns are positively correlated across horizons, whereas their correlations with initial returns are small. Among the AI-related variables, *AIWord*, *AIPatent*, and *AIEmployee* are positively correlated with one another, consistent with the idea that AI disclosure and observable AI inputs are related but not identical. Panel B2 reports correlations among IPO and firm characteristics. Overall, the pairwise correlations are generally moderate in magnitude and do not indicate serious collinearity concerns.

Panel C reports univariate comparisons of AI-washing IPOs with non-AI and AI-investment IPOs. The evidence does not imply a monotonic ordering of observable characteristics across the three groups. Instead, the relative position of AI-washing IPOs is benchmark-dependent. For several measures of firm scale, industry composition, and intermediary involvement, their sample means fall between those of non-AI and AI-investment IPOs. A small set of differences is more persistent: relative to both benchmarks, AI-washing IPOs are less likely to engage Big 4 auditors or conduct pure primary offerings and tend to exhibit greater overhang. In contrast, firm age, total assets, and the fraction of secondary shares do not differ systematically across the comparisons. Taken together, these univariate patterns provide limited support for an explanation based solely on observable fundamentals. Subsequent multivariate specifications include the full set of firm- and offering-level characteristics as controls.

[Table 1 here]

3 AI-Washing

We develop a measurement framework that identifies AI-washing by comparing firms' AI disclosure in IPO prospectuses with observable AI inputs before listing. The framework combines three sources of information: S-1 prospectuses for AI disclosure, LinkedIn profiles for AI human capital, and the USPTO Artificial Intelligence Patent Dataset for AI patent filings. This section describes the construction and validation of the AI-washing measure.

3.1 The relation between AI-talk and AI-walk

We first validate the measurement framework by examining whether AI disclosure is associated with observable AI inputs before listing. If firms' AI-related disclosure reflects underlying AI activities on average, firms with more AI employees or AI patent filings should mention AI more frequently in their S-1 prospectuses.

We estimate the following Poisson regression:

$$AIWord_{i,t} = \beta_1 AIEmployee_{i,t} + \beta_2 AIPatent_{i,t} + FE_{i,t} + \varepsilon_{i,t} \quad (1)$$

where *AIWord* is the frequency of AI-related keywords in the S-1 prospectus, while *AIEmployee* and *AIPatent* capture pre-IPO AI human capital and AI patent filings, respectively. We include two-digit SIC industry and IPO year fixed effects to control for industry-specific and time-varying factors that may affect firms' AI disclosure. Standard errors are adjusted for heteroskedasticity and clustered by industry.

[Table 2 here]

Table 2 reports the results. The coefficients on *AIEmployee* are positive and statistically significant across specifications. In the full specification, the coefficient is 0.042, implying that a one-standard-deviation increase in pre-IPO AI employees is associated with an approximately 10.0%[9] increase in AI-related keyword mentions. The coefficients on *AIPatent* are positive when

[9] The economic magnitude is calculated as $e^{0.042 \times 2.280} - 1 = 0.100$.

patent filings are included separately but become smaller and statistically insignificant after controlling for AI employees and fixed effects. These results indicate that AI disclosure is related to observable AI inputs on average, but the relation is far from one-to-one. The explanatory power remains moderate, and the coefficient on AI patent filings weakens after controlling for AI human capital and fixed effects.

3.2 AI-washing measurement

The imperfect alignment between AI disclosure and observable AI inputs motivates our AI-washing classification. Following the logic of Baker et al. (2024), we identify AI-washing from a mismatch between firms' IPO disclosure and observable AI inputs.

A firm is classified as an AI-washing IPO if the following conditions are satisfied:

(1) AI disclosure in the S-1 prospectus is above the *95th percentile* of AI disclosure among all other IPO firms in the same Fama-French 12 industry, calculated over an expanding window from 2007 through the year before the firm's IPO;[10]

(2) The firm has *zero* AI employees as of the listing date, **and**

(3) The firm has *zero* AI patent filings from three years before the listing date through the listing date.

The criteria can be summarized as follows:

$$\begin{cases} AIWord_{i,t} > P_{95}\left(AIWord_{j,\tau}\right), \tau \in [2007, t-1], j \neq i \\ AIEmployee_{i,t} = 0 \\ AIPatent_{i,t} = 0 \end{cases} \quad (2)$$

We classify the remaining IPOs into two groups. To isolate the role of observable AI investment and use non-AI IPOs as the benchmark group, we classify an IPO as an AI-investment IPO if either of the following conditions is satisfied:

(1) The firm has *at least one* AI employee as of the listing date ($AIEmployee_{i,t} > 0$); **or**

(2) The firm has *at least one* AI patent filing from three years before the listing date through the

[10] The *95th percentile* threshold is intended to identify firms with unusually high AI disclosure relative to other IPO firms in the same industry. The results are robust to alternative nearby thresholds, such as the 90th percentile.

listing date ($AIPatent_{i,t} > 0$).

IPOs with neither observable AI inputs nor unusually high AI disclosure are classified as non-AI IPOs.

Our measurement approach is tailored to the IPO setting. First, because the number of IPO observations is substantially smaller than in broad firm-year samples, a two-dimensional sorting approach similar to Baker et al. (2024) is less suitable. We therefore adopt a rule-based classification that can be implemented consistently in the IPO sample. Second, we use two complementary measures of observable AI inputs: AI human capital and AI patent filings. These measures capture different aspects of firms' AI activities before going public. Third, the AI-washing definition imposes a stringent zero-input requirement: a firm must have both zero AI employees and zero AI patent filings to be classified as AI-washing. This restriction reduces the likelihood of classifying firms with some observable AI activity as AI-washing.

3.3 Distributions

We next describe the composition of the AI-washing classification to show how IPO firms are distributed. Figure 1 reports the distribution of IPO firms based on the interaction between AI disclosure, or "AI talk," and observable AI inputs, or "AI walk." The sample is divided into four groups: AI-aligned IPOs, AI-quiet IPOs, AI-washing IPOs, and non-AI IPOs. Among the 1,671 IPOs in the classification sample, 393 firms, or 23.5%, have observable AI inputs and are classified as AI-investment IPOs. Within this group, 232 firms, or 13.9% of the sample, are AI-aligned IPOs, which combine observable AI inputs with high AI disclosure. Another 161 firms, or 9.6%, are AI-quiet IPOs, which have observable AI inputs but relatively low AI disclosure. In contrast, 107 firms, or 6.4%, are classified as AI-washing IPOs because they have high AI disclosure but no observable AI employees or AI patent filings. The remaining 1,171 firms, or 70.1%, are classified as non-AI IPOs.

[Figure 1 here]

Figure 2 Panel A presents the time-series distribution of these categories from 2010 to 2024. AI-investment IPOs become more common over time in terms of counts, rising from 16 IPOs in 2010 to a peak of 71 IPOs in 2021, although their annual share fluctuates across years. AI-washing IPOs also appear throughout the sample period and become more visible after 2018. Their annual share is higher after 2018, reaching 9.6% in 2018, 10.3% in 2020, 8.8% during the 2021 IPO wave, and a peak of 14.3% in 2023. These patterns suggest that both observable AI inputs and high AI disclosure became increasingly salient in the IPO market as AI technologies diffused across firms.

Panel B reports the composition of IPO classifications within each Fama–French 12 industry. AI-investment IPOs account for particularly large shares of IPO activity in technology-intensive industries, most notably Business Equipment and Telephone and Television Transmission. AI-washing IPOs, by contrast, appear across a broader range of industries. Their within-industry shares are relatively high in Chemicals and Allied Products and Consumer Nondurables, and they are also present in Healthcare, Finance, Manufacturing, and Wholesale, Retail, and Some Services. This cross-industry presence indicates that the disclosure–input mismatch captured by our measure is not confined to conventional technology sectors.

Taken together, the evidence indicates that AI disclosure and observable AI inputs align for many IPO firms, while a nontrivial subset emphasizes AI without corresponding observable investment. Moreover, this mismatch recurs across time and industries rather than being driven by a particular year or sector. These patterns provide the basis for the subsequent analysis of post-IPO performance.

[Figure 2 here]

3.4 AI-Washing measurement validation

3.4.1 Post-IPO AI substantial engagement

To further validate our AI-washing measure, we examine firms' AI-related activities after the IPO. This analysis uses post-IPO outcomes that are distinct from the S-1 prospectus and allows us to assess whether the AI-washing classification captures a disclosure pattern characterized by high IPO-stage AI emphasis but limited observable AI inputs.

Figure 3 plots post-IPO AI-related disclosure, hiring, and patenting activities for AI-washing, AI-investment, and non-AI IPOs. Panel A presents the average number of AI-related keyword mentions in earnings calls from the first to the eighth quarter after the IPO. Non-AI IPOs rarely mention AI in post-IPO earnings calls. AI-investment IPOs generally exhibit higher and more persistent discussion. AI-washing IPOs mention AI more frequently than non-AI IPOs in the early post-IPO quarters, indicating that they are not identified solely from isolated keyword usage in S-1 filings. However, their AI-related discussion declines substantially after the fourth post-IPO quarter and remains below that of AI-investment IPOs in later quarters.

Panels B and C provide additional validation using observable AI inputs. Panel B shows that AI-investment IPOs consistently hire more AI-related employees after the IPO, whereas AI-washing and non-AI IPOs make few AI-related hires. Panel C shows a similar pattern for AI patent filings: AI-investment IPOs account for substantially more AI patenting activity, while AI-washing IPOs remain close to non-AI IPOs. Together, these patterns support the validity of the AI-washing measure. Firms classified as AI-washing display elevated AI-related disclosure around the IPO but do not sustain comparable post-IPO AI discussion or demonstrate the same level of observable AI investment as firms with AI-related inputs.

[Figure 3 here]

3.4.2 AI job postings dynamic around expiration

A potential concern is that our measure may fail to capture some forms of genuine AI engagement due to data limitations. As a result, firms classified as AI-washing firms may in fact be AI-failure firms, that genuinely attempted to invest in AI but lacked sufficient inputs, leading

to unsuccessful AI implementation and substantial losses. To address this competing AI-failure explanation, we use the lock-up expiration setting to examine the job posting dynamic.

Prior to lockup expiration, corporate insiders have strong incentives to sustain favorable and potentially promotional disclosures, because maintaining high secondary-market valuations can increase the proceeds from post-lockup share sales (Aggarwal, Krigman, and Womack 2002). Once insiders can sell and realize the value of their holdings, their incentives to continue supporting such narratives should decline sharply. Therefore, if the AI disclosures are primarily promotional, we expect AI-washing firms to reduce AI-related hiring after lockup expiration. In contrast, if these firms are genuinely pursuing AI but fail because of insufficient inputs, we should continue to observe meaningful AI-related hiring after the lockup expires.

We estimate a dynamic Poisson regression using an event window of [-3, +3] quarters around the expiration day, focusing on the subsample of AI-washing and AI-investment IPOs as follows:

$$AIJobPost_{i,t} = \sum_{\substack{k\neq -1 \\ k=-3}}^{3} \beta_k \left(AIWashing_i \times 1(rel_quarter = k) \right) + \gamma AIWashing_i + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (3)$$

where k refers to the relative quarters to the expiration day. $Controls_i$ include a set of control variables the same as the baseline test. We control for fixed effects at the two-digit industry× year_quarter and relative quarter levels. Standard errors are adjusted for heteroskedasticity and clustering by industry.

Figure 4 Panel A plots the dynamic effects of AI-washing on AI job postings around the expiration of the IPO lockup period, and Table IA5 reports the estimated coefficients. The coefficients represent the marginal effects of *AIWashing* × *rel_month* relative to the baseline period (*quarter_−1*). The estimates before the expiration date are generally close to zero and statistically insignificant, suggesting that AI-washing IPOs do not exhibit systematical trend prior to the lockup expiration compared with AI-investment IPOs. After the expiration date, the coefficient for quarter 2 and 3 is significantly negative, indicating a substantial reduction in AI-related job postings among AI-washing firms relative to AI-investment firms two quarters after

the lock-up expiration. Panel B shows additional comparasion between AI-washing IPOs and AI-investment IPOs AI job posting dynamics with non-AI IPOs as the benchmark.

This pattern is more consistent with AI-washing instead of AI-failures. If AI-washing firms were genuinely attempting but failing to build AI capabilities, we would expect there are no significant changes before and after the lock-up expirations. Instead, AI-related job postings decline significantly in the post-expiration period. Overall, the evidence suggests that AI-washing IPOs under our measurement is more likely to strategically emphasize AI instead of real engagements in continuous genuine AI investment.

[Figure 4 here]

3.4.3 Pre-filing regulatory scrutiny

Another validation uses the pre-filing regulatory environment as an institutional benchmark. If the AI-washing measure captures a strategic mismatch between firms' AI-related disclosure and their observable AI inputs, AI-washing should be less prevalent when issuers face stronger prior regulatory scrutiny. In contrast, there is no corresponding reason for pre-filing scrutiny to predict AI investment, which is defined using observable human-capital and patent inputs. This asymmetric prediction provides a discriminant validation of the AI-washing measure.

To examine this prediction, we use logit regression to investigate whether regulatory scrutiny in the IPO market is associated with firms' propensity to engage in AI-washing:

$$AI_Washing/AI_Investment_{i,t} = \beta_1 ComLetter_{i,t} + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (4)$$

where the dependent variable is either *AIWashing*, a dummy variable equal to one for AI-washing IPOs or *AIInvestment,* a dummy variable equal to one for AI-investment IPOs. The focal explanatory variable, *ComLetter*, is the log of one plus the number of SEC comment letters related to IPO filings issued between one year and six months prior to the firm's filing date. $CONTROLS_i$ include the firm-level and IPO-level characteristics defined in Appendix A, and $FE_{i,t}$ represents

two-digit SIC industry and IPO year fixed effects. Standard errors are adjusted for heteroskedasticity and clustered by industry.

Table 3 reports the regression results. Columns (1) and (2) use *AIWashing* as the dependent variable, while Columns (3) and (4) use *AIInvestment* as the dependent variable. The coefficients on *ComLetter* are negative and statistically significant in both specifications. In column (2), which includes the full set of controls, the coefficient is -0.249. This result suggests that stronger prior regulatory scrutiny in the IPO market is associated with a lower likelihood of AI-washing. By contrast, *ComLetter* is not significantly associated with AIInvestment in either specification.

Taken together, these findings show that regulatory scrutiny is negatively associated with AI-washing but not with AI investment. The asymmetric association provides institutional validation of the AI-washing measure and is consistent with stronger oversight constraining aggressive AI-related disclosure.

[Table 3 here]

4 AI-Washing and Post-IPO Performance

4.1 AI-washing IPOs long-term returns

4.1.1 Baseline findings

In the IPO setting, whether the distinction between AI-washing and AI investment is reflected in IPO pricing and subsequent stock returns. We consider both initial returns and long-run abnormal returns. We estimate the following OLS specification:

$$RET_{i,t} = \beta_1 AIWashing_i + \beta_2 AIInvestment_i + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (5)$$

where i and t refer to firm and year, respectively. The dependent variable, $RET_{i,t}$, comprises both short-term and long-term return measures. Specifically, it includes the initial return ($InitialRet_{i,t}$) and style-adjusted buy-and-hold abnormal returns ($BHAR_{i,t}$) measured over one-year (*BHAR[0,1yr]*), two-year (*BHAR[0,2yr]*), and three-year (*BHAR[0,3yr]*) horizons following

the IPO. The main explanatory variables are *AIWashing* and *AIInvestment*, which indicate whether an IPO is AI-washing or AI investment. The control variables include a set of IPO characteristics and firm fundamentals that have been shown to affect post-IPO performance in the literature. Specifically, we control for high-reputation underwriters (*HighUW*), venture capital backing (*VC*), and Big 4 auditors (*Big4*) to capture the certification role of financial intermediaries. We also include IPO issuance characteristics such as gross proceeds (*Proceeds*), the proportion of secondary shares (*Second*), pure primary offerings (*Primary*), carve-out IPOs (*Carve*), and the overhang ratio (*Overhang*), which reflect ownership structure and supply conditions in the offering. At the firm level, we control for firm age (*Age*), pre-IPO assets (*Assets*), and pre-IPO sales (*Sales*) to capture firm size and maturity. We further include indicators for *HighTech*, *Biotech*, and *Internet* industries to account for sectoral differences in growth opportunities, as well as pre-IPO profitability (*PositiveEPS*) and initial IPO returns (*InitialReturn*) to control for firm performance and investor demand at the time of listing. For short-term returns, as the dependent variable is the initial IPO returns (*InitialReturn*), we replace the control with price revision during the IPO process (*PrcRev*). We include two-digit SIC industry fixed effects and year fixed effects ($FE_{i,t}$) to control for industry-specific and time-varying factors that may affect firms' disclosure behavior (Liu and Wu 2021). Standard errors are adjusted for heteroskedasticity and clustered by industry.

[Table 4 here]

Table 4 Panel A reports the baseline results. Columns (1)–(3) present regressions using one-, two-, and three-year buy-and-hold post-IPO abnormal returns (*BHAR*) as the dependent variables, while Column (4) examines IPO initial returns. The coefficients on *AIWashing* are negative across all long-run specifications and become statistically significant over longer horizons, indicating that AI-washing firms underperform otherwise comparable IPO firms by 8.8 percentage points over two years and 13.9 percentage points over three years. In contrast, the coefficient on *AIInvestment* is positive and statistically significant in the two- and three-year specifications, suggesting that

firms with observable AI investments outperform their peers by 10.2 percentage points and 22.0 percentage points, respectively. The difference between the coefficients on *AIWashing* and *AIInvestment* is statistically significant at the 1% level in both specifications. The baseline findings are also robust to alternative benchmark-adjusted BHAR measures based on value-weighted and equal-weighted market returns, as well as to alternative definitions of AI disclosure thresholds (Internet Appendix Table IA6 and Internet Appendix Table IA7).

By contrast, Column (4) shows that *AIWashing* is positively associated with IPO initial returns, whereas *AIInvestment* has no significant effect, suggesting that AI-washing narratives are rewarded at the IPO stage and investors can not distinguish AI-washing IPOs from AI-investment IPOs in the initial stage. As an alternative measure of AI washing, we use the residual from Equation (1) as a continuous AI-washing index. The results, reported in Internet Appendix Table IA8, are qualitatively unchanged and confirm the robustness of our baseline findings.

Figure 5 plots the style-adjusted buy-and-hold abnormal returns over the three years following the IPO. AI-washing IPOs exhibit a persistent decline in abnormal returns throughout the post-IPO period, whereas AI-investment IPOs experience a gradual improvement and ultimately generate positive abnormal returns. The performance gap between the two groups widens steadily over time, providing graphical evidence consistent with the baseline regression results. Overall, the evidence indicates that the AI disclosure–input mismatch, our empirical proxy for AI-washing, predicts a significant long-run underperformance post-IPO. This pattern is consistent with a gradual deflation of AI-driven premiums post-listing, as the public market progressively uncovers firms' underlying AI technological realities and discounts unbacked disclosures.

[Figure 5 here]

As shown in Panel C1 and C2 of Table 1, differences between serveal characteristics among groups are significant. To further alleviate concerns that the baseline results are driven by observable differences between AI-washing and other IPO firms, we repeat the analysis using

propensity score matching (PSM) and entropy balancing. Internet Appendix Figure IA2 shows that the matching procedure substantially improves covariate balance, with the standardized bias of nearly all matching variables falling within the conventional 10% threshold after matching. As reported in Panel B of Table 4, the results remain qualitatively unchanged. Under propensity score matching, AI-washing IPOs underperform matched firms by 16.1 and 18.8 percentage points over the two- and three-year horizons, respectively. Entropy balancing yields economically similar estimates of 12.8 and 19.6 percentage points. At the same time, *AIInvestment* continues to predict superior long-run performance, particularly over the three-year horizon. The consistency of the estimates across both balancing approaches suggests that the baseline findings are unlikely to be driven by observable differences in firm characteristics, IPO attributes, or intermediary certification.

4.1.2 Calendar-time portfolio tests

A potential concern with buy-and-hold abnormal returns is that statistical inference may be affected by cross-sectional dependence among IPOs and overlapping return horizons (Fama 1998; Mitchell and Stafford 2000). To address this concern, we complement with calendar-time portfolio regressions. Specifically, in each calendar month, we form portfolios consisting of AI-washing IPOs that remain within the designated post-IPO holding window and estimate abnormal portfolio returns using the CAPM, Fama–French three-factor (FF3), and Fama–French five-factor (FF5) models. The intercept from each regression represents the monthly abnormal return after adjusting for common risk factors.

Table 5 reports the results. Panel A shows that AI-washing IPO portfolios generate significantly negative abnormal returns across all asset pricing models. For the 24-month holding window, the estimated monthly alpha ranges from −1.49% to −1.95%, depending on the factor specification. Similar results are obtained for the 36-month holding window, where monthly abnormal returns range from −1.13% to −1.69%. These estimates remain statistically significant

after controlling for size, value, profitability, and investment factors, suggesting that the long-run underperformance of AI-washing IPOs cannot be explained by their exposure to standard systematic risk factors.

Panels B and C further compare AI-washing IPOs with alternative benchmark portfolios. Panel B constructs long–short portfolios that are long AI-washing IPOs and short AI-investment IPOs, while Panel C uses all non-AI IPOs as the benchmark portfolio. Across both portfolio constructions, the estimated alphas remain negative and statistically significant under all three asset pricing models. In particular, the monthly abnormal return ranges from −1.26% to −1.65% relative to AI-investment IPOs and from −0.85% to −1.30% relative to non-AI IPOs. These findings show that the post-IPO underperformance of AI-washing firms persists in calendar-time portfolio tests and is not attributable to conventional risk factors or the methodology used to measure long-run abnormal returns.

[Table 5 here]

4.1.3 Accounting for external AI capabilities

Our *AIWashing* measure identifies firms whose AI-related disclosures are not supported by observable internal AI investment. A potential concern, however, is that some firms may acquire AI capabilities through external providers rather than building them in-house. As emphasized by Babina (2026), firms' AI efforts encompass multiple dimensions, including internal AI development and the adoption of outsourced AI solutions, and different measures capture different aspects of firms' AI capabilities. Consequently, firms relying primarily on external AI vendors or strategic collaborations may be incorrectly classified as AI-washing if their AI capabilities are not reflected in our internal AI investment measure.

To address this concern, Table 6 successively excludes AI-washing IPOs with observable third-party AI activities and re-estimates the baseline regressions. Columns (1) and (2) remove firms whose prospectuses explicitly disclose outsourced AI development or the use of third-party

AI tools and services based on keyword searches. Columns (3) and (4) further exclude firms identified as AI customers using FactSet Supply Chain data, while Columns (5) and (6) remove firms with identifiable AI collaboration partners based on RavenPack news data.

[Table 6 here]

The results remain remarkably stable across all specifications. After excluding these firms, the coefficient on *AIWashing* continues to be negative and statistically significant for both the two-year and three-year buy-and-hold abnormal returns. In contrast, *AIInvestment* remains positively associated with subsequent long-run stock performance, and the difference between the coefficients on *AIWashing* and *AIInvestment* remains statistically significant in every specification. The economic magnitudes are also comparable to those reported in the baseline regressions.

Overall, these findings indicate that our results are unlikely to be driven by firms that obtain AI capabilities through observable outsourcing arrangements, supplier relationships, or strategic AI partnerships. Instead, the evidence suggests that the inferior long-run performance is concentrated among firms whose AI-related disclosures are unsupported by either internal AI investment or identifiable external AI activities. This strengthens the construct validity of our AI-washing measure.

4.2 Mechanisms: Why do AI-washing IPOs perform worse than others in the long-run?

In this section, we focus on the economic mechanisms underlying the long-run stock price reversals of AI-washing IPOs. We focus on three complementary channels: *valuation premium channel*, *verification channel* and *litigation risk channel*. Although we examine these mechanisms separately, they are not mutually exclusive; rather, they represent different stages of the same valuation cycle through which AI are initially capitalized into IPO valuations and later corrected as information is resolved.

4.2.1 Valuation premium channel

Innovation generally enhances firm value by improving firms' productivity and expanding future growth opportunities. Artificial intelligence represents a particularly distinctive form of innovation because, as a general-purpose technology, it can be deployed across a broad range of business activities, combined with firms' existing data and organizational capabilities, and scaled at relatively low marginal cost, allowing firms to generate broad and potentially transformative growth opportunities (Bresnahan and Trajtenberg 1995; Cockburn, Henderson, and Stern 2018; Aghion, Jones, and Jones 2017). Consequently, firms undertaking valuable innovation tend to command higher market valuations because AI-investing firms experience higher sales growth, employment growth, and market valuations through product innovation (Babina et al. 2024).

However, AI-washing IPOs may also benefit from the valuation premium attached to AI innovation. Despite lacking substantive AI capabilities, these firms may receive similarly favorable valuations at issuance when investors do not fully distinguish AI-related disclosures from underlying AI capabilities. Consequently, AI-washing IPOs are expected to receive higher valuations at issuance, followed by weaker long-run stock performance as these initially elevated valuations are corrected over time.

To further examine this mechanism, we exploit cross-industry variation in the value creation of AI. Because the value realization of AI depends on complementary assets, organizational adjustments, and application-specific implementation, the economic value generated by AI technologies is unlikely to be uniform across industries (Brynjolfsson, Rock, and Syverson 2017). In industries where AI is expected to create greater value, investors may assign a higher valuation premium to AI disclosures. If the long-run underperformance of AI-washing IPOs reflects investors' over-extrapolation of AI value, the effect should be stronger in industries with higher AI valuation premiums.

Following this logic, we construct a measure of AI value creation across industries (Hall, Jaffe, and Trajtenberg 2005; Kogan et al. 2017). First, for each firm-year observation, we construct AI

patent stock (*AIStock*) as the depreciated stock of AI patents accumulated over the past ten years. Specifically, it is calculated as:

$$AIStock_{i,t-1} = \sum_{k=0}^{10} (1 - \delta)^k \, AIPatNum_{i,t-k} \tag{6}$$

where $AIPatNum_{i,t-k}$ denotes the number of AI patents filed by firm i in year $t-k$, and δ represents the annual depreciation rate of AI capital. We set the depreciation rate to 15%, implying a retention rate of 0.85.

Second, within each Fama–French 12 industry group, we estimate the extent to which AI innovation is capitalized into firm valuation. Specifically, for each industry, we regress firm market value on AI patent stock using firm-year observations over the three years preceding the IPO:

$$MV_{i,t} = \beta_{IndAIVal} AIStock_{i,t-1} + CONTROLS_{i,t-1} + FE_{i,t} + \varepsilon_{i,t} \tag{7}$$

where $MV_{i,t}$ denotes the log of firm market value. $CONTROLS_{i,t-1}$ include firm size ($\ln(AT)_{i,t-1}$), profitability ($\mathrm{ROA}_{i,t-1}$), R&D intensity ($XRD/AT_{i,t-1}$), cash holdings ($Cash/AT_{i,t-1}$), advertising intensity ($XAD/AT_{i,t-1}$) and sales growth ($\Delta ln(Sales)_{i,t-1}$). We include fiscal-year fixed effects ($FE_{i,t}$) within each estimation window and cluster standard errors by fiscal year.

The estimated coefficient $\beta_{IndAIVal}$ valuation sensitivity to AI innovation within each Fama–French 12 industry-year window and serves as our measure of industry AI value creation, denoted as *AIVal*. We then estimate whether the valuation and subsequent long-run performance of AI-washing IPOs vary with the industry AI valuation premium using the following specification:

$$\begin{aligned} Y_{i,t} = {} & \beta_1 AIWashing \times AIVal_{i,t} + \beta_2 AIWashing_i + \beta_3 AIInvestment_i \\ & + \beta_4 IndAIVal_{i,t} + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{8}$$

where *Y* includes IPO price-to-sales ratio (*P/S*), the two-year *BHAR*, or the three-year *BHAR*. The key variable of interest is the interaction term *AIWashing* × *AIVal*, which captures whether the long-run performance of AI-washing IPOs varies with the extent to which AI innovation is valued in the firm's industry.

The results are reported in Table 7. Column (1) and (2) use price-to-sales ratio (*P/S*) as the dependent variable, measured by Purnanandam and Swaminathan (2004). The interaction term *AIWashing* × *AIVal* is positive and statistically significant in Column (2), indicating that AI-washing IPOs receive higher valuation multiples in industries where investors place larger valuations on genuine AI innovation. This finding is consistent with the prediction that AI-washing firms benefit more from the valuation premium attached to AI innovation.

Columns (3)–(6) examine whether the higher IPO valuations documented above are subsequently corrected in the aftermarket. The interaction term *AIWashing* × *AIVal* is negative and statistically significant for both two-year and three-year buy-and-hold abnormal returns. For two-year abnormal returns, the estimated coefficient equals −6.053 in Column (4). Similar evidence is obtained for three-year abnormal returns, where the estimated coefficients is −5.407 significantly at 10% level. Together with the evidence on IPO valuation, these results indicate that the higher valuations initially enjoyed by AI-washing IPOs are followed by stronger long-run return reversals in industries where investors place larger valuations on AI innovation.

Overall, the results support the AI value creation channel: when an industry places a higher valuation premium on AI-related innovation, AI-washing firms appear more able to benefit from inflated AI narratives at the IPO stage but subsequently experience stronger long-run performance reversals.

[Table 7 here]

4.2.2 Verification channel

Genuine AI capability is particularly difficult for external investors to verify because it is rarely embodied in a single, observable asset. Instead, effective AI adoption requires the integration of proprietary data, specialized human capital, and complementary organizational capabilities. In particular, data accumulated through firm activities are inherently firm-specific and largely unobservable to external investors (Veldkamp 2023). Moreover, AI alters the innovation

process by combining data, algorithms, and domain expertise, making its realized value contingent on how these complementary assets are integrated within the organization (Cockburn, Henderson, and Stern 2018; Abis and Veldkamp 2023).

These characteristics create a verification challenge in the IPO market , where investors have limited operating histories and rely heavily on prospectus disclosures when evaluating firms' technological capabilities. As a result, investors cannot easily distinguish substantive AI investment from symbolic AI narratives at issuance. When firms' underlying AI capabilities are more difficult to verify, investors are more likely to mistake symbolic AI claims for substantive AI capabilities, leading to more severe initial mispricing. Consequently, AI-washing IPOs facing greater verification difficulty are expected to exhibit poorer long-run stock performance when this initial mispricing is eventually corrected.

To empirically test the AI verification difficulty channel, we construct a measure of Business–AI Distance (*BusinessAIDistance*) using a text-based semantic distance approach. Specifically, we extract the Business section from each IPO firm's S-1 prospectus and embed the text using a sentence-transformer model. We compute the cosine similarity between each IPO firm's Business section embedding and the centroid embedding of the AI patent corpus.[11] Business–AI Distance is defined as one minus this cosine similarity. A higher value indicates that the firm's core business is semantically farther from the AI technological domain.

To test this mechanism, we estimate the following regression:

$$BHAR_{i,t} = \beta_1 AIWashing \times AIDistance_{i,t} + \beta_2 AIDistance_{i,t} + \beta_3 AIWashing_i + \beta_4 AIInvestment_i + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (9)$$

where the key variable of interest is the interaction term between *AIWashing* and *AIDistance*. The results are reported in Table 8. Columns (1) and (2) examine two-year abnormal returns, while Columns (3) and (4) examine three-year abnormal returns. The coefficient on the interaction term

[11] We use the sentence-transformers/all-mpnet-base-v2 model to convert each firm's Business section and each AI patent text into dense embedding vectors.
We construct an AI reference corpus using AI-related patent texts from the istat-ai/ai-patents dataset on HuggingFace, which contains patents related to artificial intelligence, machine learning, deep learning, neural networks, natural language processing, and large language models, available at https://huggingface.co/datasets/istat-ai/ai-patents.

AIWashing × *AIDistance* is negative and statistically significant in both specifications. In Column (2), the estimated coefficient is −1.313 and is significant at the 5% level. Similarly, in Column (4), the coefficient equals −1.425 and is significant at the 10% level. These results indicate that the long-run underperformance of AI-washing IPOs is more pronounced when firms' underlying AI capabilities are more difficult to verify.

The coefficient on *AIDistance* itself is not statistically significant, suggesting that verification difficulty alone does not independently predict long-run IPO performance. Rather, verification difficulty matters primarily in conjunction with AI-washing. This finding is consistent with the verification channel: when firms' underlying AI capabilities are more difficult to verify, investors are less able to distinguish symbolic AI claims from substantive AI capabilities, resulting in more severe initial mispricing. As this initial mispricing is eventually corrected, AI-washing IPOs experience poorer long-run stock performance.

[Table 8 here]

4.2.3 Litigation risk channel

The persistence of the initial mispricing associated with AI-washing depends on the expected cost of maintaining unsupported AI narratives after the IPO. While AI-washing may generate short-term valuation benefits, these benefits become increasingly difficult to sustain once firms enter the public market. As firms become increasingly accountable for the accuracy of their AI-related disclosures after going public, unsupported AI narratives expose issuers to growing legal costs. Consequently, higher expected litigation risk increases the likelihood that the initial mispricing associated with AI-washing will eventually be corrected.

This mechanism is reinforced by the institutional features of the U.S. IPO market. Because the Private Securities Litigation Reform Act of 1995 (PSLRA) safe harbor for forward-looking statements does not apply to IPO-related communications, firms face heightened legal exposure for misleading disclosures made during the offering process (Johnson, Kasznik, and Nelson

2001).[12] Consequently, unsupported AI narratives disclosed during the IPO can expose issuers to substantial legal liability if subsequent evidence fails to substantiate those claims (Field, Lowry, and Shu 2005). Anticipating these potential costs, firms have weaker incentives to maintain unsupported AI narratives once they become publicly accountable.

After listing, firms enter a continuous mandatory disclosure regime that progressively reveals information about their technological development and business activities (Healy and Palepu 2001). Periodic reporting obligations, including Forms 10-K, 8-K, earnings conference calls, and realized operating performance, provide increasingly verifiable evidence against which investors, regulators, and litigants can assess the credibility of earlier AI disclosures (Leuz and Wysocki 2016). As this evidence accumulates, unsupported AI narratives become increasingly difficult to sustain. Accordingly, AI-washing firms are more likely to scale back unsupported AI-related disclosures over time, thereby narrowing the gap between symbolic AI narratives and substantive AI investment and facilitating the eventual correction of the initial mispricing (Skinner 1994; Rogers and Van Buskirk 2009).

To examine this mechanism, we employ two alternative measures of estimated ex-post litigation risk one year after the IPO. The first measure follows Kim and Skinner (2012).[13] The litigation risk proxy is estimated by the following model:

$$\begin{aligned} LitRisk_t = {} & 0.567(FPS_t) + 0.523(LNASSET_{t-1}) + 0.896(SALES\ GROWTH_{t-1}) \\ & +0.419(RETURN_{t-1}) - 0.101(RETURN\ SKEWNESS_{t-1}) \\ & +25.254(RETURN\ STD\ DEV_{t-1}) + 0.00004(TURNOVER_{t-1}) \end{aligned} \quad (10)$$

where the independent variables include high-litigation-risk industries (*FPS*), firm size (*LNASSET*), sales growth (*SALES GROWTH*), one-year stock return (*RETURN*), one-year return skewness (*RETURN SKEWNESS*), one-year return volatility (*RETURN STD DEV*), and one-year stock turnover (*TURNOVER*), measured from the listing date to one-year after the listing. The

[12] See John Coates, Acting Director, Division of Corporation Finance, U.S. Securities and Exchange Commission, "SPACs, IPOs and Liability Risk under the Securities Laws" (Apr. 8, 2021), noting that the PSLRA safe harbor for forward-looking statements is unavailable in traditional IPOs: https://www.sec.gov/newsroom/speeches-statements/spacs-ipos-liability-risk-under-securities-laws

[13] The first measure is used as a relative proxy for litigation risk rather than as a predicted probability, so the intercept term from the original model is omitted.

second measure, following Donelson, Flam, and Yust (2022), we proxy for industry-level litigation risk using the number of disclosure-related securities class actions filed within the firm's Fama–French 12 industry over the first year following the IPO.

Second, to examine the role of litigation risk, we incorporate the litigation risk measure into the baseline regression. Our primary variable of interest is the interaction term *AIWashing* × *LitRisk*, which captures whether the post-IPO performance of AI-washing firms varies with the level of expected litigation risk. The estimation specification is as follows:

$$BHAR[1,2yr]_{i,t} = \beta_1 AIWashing \times LitRisk_{i,t} + \beta_2 AIWashing_i + \beta_3 AIInvestment_i + \beta_4 LitRisk_{i,t} + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (11)$$

where the regressions include the same set of control variables in the baseline OLS regressions, together with industry and year fixed effects. Table 9 reports the results for the litigation risk channel. Across both specifications, the interaction between AIWashing and litigation risk is significantly negative. Specifically, the coefficients on *AIWashing* × *LitRiskKS* and *AIWashing* × *LitRiskDFY* are −0.034 (t = −3.10) and −0.048 (t = −2.57), respectively. These results indicate that AI-washing IPOs experience poorer second-year abnormal stock performance when expected litigation risk is higher, consistent with the prediction that greater expected litigation costs accelerate the correction of the initial mispricing associated with AI-washing.

[Table 9 here]

To provide direct evidence on the disclosure adjustment underlying the litigation risk channel, we next examine whether litigation risk influences firms' post-IPO AI disclosure behavior. We construct two measures of the post-IPO AI disclosure gap (*AIGap*). The first measure captures the change in AI-related disclosures between the IPO prospectus (S-1) and the first annual report (Form 10-K), calculated as the number of AI-related words in the Form 10-K minus those in the S-1 prospectus. The second measure captures the change in AI disclosure intensity between earnings conference calls during the second post-IPO year (from the fifth to the eighth fiscal quarter) and the S-1 prospectus. More negative values of AIGap indicate larger reductions in AI-

related disclosures following the IPO. The AI words use the same keyword list in Internet Appendix Table IA2.

We estimate the following regression:

$$AIGap_{i,t} = \beta_1 AIWashing \times LitRisk_i + \beta_2 AIWashing_i + \beta_3 AIInvestment_i + \beta_4 LitRisk_i + CONTROLS_i + FE_{i,t} + \varepsilon_{i,t} \quad (12)$$

where the regressions include the same set of control variables in the baseline OLS regressions, together with industry and year fixed effects.

Internet Appendix Table IA9 reports the results. Panel A measures the post-IPO AI disclosure gap using the difference in AI-related disclosures between the first Form 10-K and the IPO prospectus (*AIGapTenK*), while Panel B measures the disclosure gap using earnings conference calls relative to the prospectus (*AIGapCall*). Across both disclosure measures, the interaction between AI-washing and litigation risk is consistently negative under both *LitRiskKS* and *LitRiskDFY*. These findings indicate that AI-washing firms are more likely to scale back unsupported AI-related disclosures when expected litigation risk is higher, thereby narrowing the gap between symbolic AI narratives and substantive AI investment.

Overall, the evidence supports the litigation risk mechanism. When expected legal exposure is higher, firms that relied heavily on AI disclosure during the IPO process are more likely to scale back such narratives in subsequent disclosures. Consistent with this disclosure adjustment, AI-washing firms also experience more pronounced long-term post-IPO underperformance when litigation risk is elevated.

5 Conclusion

This paper examines how disclosures about emerging technologies affect valuation in the IPO market. We focus on artificial intelligence, a general-purpose technology with substantial economic potential but underlying capabilities that are difficult for outside investors to verify. Using a measurement framework that compares AI-related disclosure in S-1 registration statements with observable AI inputs from employee profiles and patent filings, we identify issuers

that emphasize AI in their prospectuses despite having no observable AI-related human capital or patenting activity.

We find that this disclosure-input mismatch is economically meaningful. Over longer horizons, AI-washing firms significantly underperform both non-AI IPOs and IPOs with verifiable AI investment. The evidence is consistent with a correction of AI-related valuation premiums as public-market investors gradually learn about firms' underlying technological capabilities after listing. Cross-sectional evidence further supports this interpretation. The underperformance of AI-washing firms is stronger in industries where AI innovation commands larger valuation premiums and when firms' core businesses are farther from the AI technological domain. In addition, AI-washing firms facing higher expected litigation risk subsequently scale back AI-related disclosure in post-IPO filings, consistent with legal exposure constraining the continuation of aggressive technology narratives after going public.

Our findings also have implications for securities regulators such as the SEC. The evidence suggests that AI-related claims in IPO prospectuses may warrant greater ex ante scrutiny, because unsupported technology narratives can be incorporated into valuation before investors are able to verify firms' underlying capabilities. Ex post disclosure corrections, litigation, or enforcement actions may help discipline misleading claims, but they are unlikely to fully reverse valuation distortions or investor losses once inflated narratives have affected IPO pricing.[14] A more effective disclosure review process may therefore place greater emphasis on whether AI claims are specific, verifiable, and consistent with observable technological inputs.

[14] As shown in Internet Appendix IA10, AI expertise among venture capitalists, underwriters, and auditors is not significantly associated with a lower likelihood of AI-washing, suggesting that traditional market-based discipline alone may be insufficient to deter unsupported AI-related disclosures at the IPO stage.

Reference

Abis, Simona, and Laura Veldkamp, 2023, The Changing Economics of Knowledge Production, *The Review of Financial Studies* 37, 89–118.

Aggarwal, Rajesh K, Laurie Krigman, and Kent L Womack, 2002, Strategic IPO underpricing, information momentum, and lockup expiration selling, *Journal of Financial Economics* 66, 105–137.

Aghion, Philippe, Benjamin F Jones, and Charles I Jones, 2017, Artificial intelligence and economic growth, *NBER Working Paper* 23928.

Babina, Tania, 2026, Understanding Firms' AI Efforts and Their Economic Impact, *NBER Working Paper 35123*.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson, 2024, Artificial intelligence, firm growth, and product innovation, *Journal of Financial Economics* 151, 103745.

Baker, Andrew C., David F. Larcker, Charles G. McClure, Durgesh Saraph, and Edward M. Watts, 2024, Diversity washing, *Journal of Accounting Research* 62, 1661–1709.

Barber, Brad M., and John D. Lyon, 1997, Detecting long-run abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics* 43, 341–372.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, 2021, The diffusion of disruptive technologies, NBER Working Paper.

Bresnahan, Timothy F., and Manuel Trajtenberg, 1995, General purpose technologies 'Engines of growth'?, *Journal of Econometrics* 65, 83–108.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, 2017, Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics, *NBER Working Paper* 24001.

Cheng, Stephanie F., Gus De Franco, Haibo Jiang, and Pengkai Lin, 2019, Riding the blockchain mania: public firms' speculative 8-K disclosures, *Management Science* 65, 5901–5913.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern, 2018, The impact of artificial intelligence on innovation: An exploratory analysis, *NBER Working Paper* 24449.

Delmas, Magali A., and Vanessa Cuerel Burbano, 2011, The drivers of greenwashing, *California Management Review* 54, 64–87.

Donelson, Dain C., Rachel W. Flam, and Christopher G. Yust, 2022, Spillover Effects in Disclosure-Related Securities Litigation, *The Accounting Review* 97, 275–299.

Duchin, Ran, Janet Gao, and Qiping Xu, 2025, Sustainability or greenwashing: Evidence from the asset market for industrial pollution, *Journal of Finance* 80, 699–754.

Fama, Eugene F, 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283—306.

Field, Laura, Michelle Lowry, and Susan Shu, 2005, Does disclosure deter or trigger litigation?, *Journal of Accounting and Economics* 39, 487–507.

Giczy, Alexander V., Nicholas A. Pairolero, and Andrew A. Toole, 2022, Identifying artificial intelligence (AI) invention: A novel AI patent dataset, *Journal of Technology Transfer* 47, 476–505.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, 2005, Market value and patent citations, *RAND Journal of Economics* 36, 16–39.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2010, The information content of IPO prospectuses, *Review of Financial Studies* 23, 2821–2864.

Healy, Paul M, and Krishna G Palepu, 2001, Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature, *Journal of Accounting and Economics* 31, 405–440.
Henderson, Rebecca, 2018, The impact of artificial intelligence on innovation, NBER Working Paper.
Johnson, Marilyn F., Ron Kasznik, and Karen K. Nelson, 2001, The impact of securities litigation reform on the disclosure of forward-looking information by high technology firms, *Journal of Accounting Research* 39, 297–327.
Kim, Irene, and Douglas J. Skinner, 2012, Measuring securities litigation risk, *Journal of Accounting and Economics* 53, 290–310.
Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *Quarterly Journal of Economics* 132, 665–712.
Leuz, Christian, and Peter D. Wysocki, 2016, The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research, *Journal of Accounting Research* 54, 525–622.
Liu, Xiaotao (Kelvin), and Biyu Wu, 2021, Do IPO firms misclassify expenses? Implications for IPO price formation and post-IPO stock performance, *Management Science* 67, 4505–4531.
Ljungqvist, Alexander, and William J. Wilhelm, 2003, IPO pricing in the Dot-com bubble, *Journal of Finance* 58, 723–752.
Lou, Bowen, and Lynn Wu, 2021, AI on drugs: Can artificial intelligence accelerate drug development? Evidence from a large-scale examination of Bio-Pharma firms, *MIS Quarterly* 45, 1451–1482.
Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.
Lowry, Michelle, Roni Michaely, and Ekaterina Volkova, 2017, Initial public offerings: A synthesis of the literature and directions for future research, *Foundations and Trends in Finance* 11, 154–320.
Luong, Hoang, Fariborz Moshirian, Lily Nguyen, Xuan Tian, and Bohui Zhang, 2017, How do foreign institutional investors enhance firm innovation?, *Journal of Financial and Quantitative Analysis* 52, 1449–1490.
Mitchell, Mark L., and Erik Stafford, 2000, Managerial Decisions and Long-Term Stock Price Performance, *The Journal of Business* 73, 287–329.
Ofek, Eli, and Matthew Richardson, 2003, DotCom mania: The rise and fall of Internet stock prices, *Journal of Finance* 58, 1113–1137.
Purnanandam, Amiyatosh K., and Bhaskaran Swaminathan, 2004, Are IPOs Really Underpriced?, *Review of Financial Studies* 17, 811–848.
Qian, Yiming, Xinjian Shao, and Jingchi Liao, 2024, Pre-IPO hype by affiliated analysts: Motives and consequences, *Journal of Corporate Finance* 89, 102648.
Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3–27.
Ritter, Jay R., and Ivo Welch, 2002, A Review of IPO Activity, Pricing, and Allocations, *Journal of Finance* 57, 1795–1828.
Rogers, Jonathan L., and Andrew Van Buskirk, 2009, Shareholder litigation and changes in disclosure behavior, *Journal of Accounting and Economics* 47, 136–156.
Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* 32, 38–60.

Teoh, Siew Hong, Ivo Welch, and T.J. Wong, 1998, Earnings management and the long-run market performance of initial public offerings, *The Journal of Finance* 53, 1935–1974.
Veldkamp, Laura, 2023, Valuing Data as an Asset, *Review of Finance* 27, 1545–1562.
WIPO, 2019, WIPO technology trends 2019: Artificial intelligence, *Geneva: World Intellectual Property Organization*.
Wu, Lynn, Bowen Lou, and Lorin M. Hitt, 2024, Innovation strategy after IPO: How AI analytics spurs innovation after IPO, *Management Science* 71, 2360–2389.

Appendix A. Variable Definitions

Variable	Acronym	Definition	Time Period	Data Source
AI words	*AIWord*	Total number of occurrences of the 57 AI-related terms in the firm's S-1 registration statement.	2010–2024	SEC EDGAR (via WRDS)
AI patents	*AIPatent*	Number of patent filings classified as AI-related by the Artificial Intelligence Patent Dataset, including granted patents and pre-grant publications, from three years before the IPO through the listing date.	2010–2023	USPTO AIPD; KPSS patent links
AI employees	*AIEmployee*	Number of employees as of the IPO listing date who satisfy the AI-skill classification, excluding interns and applying the STEM-education and AI-skill-share screens.	2010–2024	Revelio Labs (via WRDS)
AI-washing IPOs	*AIWashing*	Indicator equal to one if *AIWord* exceeds the 95th percentile among IPO firms in the same Fama–French 12 industry, calculated using an expanding window from the beginning of the sample through the year preceding the IPO, and both *AIEmployee* and *AIPatent* equal zero, and zero otherwise.	2010–2024	SEC EDGAR; Revelio Labs; USPTO AIPD
AI-investment IPOs	*AIInvestment*	Indicator equal to one if the firm has at least one AI employee or one AI patent filing before the IPO, and zero otherwise.	2010–2024	Revelio Labs; USPTO AIPD
AI-washing index	*AIWashingIndex*	Standardized residual from a Poisson regression of *AIWord* on *AIEmployee* and AIPatent, including two-digit SIC industry and IPO-year fixed effects. Higher values indicate AI disclosure that is large relative to observable AI inputs.		
First-day return	*InitialReturn*	First-day closing price minus the offer price, divided by the offer price.	2010–2024	CRSP; SDC Platinum
One-year abnormal return	*BHAR[0,1yr]*	Buy-and-hold return from the first trading day through the first post-IPO year minus the buy-and-hold return on the corresponding Fama–French 5×5 size–book-to-market portfolio.	2010–2024	CRSP; Kenneth French Data Library
Second-year abnormal return	*BHAR[1,2yr]*	Buy-and-hold return from the end of the first post-IPO year through the end of the second year minus the buy-and-hold return on the corresponding Fama–French 5×5 size–book-to-market portfolio.	2010–2024	CRSP; Kenneth French Data Library
Two-year abnormal return	*BHAR[0,2yr]*	Buy-and-hold return from the first trading day through the second post-IPO year minus the buy-and-hold return on the corresponding Fama–French 5×5 size–book-to-market portfolio.	2010–2024	CRSP; Kenneth French Data Library

Variable	Acronym	Definition	Time Period	Data Source
Three-year abnormal return	*BHAR[0,3yr]*	Buy-and-hold return from the first trading day through the third post-IPO year minus the buy-and-hold return on the corresponding Fama–French 5×5 size–book-to-market portfolio.	2010–2024	CRSP; Kenneth French Data Library
High-reputation underwriter	*HighUW*	Indicator equal to one if the lead underwriter is classified as high reputation using Jay Ritter's underwriter rankings, and zero otherwise.	2010–2024	Jay Ritter's website
Venture capital backing	*VC*	Indicator equal to one if the IPO firm is venture capital backed, and zero otherwise.	2010–2024	Jay Ritter's website
Big 4 auditor	*Big4*	Indicator equal to one if the IPO firm's auditor is Deloitte, Ernst & Young, KPMG, or PricewaterhouseCoopers, and zero otherwise.	2010–2024	SDC Platinum
Gross proceeds	*Proceeds*	Natural logarithm of gross proceeds, calculated as the offer price multiplied by the number of shares offered.	2010–2024	SDC Platinum
Firm age	*Age*	Natural logarithm of one plus the firm's age in years at the IPO date.	2010–2024	Jay Ritter's website; SDC Platinum
Pre-IPO assets	*Assets*	Natural logarithm of pre-IPO total assets measured in USD millions.	2010–2024	SDC Platinum
Pre-IPO sales	*Sales*	Natural logarithm of one plus pre-IPO sales measured in CPI-adjusted USD millions.	2010–2024	SDC Platinum; U.S. Bureau of Labor Statistics
Carve-out	*Carve*	Indicator equal to one if the IPO is an equity carve-out, and zero otherwise.	2010–2024	SDC Platinum
High-tech firm	*HighTech*	Indicator equal to one if the IPO firm operates in a high-technology industry, and zero otherwise.	2010–2024	SDC Platinum
Biotech firm	*Biotech*	Indicator equal to one if the IPO firm has an SIC code from 2830 through 2839 or 8731, and zero otherwise.	2010–2024	CRSP–Compustat
Internet firm	*Internet*	Indicator equal to one if the IPO firm is classified as an Internet company, and zero otherwise.	2010–2024	Jay Ritter's website
Price revision	*PrcRev*	Offer price minus the midpoint of the initial filing range, divided by that midpoint.	2010–2024	SDC Platinum
Overhang	*Overhang*	Shares retained by pre-IPO shareholders divided by shares offered, calculated as (post-IPO shares outstanding minus shares offered) divided by shares offered.	2010–2024	SDC Platinum
Pure primary offering	*Primary*	Indicator equal to one if all shares offered in the IPO are newly issued primary shares, and zero otherwise.	2010–2024	SDC Platinum
Positive earnings	*PositiveEPS*	Indicator equal to one if earnings per share over the 12 months preceding the IPO are positive, and zero otherwise.	2010–2024	SDC Platinum

Variable	Acronym	Definition	Time Period	Data Source
Secondary shares	*Second*	Number of secondary shares divided by the total number of shares offered in the IPO.	2010–2024	SDC Platinum
R&D before the IPO	*R&D (before IPO)*	Research and development expense for the fiscal year immediately preceding the IPO, scaled by total assets for the same fiscal year.	2010–2024	Compustat
R&D after the IPO	*R&D (after IPO)*	Research and development expense for the first fiscal year following the IPO, scaled by total assets for the same fiscal year.	2010–2024	Compustat
Pre-filing comment letters	*ComLetter*	Natural logarithm of one plus the number of SEC comment letters concerning IPO filings issued from twelve to six months before the firm's filing date.	2010–2025	Ideagen Audit Analytics
Price-to-sales ratio	*P/S*	Offer price divided by pre-IPO sales per share, constructed following Purnanandam and Swaminathan (2004).	2010–2024	SDC Platinum; CRSP; Compustat
AI patent stock	*AIStock*	Depreciated stock of AI patents accumulated over the ten years preceding each firm-year observation.		
AI valuation premium	*AIVal*	Coefficient on AIStock from firm-value regressions estimated within each Fama–French 12 industry using firm-year observations over the three years preceding the IPO.		
Business–AI distance	*AIDistance*	One minus the cosine similarity between the embedding of the firm's Business section in its S-1 registration statement and the centroid embedding of the AI patent corpus, scaled by the sample mean.	2010–2024	SEC EDGAR; USPTO AIPD
AI job postings	*AIJobPost*	Number of AI-related job postings associated with the firm in an event quarter around the expiration of the IPO lockup period.	2010–2024	LightCast (Burning Glass)
Litigation risk (KS)	*LitRiskKS*	Ex ante probability of securities litigation constructed following Kim and Skinner (2012).	2010–2024	CRSP–Compustat
Litigation risk (DFY)	*LitRiskDFY*	Ex ante probability of securities litigation constructed following Donelson, Flam, and Yust (2022).	2010–2024	Ideagen Audit Analytics
10-K AI disclosure gap	*AIGapTenK*	Total AI-keyword occurrences in the Business, Risk Factors, and MD&A sections of the first Form 10-K filed at least one year after the IPO minus AI-keyword occurrences in the firm's S-1 registration statement.	2010–2025	SEC EDGAR (via WRDS)
Conference-call AI disclosure gap	*AIGapCall*	AI-word intensity in earnings-call transcripts during post-IPO quarters 5 through 8 minus AI-word intensity in the firm's S-1 registration statement.	2010–2025	CapitalIQ earnings-call transcripts (via WRDS)

Variable	Acronym	Definition	Time Period	Data Source
Venture capitalist AI skills	*VCAISkill*	Indicator equal to one if the lead venture capitalist has AI-related skills identified from employee profiles, and zero otherwise.	2010–2024	Revelio Labs
Underwriter AI experience	*UWAIExp*	Indicator equal to one if the lead underwriter completed at least one AI-investment IPO before the focal IPO, and zero otherwise.	2010–2024	
Auditor AI experience	*AuditorAIExp*	Indicator equal to one if the auditor audited at least one AI-investment IPO before the focal IPO, and zero otherwise.	2010–2024	

Appendix B. AI Keyword List

The appendix presents the final AI keyword dictionary used in the textual analysis. The initial dictionary consists of 97 AI-related terms compiled from prior literature and industry sources. Of these, 57 keywords appear in the IPO prospectus corpus and are used throughout the analysis to identify AI-related disclosures and construct the AI-washing measure.

Keywords	Keywords	Keywords	Keywords (not found)	Keywords (not found)
3D Imaging	Feature Selection	Pattern Recognition	AI ChatBot	Matrix Factorization
Acoustic Model	Gen AI	Predictive Analytics	Apache UIMA	Maximum Entropy Classifier
Active Learning	Generative Artificial Intelligence	Random Forest	Autoencoders	Microsoft Cognitive Toolkit
AI Strategy	Gesture Recognition	Recommender System	Bayesian Networks	MLPACK
AI Training	H2O.ai	Regression Model	Caffe Deep Learning Framework	MXNet
AIOps	Hierarchical Clustering	Search Algorithm	Computational Linguistics	Naive Bayes
Artificial Intelligence	Image Matching	Semantics	DBSCAN	Natural Language Toolkit
Augmented Reality	Image Recognition	Semi-supervised Learning	Deeplearning4j	OpenCV
Automatic Speech Recognition	Intelligent Agent	Sensor Network	Expectation-Maximization Algorithm	OpenNLP
Autonomous System	Language Model	Sentiment Analysis	Hidden Markov Model	Pybrain
Azure ML	Machine Learning	Signal Processing	Information Extraction	Scikit-learn
Bayesian Optimization	Maximum Likelihood	Supervised Learning	Jung Framework	Sentiment Classification
Classification Algorithm	Model Selection	Support Vector Machine	Keras	Stochastic Gradient Descent
Cluster Analysis	Motion Capture	Tensor Flow	Kernel Methods	Theano
Collaborative System	Natural Language Processing	Tokenization	K-Means	Unstructured Information Management Architecture
Computer Vision	Network Analysis	Torch	Latent Dirichlet Allocation	Unsupervised Learning
Data Fusion	Neural Network	Virtual Agent	Latent Semantic Analysis	Vowpal
Deep Learning	Object Oriented		Libsvm	Wabbit
Dimensionality Reduction	Object Recognition		Long Short-Term Memory	Word2Vec
Facial Recognition	Optimal Search		Machine Translation	Xgboost

Figure 1. 2×2 Classification Matrix

This figure classifies the 1,671 IPO firms in our 2010–2024 sample along two dimensions: AI disclosure ("AI talk") and observable AI inputs ("AI walk"). AI-aligned IPOs have both AI talk and observable AI walk. AI-quiet IPOs have observable AI walk but without AI disclosure. AI-washing IPOs have high AI disclosure but have neither observable AI employees nor AI patent filings. Non-AI IPOs have no observable AI inputs and do not meet the high AI disclosure threshold. AI-investment IPOs comprise the AI-aligned and AI-quiet groups. Reported numbers are firm counts, with percentages in parentheses denoting shares of the full classification sample.

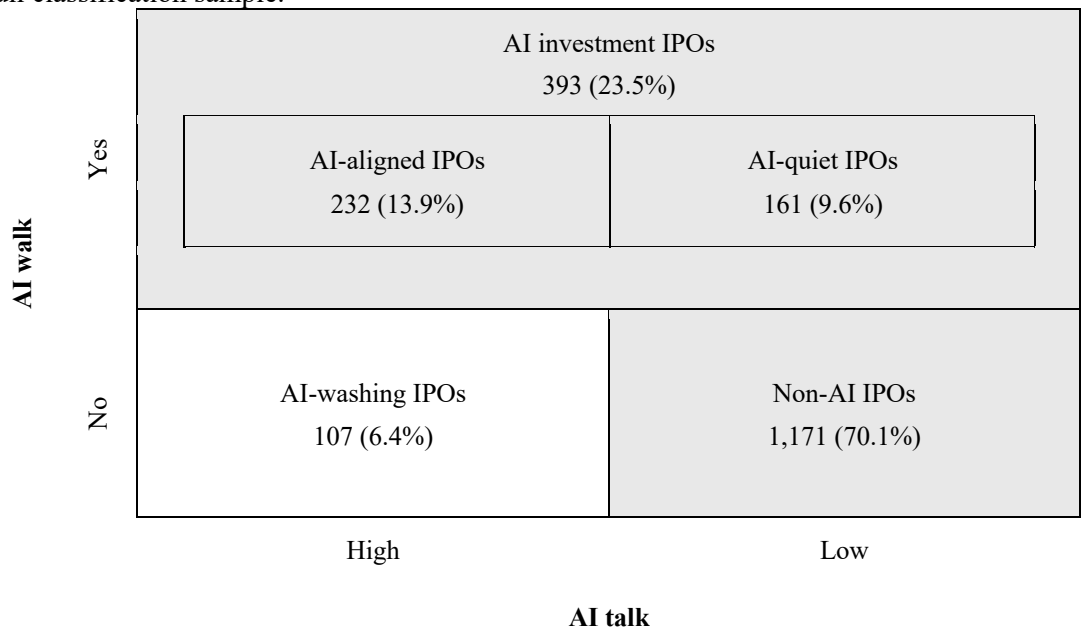
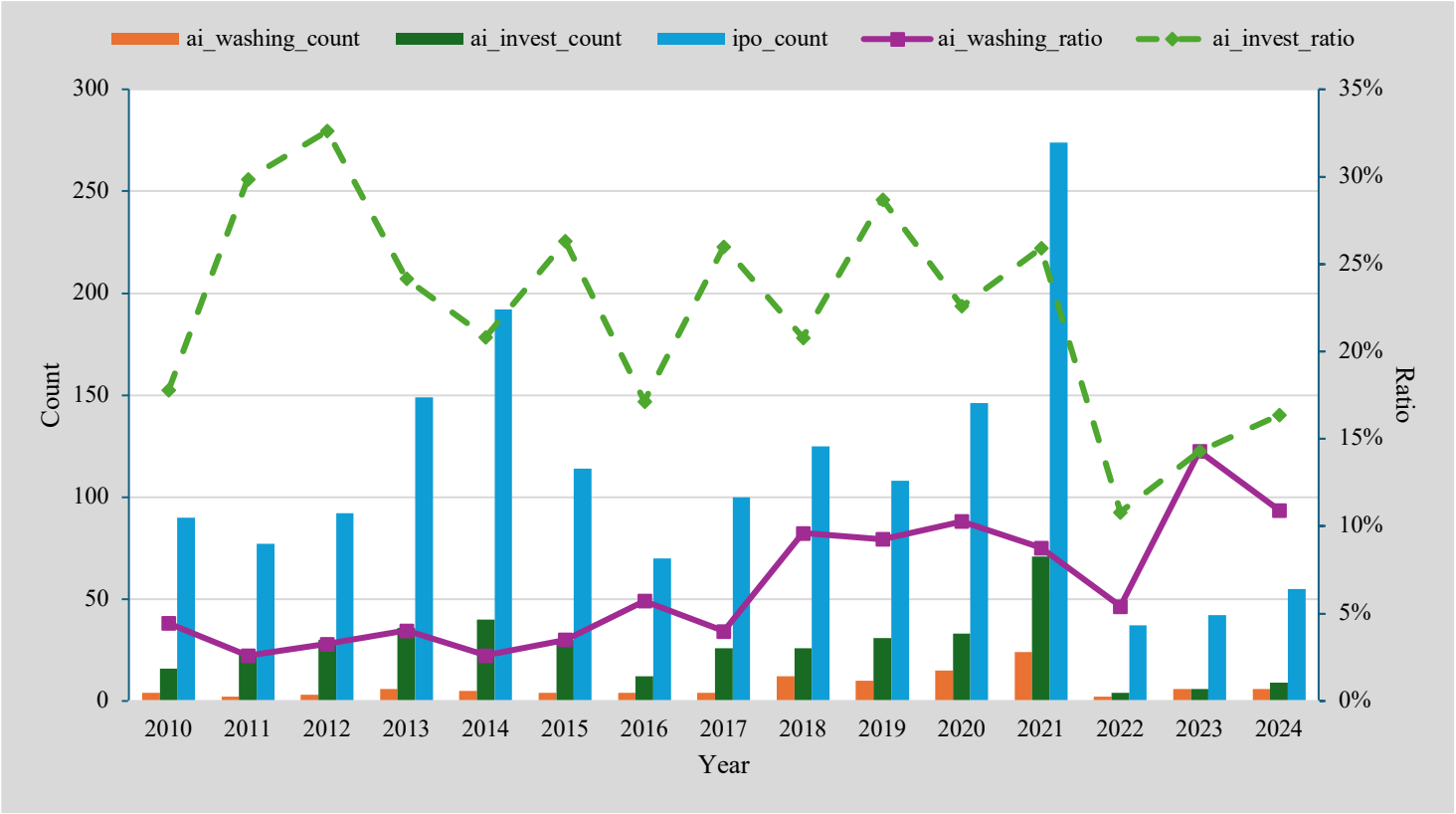


Figure 2. The Distributions of AI-washing IPOs

This figure shows the distributions of AI-related IPOs. Panel A plots the annual number of AI-washing and AI investment IPOs. Panel B reports their distributions across the Fama-French 12 industries. The sample period is from January 1, 2010, to December 31, 2024.

Panel A. Time Series Distribution



Panel B. Industry Distribution

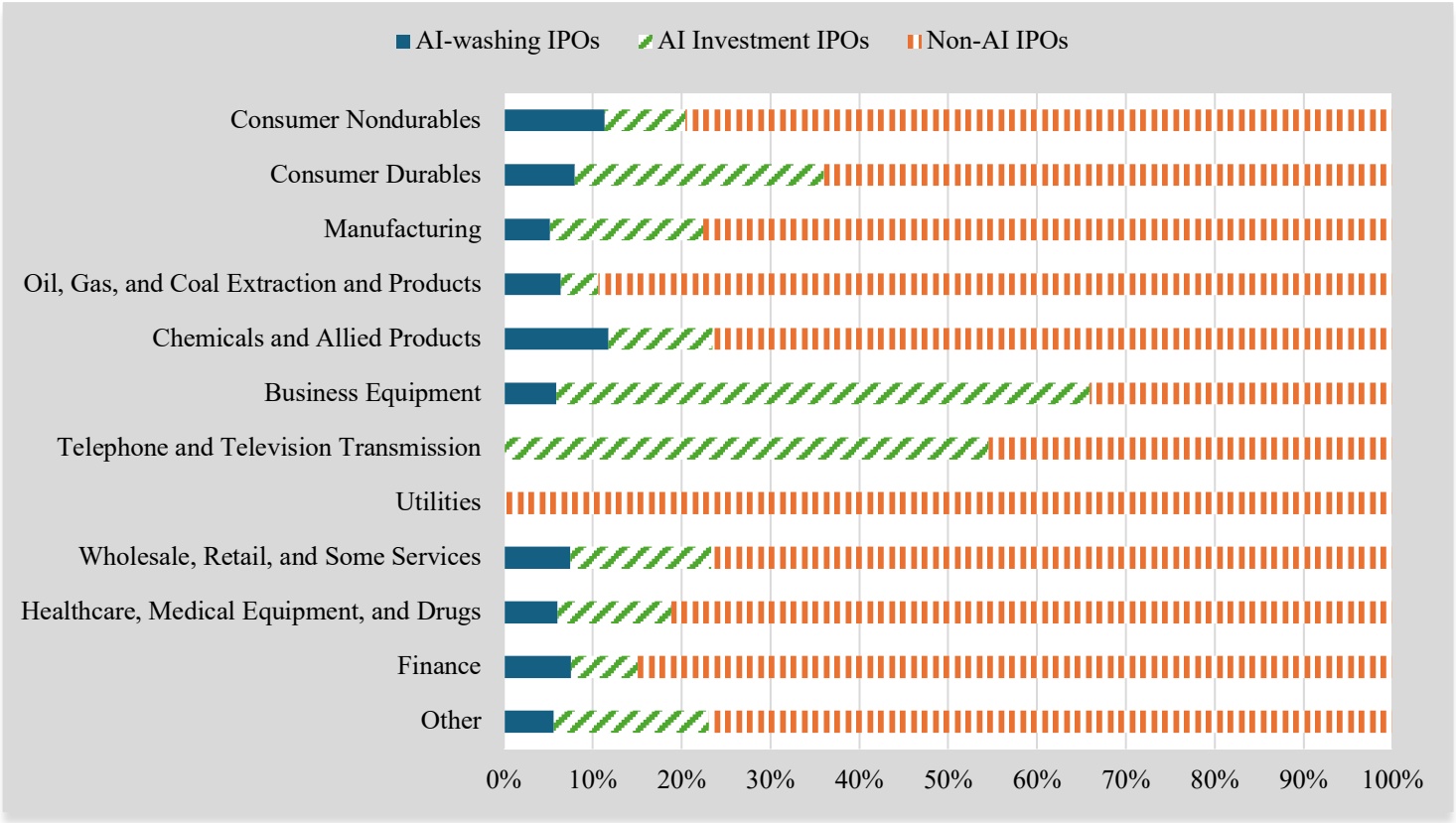
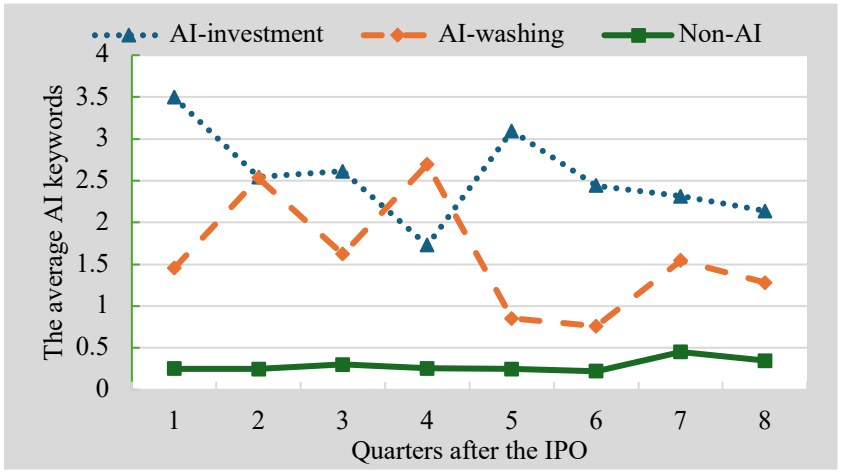


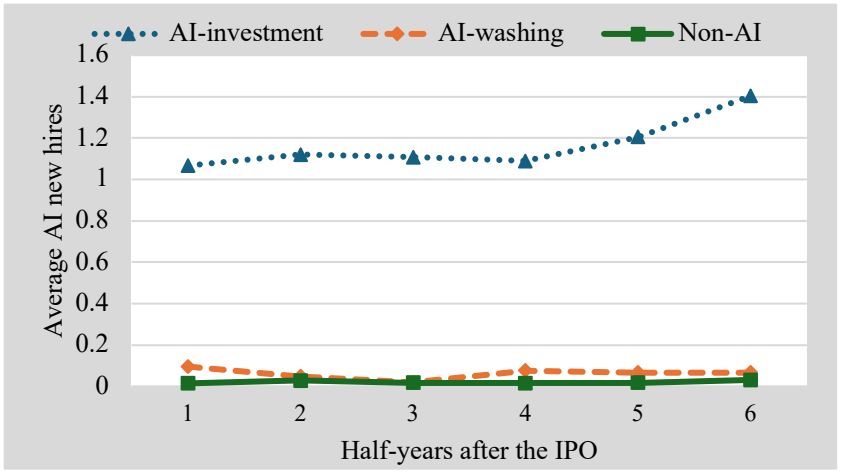
Figure 3. Average AI Talking, Hiring and Patent Filings After IPOs

This figure compares post-IPO AI-related activities across three groups of firms: AI-investment firms, AI-washing firms, and non-AI firms. Panel A reports the average number of AI-related keyword mentions in earnings calls from the first to the eighth quarter after the IPO. Panel B reports the average number of AI-related new hires from the first to the sixth half-year after the IPO. Panel C reports the average number of AI-related patent filings from the first to the sixth half-year after the IPO. The sample period is from January 1, 2010, to December 31, 2024.

Panel A. Post-IPO AI Keywords in Earnings Calls



Panel B. Post-IPO AI-Related New Hires



Panel C. Post-IPO AI Patent Filings

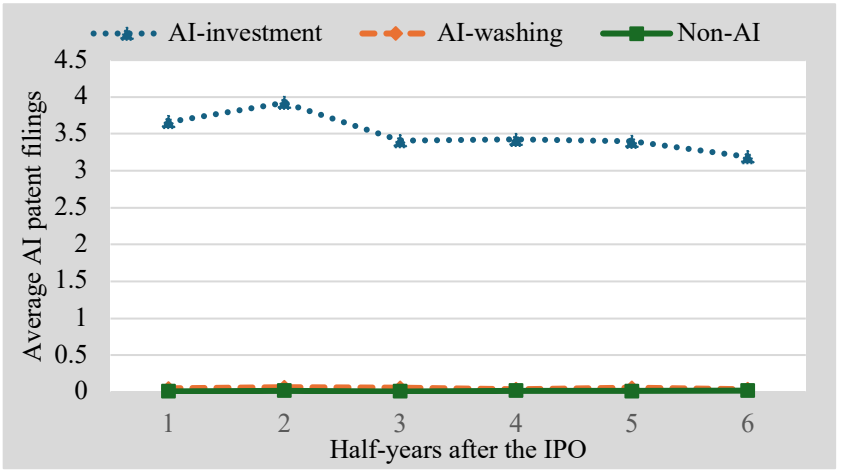
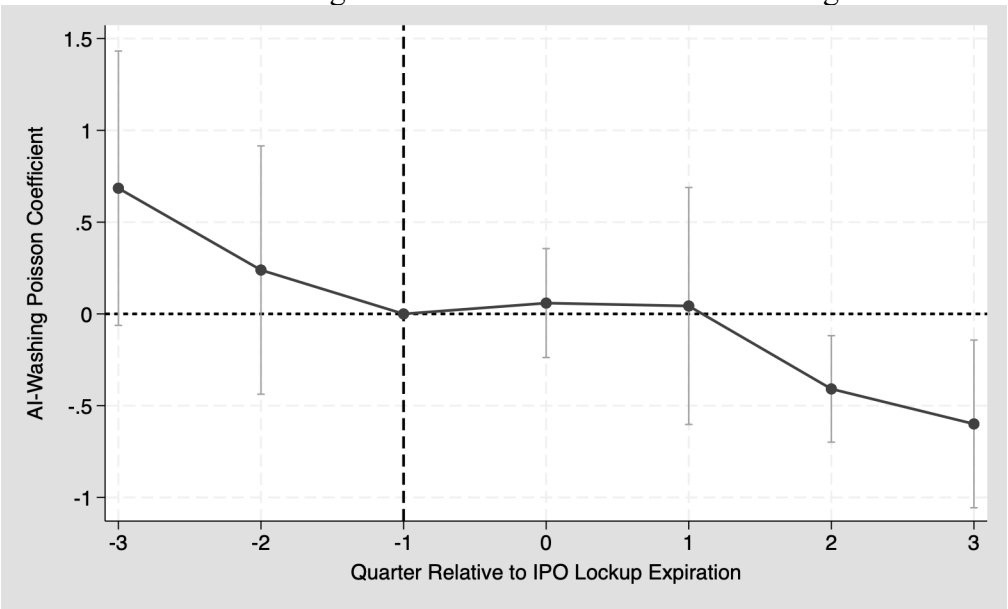


Figure 4. AI Job Postings around the Expiration Day

This figure presents event-study estimates from Poisson regressions around IPO lockup expiration. Event time is measured in quarters relative to the lockup expiration quarter, with quarter −1 omitted as the reference period. The event window spans quarters −3 to +3. Panel A compares AI-washing firms with AI-investment firms. Panels B1 and B2 compare AI-washing and AI-investment firms, respectively, with non-AI firms. Panel C plots the event-time coefficients for AI-washing and AI-investment firms, with each series normalized to its own quarter −1. The figure reports 95% confidence intervals. All regressions include two-digit industry × year-quarter fixed effects, relative-quarter fixed effects, and the full set of control variables. Standard errors are heteroskedasticity-robust and clustered at the industry level. The sample period is from January 1, 2010, to December 31, 2024.

Panel A. Marginal Effects on AI-Related Job Postings



Panel B. AI-Investment and AI-Washing Firms Relative to Non-AI Firms

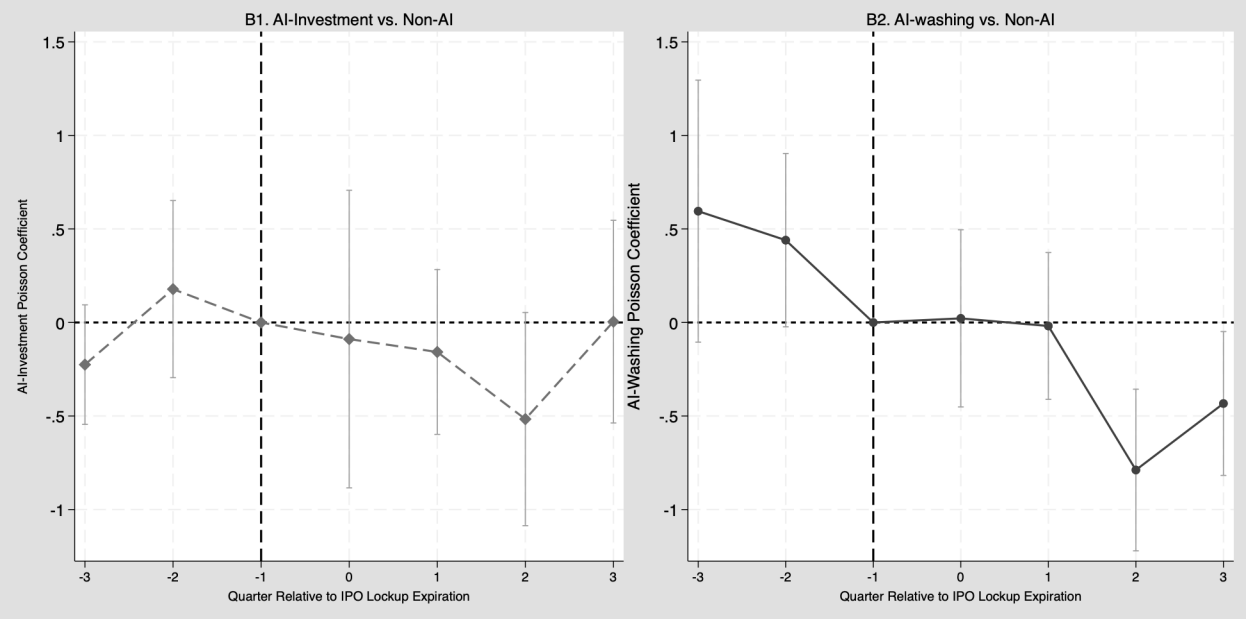


Figure 5. Post-IPO Buy-and-hold Abnormal Returns

This figure plots cumulative buy-and-hold returns for different IPO portfolios. Returns are equal-weighted across firms and adjusted by the Fama-French 5×5 portfolio returns. The solid line with squares represents all IPOs. The dotted line represents IPOs excluding AI-related IPOs. The solid line with triangles represents AI investment IPOs. The dashed line represents AI-washing IPOs. The sample period is from January 1, 2010, to December 31, 2024.

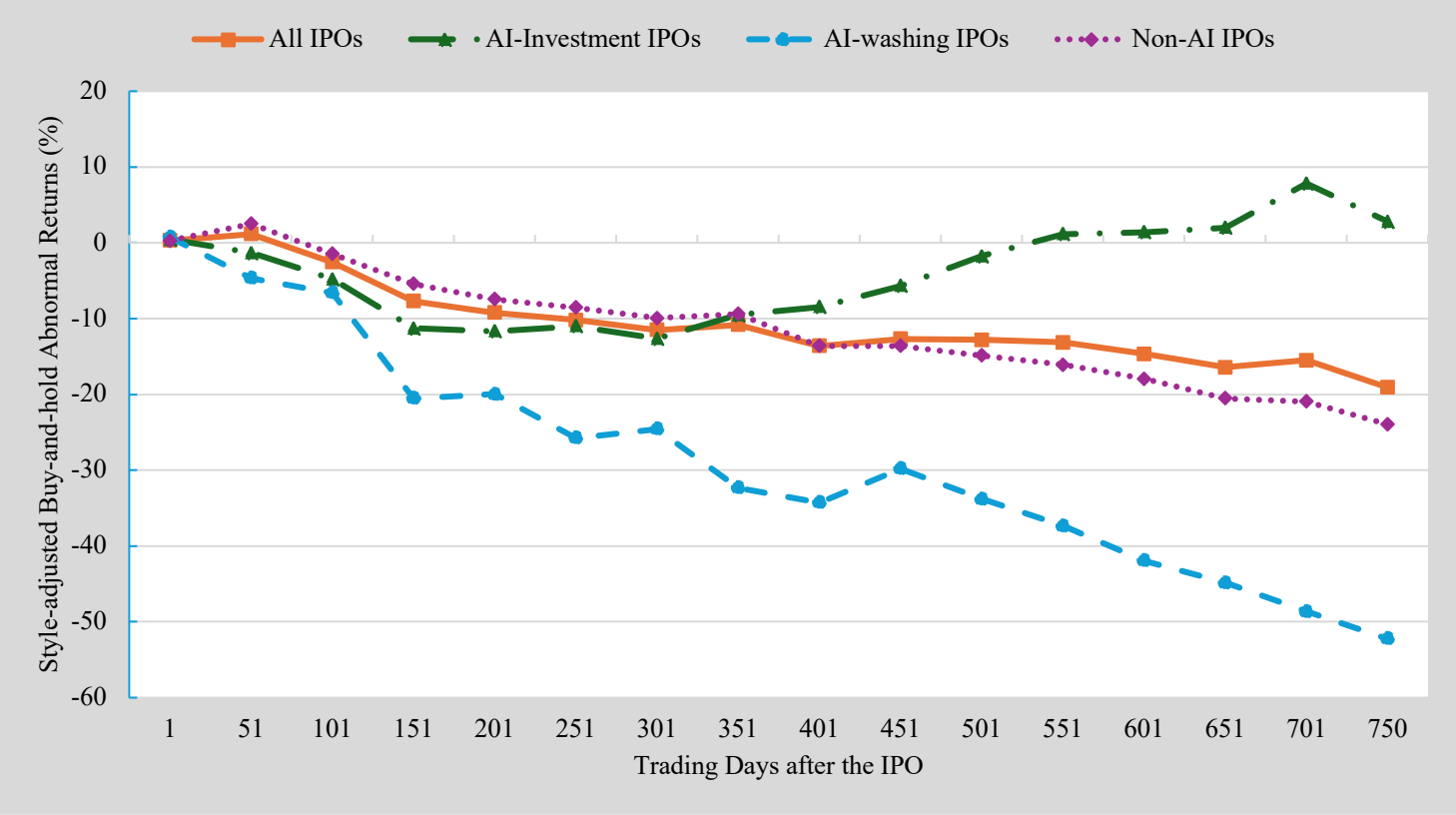


Table 1. Summary Statistics

This table presents descriptive statistics, correlations, and univariate analyses for 1,671 U.S. IPOs from 2010 to 2024. Panel A1 reports summary statistics for initial returns and buy-and-hold returns adjusted by the corresponding Fama–French 5×5 portfolio returns. Panel A2 reports summary statistics for AI-related characteristics. Panel A3 reports summary statistics for other IPO characteristics, presenting the variables in their raw (pre-log-transformation) values. Panel B1 reports correlation coefficients for main variables, including dependent and independent variables. Panel B2 reports correlation coefficients for other control variables. Panel C1 compares AI-washing IPOs with non-AI IPOs, while Panel C2 compares AI-washing IPOs with AI-investment IPOs. All variables are defined in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Panel A1. Summary Statistics for First Day and Buy-and-hold Returns

	N	Mean	STD	10%	25%	Median	75%	90%
BHAR[0,1yr]	1,600	-0.110	0.658	-0.803	-0.545	-0.196	0.194	0.677
BHAR[1,2yr]	1,491	-0.094	0.626	-0.768	-0.489	-0.179	0.170	0.562
BHAR[0,2yr]	1,556	-0.137	1.005	-1.058	-0.754	-0.351	0.225	0.948
BHAR[0,3yr]	1,515	-0.172	1.297	-1.278	-0.946	-0.490	0.235	1.255
InitialReturn	1,671	0.197	0.327	-0.083	0.000	0.109	0.308	0.603

Panel A2. Summary Statistics for AI-related Characteristics

	N	Mean	STD	10%	25%	Median	75%	90%
AIWashing	1,671	0.064	0.245	0.000	0.000	0.000	0.000	0.000
AIInvestment	1,671	0.235	0.424	0.000	0.000	0.000	0.000	1.000
AIWord	1,671	2.841	8.603	0.000	0.000	0.000	1.000	8.000
AIPatent	1,671	1.530	6.050	0.000	0.000	0.000	0.000	3.000
AIEmployee	1,671	0.500	2.280	0.000	0.000	0.000	0.000	1.000

Table 1—*continued*

Panel A3. Summary Statistics for Other Firm-level Characteristics

	N	Mean	STD	10%	25%	Median	75%	90%
HighUW	1,670	0.513	0.500	0.000	0.000	1.000	1.000	1.000
VC	1,671	0.591	0.492	0.000	0.000	1.000	1.000	1.000
Big4	1,671	0.393	0.489	0.000	0.000	0.000	1.000	1.000
Proceeds (in millions)	1,671	248.756	383.826	30.000	67.200	120.000	250.000	576.000
Age	1,665	17.920	22.697	3.000	6.000	10.000	19.000	40.000
Assets (in millions)	1,669	1414.601	4109.190	16.701	63.757	173.568	876.255	3414.723
Sales (in millions)	1,668	652.544	1764.007	0.000	0.000	85.929	395.431	1596.681
Carve	1,671	0.178	0.382	0.000	0.000	0.000	0.000	1.000
HighTech	1,671	0.550	0.498	0.000	0.000	1.000	1.000	1.000
Biotech	1,671	0.323	0.468	0.000	0.000	0.000	1.000	1.000
PrcRev	1,671	-0.013	0.139	-0.214	-0.077	0.000	0.071	0.143
Internet	1,671	0.095	0.293	0.000	0.000	0.000	0.000	0.000
Overhang	1,671	4.615	4.688	1.059	2.175	3.420	5.590	8.560
Primary	1,671	0.328	0.470	0.000	0.000	0.000	1.000	1.000
PositiveEPS	1,671	0.271	0.445	0.000	0.000	0.000	1.000	1.000
Second	1,671	0.097	0.231	0.000	0.000	0.000	0.000	0.368

Table 1—*continued*

Panel B1. Correlation Coefficients for Main Variables (Spearman for the upper-right part, highlighted; Pearson for the bottom-left part)

	BHAR[0,1yr]	*BHAR[1,2yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*InitialReturn*	*AIWashing*	*AIInvestment*	*AIWord*	*AIPatent*	*AIEmployee*
BHAR[0,1yr]	-	0.205	0.757	0.632	0.009	-0.075	-0.014	-0.127	-0.022	-0.034
BHAR[1,2yr]	0.205	-	0.716	0.599	0.029	-0.027	0.076	0.008	0.084	0.046
BHAR[0,2yr]	0.757	0.716	-	0.829	0.008	-0.062	0.023	-0.067	0.012	-0.001
BHAR[0,3yr]	0.632	0.599	0.829	-	0.007	-0.056	0.068	-0.050	0.044	0.025
InitialReturn	0.009	0.029	0.008	0.007	-	0.040	0.195	0.181	0.169	0.162
AIWashing	-0.075	-0.027	-0.062	-0.056	0.040	-	-0.145	0.438	-0.116	-0.100
AIInvestment	-0.014	0.076	0.023	0.068	0.195	-0.145	-	0.370	0.800	0.689
AIWord	-0.057	-0.033	-0.064	-0.068	0.138	0.335	0.264	-	0.327	0.361
AIPatent	-0.022	0.084	0.012	0.044	0.169	-0.116	0.800	0.327	-	0.326
AIEmployee	-0.034	0.046	-0.001	0.025	0.162	-0.100	0.689	0.361	0.326	-

Table 1—*continued*

Panel B2. Correlation Coefficients for Control Variables (Spearman for the upper-right part, highlighted; Pearson for the bottom-left part)

	HighUW	*VC*	*Big4*	*Proceeds*	*Age*	*Assets*	*Sales*	*Carve*	*HighTech*	*Biotech*	*PrcRev*	*Internet*	*Overhang*	*Primary*	*PositiveEPS*	*Second*
HighUW	-	0.023	0.256	0.481	0.150	0.404	0.410	0.136	-0.086	-0.227	0.138	0.192	0.193	0.000	0.121	0.210
VC	0.029	-	0.031	-0.035	-0.344	-0.265	-0.262	-0.289	0.379	0.293	0.156	0.135	0.051	-0.009	-0.329	-0.110
Big4	0.256	0.029	-	0.026	0.087	0.076	0.138	0.158	0.021	-0.087	-0.070	0.054	-0.045	0.492	0.066	0.152
Proceeds	0.487	0.016	0.071	-	0.250	0.712	0.536	0.188	-0.182	-0.268	0.343	0.126	0.198	-0.224	0.156	0.261
Age	0.171	-0.332	0.103	0.289	-	0.416	0.544	0.230	-0.330	-0.425	-0.045	-0.006	0.080	-0.024	0.327	0.257
Assets	0.409	-0.188	0.114	0.739	0.463	-	0.701	0.277	-0.447	-0.472	0.074	0.091	0.222	-0.110	0.416	0.338
Sales	0.397	-0.209	0.149	0.519	0.544	0.683	-	0.286	-0.434	-0.643	0.063	0.201	0.263	-0.056	0.49	0.374
Carve	0.136	-0.259	0.158	0.186	0.250	0.274	0.273	-	-0.237	-0.184	-0.081	-0.086	-0.013	0.075	0.206	0.146
HighTech	-0.086	0.317	0.021	-0.152	-0.353	-0.422	-0.429	-0.237	-	0.517	0.096	0.091	0.017	0.043	-0.409	-0.234
Biotech	-0.227	0.244	-0.087	-0.238	-0.404	-0.444	-0.656	-0.184	0.517	-	-0.074	-0.223	-0.187	0.038	-0.367	-0.357
PrcRev	0.116	0.115	-0.085	0.273	-0.053	0.072	0.056	-0.088	0.086	-0.078	-	0.137	0.206	-0.145	-0.059	0.051
Internet	0.192	0.125	0.054	0.133	-0.028	0.079	0.199	-0.086	0.091	-0.223	0.130	-	0.186	-0.034	0.001	0.092
Overhang	0.077	0.041	-0.125	0.030	0.017	0.086	0.108	-0.027	0.034	-0.076	0.147	0.099	-	-0.171	0.042	0.097
Primary	0.000	-0.014	0.492	-0.182	-0.028	-0.089	-0.047	0.075	0.043	0.038	-0.148	-0.034	-0.156	-	-0.030	-0.385
PositiveEPS	0.121	-0.291	0.066	0.154	0.332	0.402	0.488	0.206	-0.409	-0.367	-0.047	0.001	-0.015	-0.030	-	0.336
Second	0.198	-0.127	0.135	0.264	0.271	0.328	0.328	0.208	-0.239	-0.273	0.003	0.017	0.029	-0.295	0.339	-

Table 1—*continued*

Panel C1. Univariate Analyses between AI-washing and non-AI IPOs (N = 1,278)						
	Non-AI IPOs		AI-washing IPOs		Differences	
Variable	N	Mean	N	Mean	Diff.	t-Stat
BHAR[0,1yr]	1,118	-0.098	98	-0.264	0.166	2.410**
BHAR[0,2yr]	1,086	-0.161	92	-0.350	0.189	1.834*
BHAR[0,3yr]	1,051	-0.216	90	-0.500	0.284	2.133**
InitialReturn	1,171	0.160	107	0.278	-0.118	-3.623***
HighUW	1,171	0.445	106	0.575	-0.131	-2.589***
VC	1,171	0.599	107	0.701	-0.102	-1.619
Age	1,165	2.488	107	2.513	-0.026	-0.270
Proceeds	1,171	4.672	107	4.847	-0.175	-1.515
Sales	1,168	3.330	107	4.182	-0.852	-2.907***
Assets	1,169	5.226	107	5.398	-0.172	-0.794
HighTech	1,171	0.495	107	0.495	0.000	0.000
Biotech	1,171	0.398	107	0.271	0.127	2.585***
Internet	1,171	0.035	107	0.150	-0.115	-5.555***
Second	1,171	0.091	107	0.086	0.004	0.195
Big4	1,171	0.386	107	0.234	0.152	3.128***
PositiveEPS	1,171	0.289	107	0.299	-0.010	-0.227
PrcRev	1,171	-0.029	107	-0.011	-0.019	-1.332
Carve	1,171	0.204	107	0.121	0.083	2.058**
Overhang	1,171	4.014	107	6.841	-2.827	-6.058***
Primary	1,171	0.353	107	0.187	0.166	3.482***

Table 1—*continued*

Panel C2. Univariate Analyses between AI-washing and AI investment IPOs (N = 500)						
	AI Investment IPOs		AI-washing IPOs		Differences	
Variable	N	Mean	N	Mean	Diff.	t-Stat
BHAR[0,1yr]	384	-0.106	98	-0.264	0.158	2.091**
BHAR[0,2yr]	378	-0.019	92	-0.350	0.331	2.565**
BHAR[0,3yr]	374	0.031	90	-0.500	0.532	3.218***
InitialReturn	393	0.286	107	0.278	0.008	0.205
HighUW	393	0.697	106	0.575	0.122	2.376**
VC	393	0.814	107	0.701	0.113	2.067**
Age	393	2.553	107	2.513	0.04	0.525
Proceeds	393	5.34	107	4.847	0.493	3.975***
Sales	393	4.975	107	4.182	0.793	3.002***
Assets	393	5.719	107	5.398	0.321	1.611
HighTech	393	0.728	107	0.495	0.232	4.650***
Biotech	393	0.115	107	0.271	-0.157	-4.102***
Internet	393	0.257	107	0.15	0.107	2.336**
Second	393	0.119	107	0.086	0.033	1.273
Big4	393	0.455	107	0.234	0.222	4.204***
PositiveEPS	393	0.211	107	0.299	-0.088	-1.918*
PrcRev	393	0.037	107	-0.011	0.048	3.141***
Carve	393	0.115	107	0.121	-0.007	-0.2
Overhang	393	5.799	107	6.841	-1.042	-1.893*
Primary	393	0.293	107	0.187	0.106	2.190**

Table 2. AI Investment and AI Disclosures in IPO Prospectuses

This table presents Poisson regression results for AI disclosure (*AIWord*) on pre-IPO AI employees (*AIEmployee*), and AI patent filings (*AIPatent*). *AIWord* is the number of AI keywords in IPO S-1 prospectuses. *AIEmployee* is the number of AI-related employees prior to the IPO listing date. *AIPatent* is the number of AI-related patent filings from three years prior to the IPO to the listing date. The construction of these variables is detailed in Appendix A. We control for fixed effects at the two-digit industry and year levels. t-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and clustering by industry. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dep. Variable=	*AIWord*					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
AIEmployee	**0.093*** **	**0.046*** **			**0.077*** **	**0.042*** **
	(6.24)	**(5.20)**			**(7.98)**	**(6.34)**
AIPatent			**0.031*** **	**0.012*** **	**0.012*** **	**0.005**
			(4.21)	**(2.94)**	**(3.64)**	**(1.49)**
Fixed Effects	No	Ind & Year	No	Ind & Year	No	Ind & Year
Observations	1,671	1,660	1,671	1,660	1,671	1,660
Pseudo R^2	8.1%	32.1%	4.6%	31.0%	8.6%	32.2%

Table 3. Regulatory Scrutiny and Propensity to AI-washing

This table reports the Logit regression of AI-washing IPOs (*AIWashing*) and AI investment IPOs (*AIInvestment*) on the pre-filing comment letter (*ComLetter*). The pre-filing comment letter (*ComLetter*) is the logarithm of one plus the number of comment letters on IPOs one year before IPOs to half a year before IPOs. Columns with controls include the full set of control variables defined in Appendix A. We control for fixed effects at the two-digit industry and year levels. *t-statistics* shown in parentheses are based on standard errors adjusted for heteroskedasticity and clustering by industry. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dep. Variable=	*AIWashing*		*AIInvestment*	
Variable	(1)	(2)	(3)	(4)
ComLetter	**-0.264*** **	**-0.249** **	**0.082**	**-0.044**
	(-3.57)	**(-2.44)**	**(0.84)**	**(-0.58)**
Controls	No	Yes	No	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,445	1,436	1,524	1,515
Pseudo R^2	8.5%	11.4%	23.5%	33.9%

Table 4. Baseline Test: AI-washing and IPO Performances

This table presents regression results on the relation between AI-washing and IPO performance. Panel A reports results for one-year buy-and-hold abnormal returns (*BHAR[0,1yr]*), two-year buy-and-hold abnormal returns (*BHAR[0,2yr]*), three-year buy-and-hold abnormal returns (*BHAR[0,3yr]*), and initial returns (*InitialReturn*). Buy-and-hold abnormal returns are adjusted by the corresponding Fama–French 5×5 portfolio returns. *AIWashing* is an indicator equal to one if a firm's AI-related disclosure exceeds the 95th percentile of AI-related disclosure among IPO firms in the same Fama–French 12 industry, where the percentile cutoff is calculated using an expanding window from the beginning of the sample through the year preceding the firm's IPO, and both the number of AI employees and AI patent filings are zero. *AIInvestment* is an indicator equal to one if either the number of AI employees or AI patent filings is greater than zero. Panel B reports results based on one-to-five nearest-neighbor propensity-score matching with a caliper of 0.05 and entropy balancing. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Baseline Regression Results

	Long-term returns			Short-term returns
Dep. Variable=	*BHAR[0,1yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*InitialReturn*
Variable	(1)	(2)	(3)	(4)
AIWashing	**-0.077**	**-0.088****	**-0.139*****	**0.072****
	(-1.09)	**(-2.35)**	**(-2.73)**	**(2.04)**
AIInvestment	**-0.022**	**0.102****	**0.220*****	0.014
	(-0.74)	**(2.31)**	**(4.85)**	(0.59)
HighUW	-0.042	-0.031	0.048	-0.010
	(-0.81)	(-0.42)	(0.45)	(-1.08)
VC	0.061	0.037	0.083	0.028**
	(1.47)	(0.68)	(1.34)	(2.13)
Big4	-0.010	-0.075	-0.088	0.008
	(-0.12)	(-1.12)	(-0.98)	(0.70)
Proceeds	0.034	0.027	-0.021	0.035
	(1.10)	(0.43)	(-0.29)	(1.55)
Age	0.026	0.052	0.039	-0.013
	(0.70)	(1.49)	(0.71)	(-1.23)
Assets	0.013	0.007	0.006	-0.025***
	(0.61)	(0.27)	(0.20)	(-5.11)
Sales	0.015	0.025*	0.040*	0.000
	(1.67)	(1.71)	(1.93)	(0.09)
Carve	0.047	0.054	0.080	0.034**
	(0.83)	(0.93)	(1.40)	(2.54)

Table 4—*continued*

HighTech	0.004	-0.040	0.036	0.011
	(0.13)	(-0.71)	(0.31)	(0.56)
Biotech	0.275**	0.493***	0.444***	-0.002
	(2.62)	(3.74)	(2.82)	(-0.06)
Internet	-0.060	-0.009	-0.124	0.052***
	(-1.10)	(-0.14)	(-1.43)	(3.06)
Overhang	-0.003	0.001	-0.002	0.002
	(-0.68)	(0.21)	(-0.43)	(0.97)
Primary	0.029	0.083	-0.023	-0.004
	(0.56)	(1.20)	(-0.25)	(-0.26)
PositiveEPS	0.115***	0.104	0.128*	0.001
	(3.08)	(1.54)	(1.75)	(0.04)
Second	-0.083	0.108	0.011	0.007
	(-1.21)	(0.95)	(0.08)	(0.21)
InitialReturn	-0.065	-0.049	-0.015	
	(-0.91)	(-0.80)	(-0.14)	
PrcRev				0.806***
				(16.67)
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,581	1,538	1,498	1,652
Adjusted R^2	14.5%	10.0%	7.6%	22.7%
AIWashing - AIInvestment	**-0.055**	**-0.190**	**-0.360**	**0.06**
Test of difference [p-value]	**[0.43]**	**[0.00]**	**[0.00]**	**[0.08]**

Panel B. Propensity Score Matching and Entropy Balancing Results

	Propensity-scores matching		Entropy balancing	
Dep. Variable=	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*
Variable	(1)	(2)	(3)	(4)
AIWashing	**-0.161****	**-0.188***	**-0.128****	**-0.196*****
	(-2.63)	**(-1.99)**	**(-2.09)**	**(-3.34)**
AIInvestment	**0.110**	**0.310****	**0.150**	**0.227*****
	(1.52)	**(2.46)**	**(0.59)**	**(2.91)**
Controls	YES	YES	YES	YES
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	471	453	1,508	1,468
Adjusted R^2	14.8%	11.3%	18.7%	16.8%

Table 5. Calendar-time Factor Regressions for the Portfolio Excess Returns

These tables report calendar-time factor regressions for portfolios formed from AI-washing IPOs. In each calendar month, portfolios are formed based on the indicated post-IPO holding windows, and monthly portfolio excess returns are computed. Panel A reports factor regressions for AI-washing IPO portfolios. Panel B reports long–short portfolio regressions that are long AI-washing IPOs and short AI-investment IPOs. Panel C reports long–short portfolio regressions that are long AI-washing IPOs and short non-AI IPOs, where non-AI IPOs include both AI-investment IPOs and non-AI IPOs. CAPM, FF3, and FF5 denote regressions based on the market factor, the Fama–French three-factor model, and the Fama–French five-factor model, respectively. The dependent variables and factor returns are measured in percentage points per month. The intercept, α, represents the monthly abnormal return. Newey–West t-statistics based on three lags are reported in parentheses. The sample period is January 2013 to December 2025 for the 24-month holding window and January 2012 to December 2025 for the 36-month holding window. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. AI-washing portfolios

	24-month holding period			36-month holding period		
	CAPM	FF3	FF5	CAPM	FF3	FF5
	(1)	(2)	(3)	(4)	(5)	(6)
α	**-1.946*** **	**-1.553** **	**-1.491** **	**-1.694*** **	**-1.255*** **	**-1.128*** **
	(-2.92)	**(-2.55)**	**(-2.58)**	**(-3.46)**	**(-2.72)**	**(-2.65)**
MKT	1.495***	1.264***	1.232***	1.605***	1.357***	1.311***
	(10.85)	(9.14)	(8.74)	(13.61)	(12.39)	(12.72)
SMB		1.144***	0.852***		1.223***	0.921***
		(5.12)	(3.23)		(6.00)	(4.18)
HML		-0.805***	-0.346		-0.764***	-0.339*
		(-4.97)	(-1.51)		(-5.80)	(-1.80)
RMW			-0.571*			-0.633**
			(-1.76)			(-2.48)
CMA			-0.835***			-0.763***
			(-2.70)			(-2.85)
Observations	154	154	154	167	167	167
Adjusted R^2	40.6%	52.4%	55.7%	50.4%	62.4%	65.7%

Panel B. Long-short portfolios: Buy AI-washing IPO stocks and sell AI-investment IPO stocks

	24-month holding period			36-month holding period		
	CAPM	FF3	FF5	CAPM	FF3	FF5
	(1)	(2)	(3)	(4)	(5)	(6)
α	**-1.646*** **	**-1.605*** **	**-1.606*** **	**-1.309*** **	**-1.259*** **	**-1.255*** **
	(-2.64)	**(-2.68)**	**(-2.66)**	**(-2.94)**	**(-2.82)**	**(-2.78)**
Observations	154	154	154	167	167	167
Adjusted R^2	1.8%	1.7%	1.3%	3.4%	3.4%	3.2%

Panel C. Long-short portfolios: Buy AI-washing IPO stocks and sell non-AI-washing IPO stocks

	24-month holding period			36-month holding period		
	CAPM	FF3	FF5	CAPM	FF3	FF5
	(1)	(2)	(3)	(4)	(5)	(6)
α	**-1.298** **	**-1.237** **	**-1.239** **	**-0.955** **	**-0.874** **	**-0.853** **
	(-2.28)	**(-2.21)**	**(-2.20)**	**(-2.37)**	**(-2.08)**	**(-2.02)**
Observations	154	154	154	167	167	167
Adjusted R^2	2.6%	1.9%	2.5%	6.1%	6.0%	6.9%

Table 6. Excluding AI-washing IPOs Associated with Third-party AI Services

This table reports regression results after excluding AI-washing IPOs associated with third-party AI activities. In Columns (1) and (2), AI-washing IPOs mentioning third-party outsourcing or AI tools/services in AI-related disclosures are excluded based on keyword searches, resulting in the exclusion of 33 firms. In Columns (3) and (4), AI suppliers identified using FactSet Supply Chain data are excluded, removing 11 AI-washing IPOs. In Columns (5) and (6), AI collaboration partners identified using RavenPack news data are excluded, removing 9 AI-washing IPOs. The dependent variables are two-year (*BHAR[0,2yr]*) and three-year (*BHAR[0,3yr]*) buy-and-hold abnormal returns, adjusted by the corresponding Fama–French 5×5 portfolio returns. *AIWashing* and *AIInvestment* are defined as in Table 4. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Exclude AI outsources and AI tools		Exclude AI suppliers		Exclude AI partnerships	
Dep. Variable=	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*
Variable	(1)	(2)	(3)	(4)	(5)	(6)
AIWashing	**-0.123***	**-0.203****	**-0.070***	**-0.141*****	**-0.096****	**-0.124****
	(-2.00)	**(-2.63)**	**(-1.80)**	**(-2.72)**	**(-2.17)**	**(-2.13)**
AIInvestment	**0.115****	**0.239*****	**0.101****	**0.222*****	**0.103****	**0.221*****
	(2.33)	**(4.88)**	**(2.26)**	**(4.78)**	**(2.43)**	**(4.91)**
Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,511	1,471	1,530	1,490	1,531	1,491
Adjusted R^2	10.0%	7.5%	10.0%	7.7%	10.0%	7.5%
AIWashing - AIInvestment	**-0.239**	**-0.442**	**-0.171**	**-0.363**	**-0.200**	**-0.346**
Test of difference [p-value]	**[0.02]**	**[0.00]**	**[0.01]**	**[0.00]**	**[0.00]**	**[0.00]**

Table 7. Valuation Premium Channel

This table reports tests of the AI valuation premium mechanism. Columns (1) and (2) use the price-to-sales ratio (P/S) as the dependent variable to examine whether AI-washing IPOs receive higher valuations at issuance, following Purnanandam and Swaminathan (2004). Columns (3) and (4) report regressions of two-year buy-and-hold abnormal returns (*BHAR[0,2yr]*), and Columns (5) and (6) report regressions of three-year buy-and-hold abnormal returns (*BHAR[0,3yr]*). Buy-and-hold abnormal returns are adjusted by the corresponding Fama–French 5×5 portfolio returns. *AIVal* measures the industry-level AI valuation premium and is estimated as the coefficient on AI patent stock (*AIStock*) from regressions of firm market value on *AIStock* using firm-year observations within each Fama–French 12 industry over the three years preceding the IPO. *AIStock* is constructed as the depreciated stock of AI patents accumulated over the previous ten years. *AIWashing × AIVal* denotes the interaction between AI-washing status and the industry-level AI valuation premium. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*P/S*		*BHAR[0,2yr]*		*BHAR[0,3yr]*	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
AIWashing×AIVal		**16.105*** **		**-6.053*** **		**-5.407***
		(2.87)		**(-3.16)**		**(-2.00)**
AIVal	**-0.921**	**-2.104**	**1.516**	**1.842**	**0.675**	**0.974**
	(-0.28)	**(-0.64)**	**(0.95)**	**(1.10)**	**(0.17)**	**(0.25)**
AIWashing	-0.356**	-0.465***	-0.086**	-0.043	-0.138**	-0.099
	(-2.55)	(-2.98)	(-2.34)	(-1.30)	(-2.59)	(-1.58)
AIInvestment	-0.403***	-0.412***	0.102**	0.106**	0.221***	0.224***
	(-3.62)	(-3.72)	(2.33)	(2.39)	(4.91)	(5.05)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	605	605	1,547	1,547	1,511	1,511
Adjusted R^2	24.0%	24.2%	10.0%	10.0%	7.4%	7.4%

Table 8. Verification Channel

This table reports tests of the AI verification difficulty mechanism. The dependent variables are two-year (*BHAR[0,2yr]*) and three-year (*BHAR[0,3yr]*) buy-and-hold abnormal returns, adjusted by the corresponding Fama–French 5×5 portfolio returns. *AIDistance* is measured as one minus the cosine similarity between the firm's Business section in its S-1 prospectus and an AI patent reference corpus, normalized by its sample mean. *AIWashing* × *AIDistance* denotes the interaction between *AI-washing* and *AIDistance*. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*BHAR[0,2yr]*		*BHAR[0,3yr]*	
Variable	(1)	(2)	(3)	(4)
AIWashing×AIDistance		**-1.313****		**-1.425***
		(-2.19)		**(-1.98)**
AIDistance	**-0.099**	**0.003**	**-0.310**	**-0.200**
	(-0.70)	**(0.03)**	**(-1.25)**	**(-0.82)**
AIWashing	-0.085**	-0.121**	-0.166***	-0.209***
	(-2.12)	(-2.37)	(-4.31)	(-4.69)
AIInvestment	0.062	0.072	0.156***	0.167***
	(1.12)	(1.38)	(3.23)	(3.44)
Controls	Yes	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,410	1,410	1,375	1,375
Adjusted R^2	8.7%	8.7%	6.2%	6.2%

Table 9. Litigation Risk Channel

This table reports tests of the litigation risk concern mechanism. The dependent variable is the second-year buy-and-hold abnormal return (*BHAR[1,2yr]*), adjusted by the corresponding Fama–French 5×5 portfolio returns. Litigation risk is measured following Kim and Skinner (2012) (*LitRiskKS*) and Donelson, Flam, and Yust (2022) (*LitRiskDFY*). *AIWashing* × *LitRisk* denotes the interaction between *AI-washing* and the corresponding litigation risk measure. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*BHAR[1,2yr]*			
Variable	(1)	(2)	(3)	(4)
AIWashing×LitRiskKS		**-0.034*** **		
		(-3.10)		
AIWashing×LitRiskDFY				**-0.048** **
				(-2.57)
LitRiskKS	**-0.003**	**-0.001**		
	(-0.32)	**(-0.08)**		
LitRiskDFY			**0.023** **	**0.025*** **
			(2.46)	**(2.81)**
AIWashing	0.002	-0.097	0.001	0.080
	(0.05)	(-1.35)	(0.01)	(1.33)
AIInvestment	0.075*	0.074*	0.069*	0.063
	(1.76)	(1.73)	(1.69)	(1.55)
Controls	Yes	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,472	1,472	1,472	1,472
Adjusted R^2	5.1%	5.1%	5.3%	5.4%

Internet Appendix

"When AI-washing Meets Initial Public Offerings"

This appendix provides supplemental analyses to the main results presented in "When AI-washing Meets Initial Public Offerings". When there is no confusion, additional tests are labelled, adding the extension "IA" for "Internet Appendix" (e.g., Table IA), while the full specifications of the tables reported in the main text are labelled with the original table name. The references to page numbers in parentheses are in relation to the position in the paper where the relevant internet appendix table is cross-referenced.

Figure IA1. Top 20 Most Frequent Keywords in S-1 Files

Figure IA2. Covariate Balance Before and After Propensity Score Matching

Table IA1. IPO Data Cleaning Process

Table IA2. LinkedIn Profiles

Table IA3. Artificial Intelligence Patent Dataset

Table IA4. R&D Expenditure Distribution by Industry

Table IA5. AI Job Postings Dynamics around the Expiration Day

Table IA6. Robustness Tests: Returns Adjusted by Market Returns

Table IA7. Robustness Tests: Different AI Keyword Thresholds

Table IA8. Robustness Tests: AI Washing Index as the Alternative Measure

Table IA9. Litigation Risk Concern and AI Disclosure Retrenchment

Table IA10. AI Expertise of IPO Intermediaries and AI-Washing

Figure IA1. Top 20 Most Frequent Keywords in S-1 Files

This figure reports the 20 most frequently occurring AI-related keywords in the S-1 filings of IPO firms during 2010–2024. Keywords are ordered from left to right by the number of occurrences. The left y-axis reports the number of occurrences of each keyword, and the right y-axis reports the cumulative share of total AI keyword occurrences accounted for by the listed keywords.

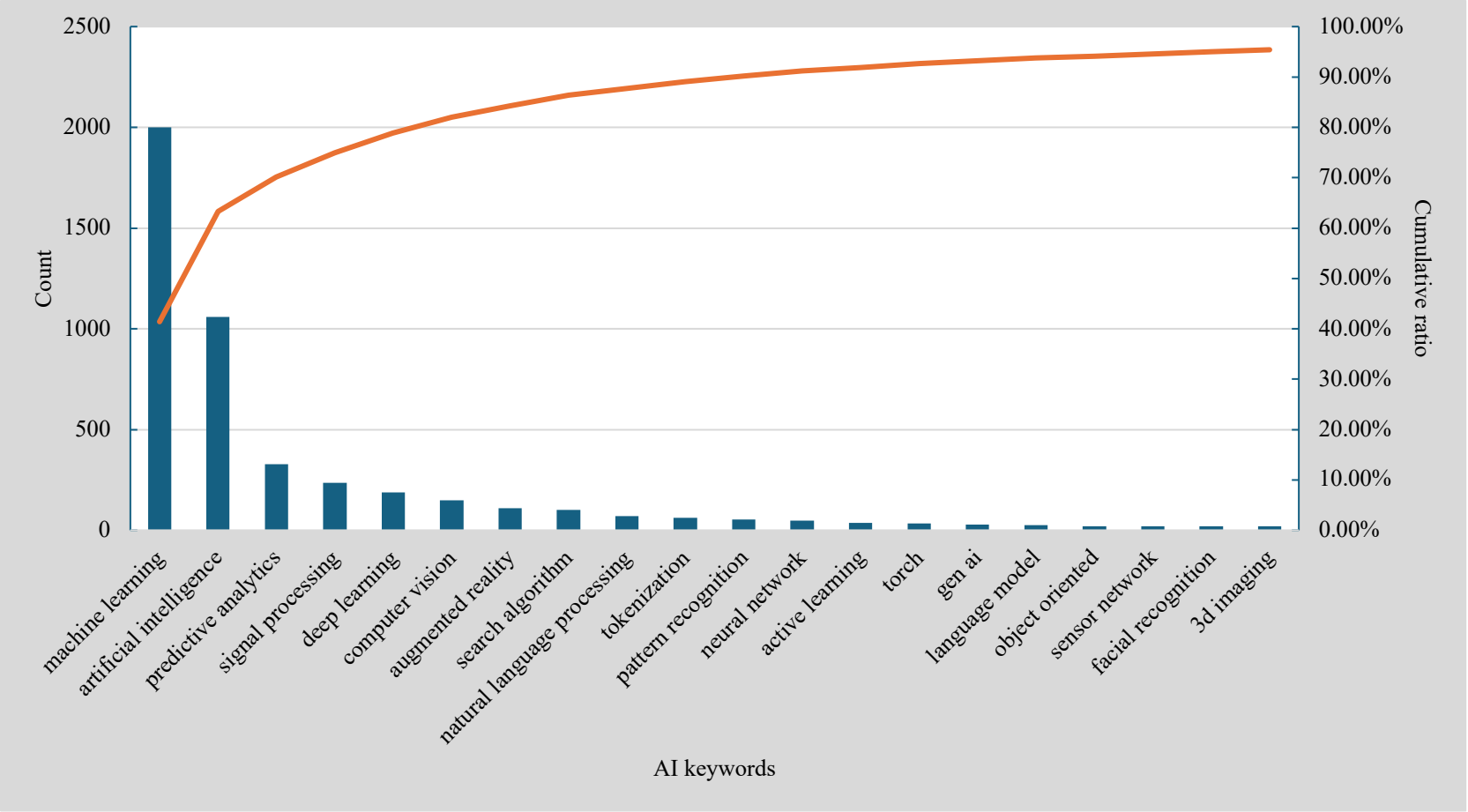


Figure IA2. Covariate Balance Before and After Propensity Score Matching

This figure presents the standardized percentage bias for each covariate before and after propensity score matching. Black circles represent the unmatched sample, whereas red crosses represent the matched sample. Values closer to zero indicate better covariate balance. The vertical dashed lines at −10% and 10% denote the conventional thresholds for acceptable balance, while the vertical line at zero indicates perfect balance. The sign of the standardized bias indicates the direction of the difference between the treatment and control groups.

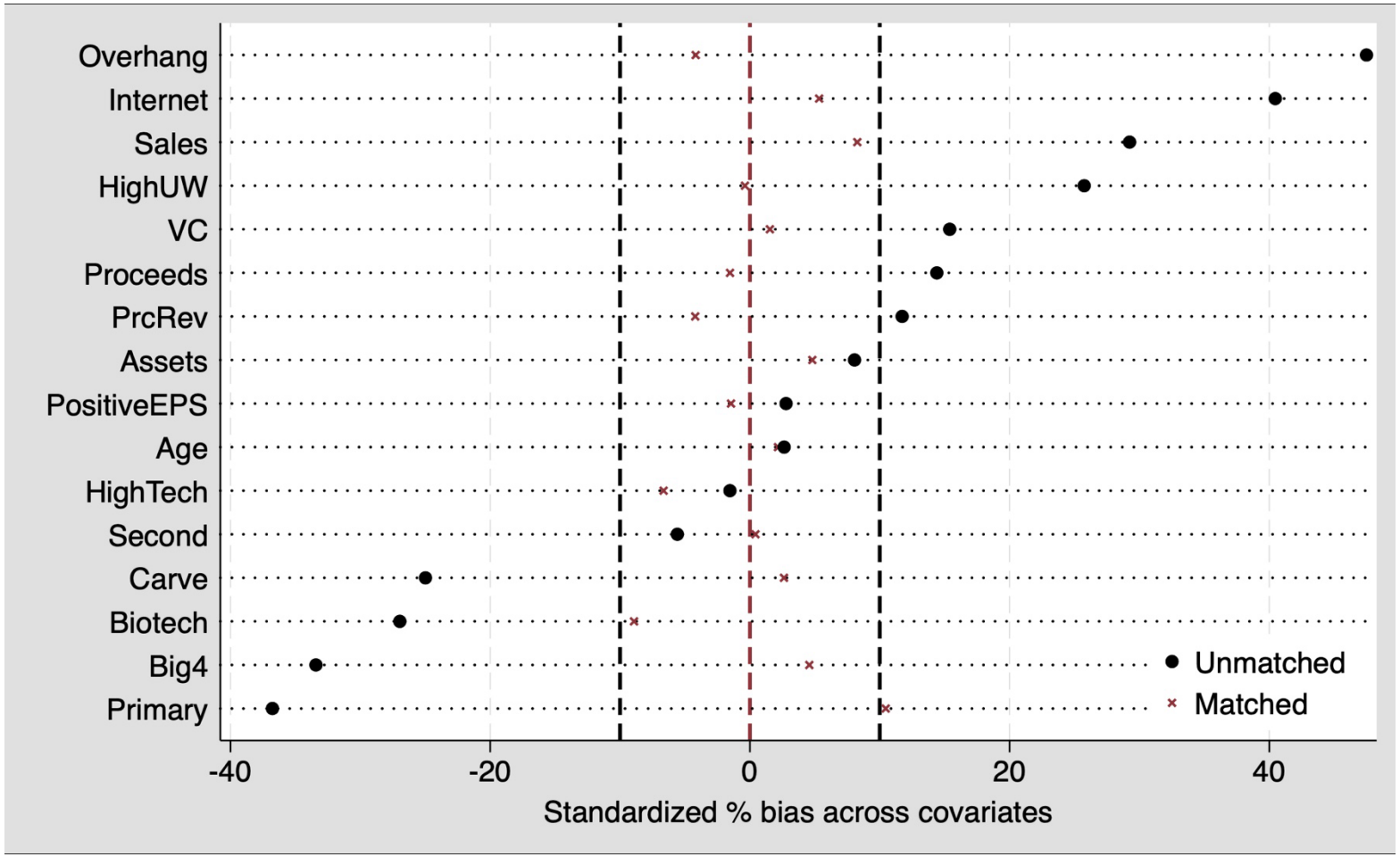


Table IA1. IPO Data Cleaning Process

This table presents each step of our cleaning process for IPO samples. We first download the original transaction records from the SDC Platinum. Following the online appendix of Lowry, Michaely, and Volkova (2017), we identify IPO firms by matching issue-level data to CRSP using *CUSIP* identifiers. We prioritize exact 8-digit CUSIP matches and use 6-digit matches only when necessary, excluding ambiguous mappings. We require the first CRSP trading date to fall within a narrow window around the issue date. To better match IPO data with *Revelio Lab LinkedIn* Data, we use a multi-key matching process (*GVKEY*, *CIK*, and *CUSIP-8*). If one issuing firm can be matched with different *RCID* codes, we choose the one whose name has the highest similarity score with the issuer's name.

IPO sample	**# of firms**
SDC domestic IPOs issued during 2010-2024	4,853
Drop: Non-common shares issues	2,947
Drop: Units, ADR, Closed-end Fund, Offer Price less than 5 dollars, and REIT IPOs	2,085
Drop: stocks not listed on CRSP (CRSP includes Amex, NYSE, and NASDAQ stocks)	1,942
Drop: IPOs are not in Ritter's Files	1,908
Drop: without file number to merge with SEC EDGAR	1,684
Drop: without links with *Revelio Lab* data	1,671

Table IA2. LinkedIn Profiles

This table presents the detailed information about *LinkedIn* profiles from *Revelio Lab*. Panel A presents summary statistics for *LinkedIn* Profiles, including the number of matched firms, positions, users, skills, and education. Panel B shows the AI skills keyword list, the same as Babina et al. (2024), excluding keywords with fewer than 4 characters.

Panel A. Numbers of Profile Information

Companies	Positions	Users	Skills	Educations
1,671	10,051,514	7,155,652	253,255,381	11,859,703

Panel B. AI Skills Keyword List

Artificial Intelligence	MLPACK	Stochastic Gradient Descent	Naive Bayes
Computer Vision	Keras	Gradient boosting	Expectation-Maximization Algorithm
Machine Learning	Theano	Dimensionality Reduction	Clustering Algorithms
Natural Language Processing	Torch	Deep Learning	Matrix Factorization
Kernel Methods	Vowpal Wabbit	DBSCAN	Object Recognition
Microsoft Cognitive Toolkit	Boosting	AI ChatBot	Classification Algorithms
Xgboost	TensorFlow	Recommender Systems	Information Extraction
Sentiment Classification	Vowpal	Random Forests	Image Recognition
Long Short-Term Memory	Convolutional Neural Network	Deeplearning4j	Bayesian Networks
Libsvm	Jung Framework	Support Vector Machines	Supervised Learning
Semi-Supervised Learning	OpenNLP	Unstructured Information Management Architecture	OpenCV
Recurrent Neural Network	Natural Language Toolkit	Apache UIMA	K-Means
Word2vec	Unsupervised Learning	Maximum Entropy Classifier	Sentiment Analysis
Mxnet	Scikit-learn	Hidden Markov Model	Machine Translation
Caffe Deep Learning Framework	Latent Semantic Analysis	PyBrain	Neural Networks
Autoencoders	Latent Dirichlet Allocation	Computational Linguistics	

Table IA3. Artificial Intelligence Patent Dataset

This table presents the detailed information about the *Artificial Intelligence Patent Dataset (AIPD)*. Panel A presents the number of patents in different datasets. Panel B shows the distribution of patent filings from 2007 to 2023.

Panel A. Number of Granted Patents (matched within KPSS)			
AIPD	**KPSS**	**Merged**	**AI Patents**
15,437,542	3,324,963	2,703,723	111,486

Panel B. Number of AI Patents (with granted and pre-granted patents)				
All granted	**All pre-granted**	**All patents (IPOs)**	**AI patents (IPOs)**	**Ratio (AI/All)**
5,050,885	6,255,085	357,206	150,641	42.17%

Table IA4. R&D Expenditure Distribution by Industry

This table reports the distribution of research and development (R&D) expenditure across industries for AI-washing firms, AI-investment firms, and non-AI firms. R&D is measured as research and development expense divided by total assets. Number (N) denotes the number of firm observations in each industry–firm category. A dash ("–") indicates that no firm observations are available in the corresponding category and therefore the average R&D cannot be calculated. Industry classifications follow the Fama–French 12-industry classification. Panel A compares average R&D intensity before and after the IPO across the three firm categories. Panel B presents the industry-level distribution of average R&D intensity and the number of firm observations for each firm category. R&D represents the average ratio of research and development expense to total assets within each industry–firm category.***, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Univariate Analyses of R&D

	Non-AI		AI-investment		AI-washing		Non-AI versus AI-washing		AI-investment versus AI-washing	
	N	Mean	N	Mean	N	Mean	Diff	t-value	Diff	t-value
R&D (before IPO)	539	0.308	278	0.209	56	0.193	0.115	1.988**	0.016	0.582
R&D (after IPO)	545	0.151	275	0.142	53	0.119	0.032	1.542	0.023	1.373

Panel B. Industrial Distribution of R&D

Industry	AI-washing		AI-investment		Non-AI	
	R&D	Num	R&D	Num	R&D	Num
Consumer Nondurables	0.015	2	-	0	0.017	11
Consumer Durables	0.044	1	0.105	5	0.035	4
Manufacturing	0.252	1	0.070	6	0.019	17
Oil, Gas, and Coal Extraction and Products	-	0	-	0	0.134	2
Chemicals and Allied Products	0.024	1	0.009	2	0.096	6
Business Equipment	0.207	16	0.205	183	0.176	80
Telephone and Television Transmission	-	0	0.258	4	0.067	1
Wholesale, Retail, and Some Services	0.047	4	0.044	12	0.000	61
Healthcare, Medical Equipment, and Drugs	0.285	23	0.306	54	0.507	333
Finance	0.046	5	0.130	3	0.066	4
Others	0.051	3	0.147	9	0.270	20

Table IA5. AI Job Postings Dynamics around the Expiration Day

This table reports the dynamics of AI-related job postings around the expiration day estimated by Poisson regressions. The dependent variable, *AIJobPost*, is calculated as the number of quarterly AI Job postings. Event quarter *Quarter_k* is defined relative to the quarter of the expiration day, where the benchmark is the *quarter_-1*. Columns with controls include the full set of control variables defined in Appendix A. All specifications include industry × year_quarter and relative-quarter fixed effects. *t-statistics* are reported in parentheses. Key results are highlighted in bold. The sample period is from January 1, 2010 to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*AIJobPost*
Variable	(1)
AIWashing×*Quarter_-3*	**-0.075**
	(-0.13)
AIWashing×*Quarter_-2*	**-0.502**
	(-1.01)
AIWashing×*Quarter_0*	**-0.689***
	(-1.74)
AIWashing×*Quarter_1*	**-0.717**
	(-1.34)
AIWashing×*Quarter_2*	**-1.154*****
	(-5.24)
AIWashing×*Quarter_3*	**-1.359*****
	(-5.91)
Controls	YES
Fixed Effects	Ind × Year_quarter & Relative_quarter
Observations	2,350
Pseudo R^2	34.8%

Table IA6. Robustness Tests: Returns Adjusted by Market Returns

This table reports regressions of long-run IPO performance on AI-washing IPOs and AI-investment IPOs with alternative returns market-return adjustments. The main dependent variables are the buy-and-hold returns across different periods, denoted by *BHAR[0,1yr]*, *BHAR[0,2yr]* and *BHAR[0,3yr]*. Abnormal returns are adjusted using value-weighted market returns in Column (1)-(3) and the equal-weighted market returns in Column (4)-(6). AI-washing IPOs (*AIWashing*) and AI investment IPOs (*AIInvestment*) are defined the same as Table 4. Columns with controls include the full set of control variables defined in Appendix A. We control for fixed effects at the Fama-French 12 industry × year level. *t-statistics* reported in parentheses are based on robust standard errors clustered at the two-digit industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Returns adjusted by value-weighted market returns			Returns adjusted by equal-weighted market returns		
Dep. Variable=	*BHAR[0,1yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*BHAR[0,1yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*
Variable	(1)	(2)	(3)	(4)	(5)	(6)
AIWashing	**-0.055**	**-0.075***	**-0.156****	**-0.052**	**-0.075***	**-0.156****
	(-0.67)	**(-1.68)**	**(-2.31)**	**(-0.63)**	**(-1.78)**	**(-2.31)**
AIInvestment	**-0.028**	**0.073****	**0.169*****	**-0.032**	**0.076****	**0.167*****
	(-1.03)	**(2.05)**	**(4.51)**	**(-1.11)**	**(2.10)**	**(4.45)**
Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	Ind × Year	Ind × Year	Ind × Year	Ind × Year	Ind × Year	Ind × Year
Observations	1,581	1,538	1,498	1,574	1,531	1,491
Adjusted R^2	12.3%	10.7%	9.7%	12.5%	9.0%	7.4%
AIWashing - AIInvestment	**-0.02**	**-0.15**	**-0.32**	**-0.03**	**-0.15**	**-0.33**
Test of difference [p-value]	**[0.77]**	**[0.00]**	**[0.00]**	**[0.71]**	**[0.00]**	**[0.00]**

Table IA7. Robustness Tests: Different AI Keyword Thresholds

This table reports robustness tests using alternative percentile cutoffs (98th, 90th, 88th) to define AI-washing IPOs. Panels A–C report regressions for one-, two-, and three-year buy-and-hold abnormal returns, respectively, adjusted by the Fama–French 5×5 portfolio returns. *AIWashing* equals one when a firm's S-1 AI disclosure exceeds the indicated cutoff, calculated using prior-year IPOs in the same Fama–French 12 industry, while both pre-IPO AI employees and AI patents are zero. All specifications include the controls defined in Appendix A and two-digit SIC industry and year fixed effects. *t-statistics* in parentheses are based on standard errors clustered by industry; p-values for the difference between *AIWashing* and *AIInvestment* are reported in brackets. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*BHAR[0,1yr]*			*BHAR[0,2yr]*			*BHAR[0,3yr]*		
	98th	90th	88th	98th	90th	88th	98th	90th	88th
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AIWashing	**-0.109**	**-0.103**	**-0.097**	**-0.110***	**-0.082****	**-0.102****	**-0.172****	**-0.153*****	**-0.162*****
	(-1.25)	**(-1.60)**	**(-1.50)**	**(-1.81)**	**(-2.08)**	**(-2.19)**	**(-2.31)**	**(-3.31)**	**(-3.32)**
AIInvestment	**-0.026**	**-0.039**	**-0.039**	**0.097****	**0.088****	**0.083***	**0.209*****	**0.191*****	**0.187*****
	(-1.02)	**(-1.57)**	**(-1.57)**	**(2.39)**	**(2.06)**	**(1.93)**	**(4.49)**	**(4.09)**	**(3.94)**
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,581	1,538	1,498	1,581	1,538	1,498	1,581	1,538	1,498
Adjusted R^2	14.5%	14.6%	14.6%	10.1%	10.1%	10.1%	7.8%	7.8%	7.9%
AIWashing - AIInvestment	**-0.08**	**-0.06**	**-0.06**	**-0.21**	**-0.17**	**-0.19**	**-0.38**	**-0.34**	**-0.35**
Test of difference [p-value]	**[0.37]**	**[0.34]**	**[0.37]**	**[0.02]**	**[0.00]**	**[0.00]**	**[0.00]**	**[0.00]**	**[0.00]**

Table IA8. Robustness Tests: AI Washing Index as the Alternative Measure

This table reports return regressions using a residual-based continuous measure of AI-washing. The AI-washing index (*AIWashingIndex*) is the standardized residual from a Poisson regression of AI disclosure (*AIWord*) on pre-IPO AI employees (*AIEmployee*) and AI patents (*AIPatent*), controlling for two-digit SIC industry and IPO year fixed effects. A higher value indicates that AI disclosure exceeds the level predicted by observable AI inputs. All specifications include the full set of control variables defined in Appendix A. *t-statistics* are reported in parentheses based on industry-clustered standard errors. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Long-term returns			Short-term returns
Dep. Variable=	*BHAR[0,1yr]*	*BHAR[0,2yr]*	*BHAR[0,3yr]*	*InitialReturn*
Variable	(1)	(2)	(3)	(4)
AIWashingIndex	**0.001**	**-0.036*****	**-0.043*****	**0.016*****
	(0.03)	**(-3.69)**	**(-2.90)**	**(3.81)**
Controls	YES	YES	YES	YES
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	1,505	1,464	1,426	1,574
Adjusted R^2	14.3%	10.3%	7.6%	23.1%

Table IA9. Litigation Risk Concern and AI Disclosure Retrenchment

This table reports regression results on the effect of AI-washing IPOs (*AIWashing*) on the post-IPO AI disclosure gap (*AIGap*). In Panel A, the post-IPO AI disclosure gap (*AIGap*) is defined as the difference between the total number of AI words (business description, risk factors and MD&A sections) in the first 10-K filings one year after the IPO, minus the number of AI words disclosed in the S-1 prospectus. In Panel B, the post-IPO AI disclosure gap (*AIGap*) is defined as the difference between the AI words intensity with the second year (the fifth quarter to the eighth quarter) after the IPO, minus the AI words intensity in the S-1 prospectus. We follow Kim and Skinner (2012) (*LitRiskKS*) and Donelson, Flam, and Yust (2022) (*LitRiskDFY*) to measure the litigation risk. *AIWashing* × *LitRisk* denotes interaction terms between the litigation risk (*LitRisk*) and the AI-washing IPOs (*AIWashing*). Columns with controls include the full set of control variables defined in Appendix A. We control for fixed effects at the two-digit industry and year levels. *t-statistics* reported in parentheses are based on robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. AI disclosure gap between the 10-K and the S-1 prospectuses

Dep. Variable=	*AIGapTenK*				
Variable	(1)	(2)	(3)	(4)	(5)
AIWashing×LitRisk_KS			**-1.328***		
			(-1.76)		
AIWashing×LitRisk_DFY					**-0.547***
					(-1.92)
LitRisk_KS		-0.001	0.146		
		(-0.01)	(1.43)		
LitRisk_DFY				-0.082	0.005
				(-0.83)	(0.04)
AIWashing	-0.136	-0.159	-3.926	-0.153	0.906
	(-0.13)	(-0.14)	(-1.35)	(-0.15)	(1.31)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	214	212	212	214	214
Adjusted R^2	1.7%	1.5%	3.6%	1.2%	1.7%

Panel B. AI disclosure gap between the conference calls and the S-1 prospectuses

Dep. Variable=	*AIGapCall*				
Variable	(1)	(2)	(3)	(4)	(5)
AIWashing×LitRisk_KS			**-0.188*****		
			(-3.28)		
AIWashing×LitRisk_DFY					**-0.118***
					(-1.83)
LitRisk_KS		0.009	0.044		
		(0.49)	(1.21)		
LitRisk_DFY				0.067*	0.078**
				(1.87)	(2.21)
AIWashing	-0.157	-0.161	-0.698***	-0.127	0.131
	(-1.05)	(-1.07)	(-3.57)	(-0.85)	(0.48)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year	Ind & Year	Ind & Year
Observations	374	373	373	374	374
Adjusted R^2	12.7%	12.3%	12.8%	13.2%	13.4%

Table IA10. AI Expertise of IPO Intermediaries and AI-Washing

This table reports Logit regressions of *AIWashing* on the AI expertise of IPO intermediaries. *VCAISkill* is an indicator equal to one if the lead venture capitalist has AI-related skills identified from Revelio Labs employee profiles. *UWAIExp* is an indicator equal to one if the lead underwriter has previously underwritten at least one AI-investment IPO before the focal IPO. *AuditorAIExp* is an indicator equal to one if the auditor has previously audited at least one AI-investment IPO before the focal IPO. All specifications include the full set of control variables defined in Appendix A, as well as two-digit industry and year fixed effects. The reported t-statistics, shown in parentheses, are based on heteroskedasticity-robust standard errors clustered at the industry level. Key results are highlighted in bold. The sample period is from January 1, 2010, to December 31, 2024. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable=	*AIWashing*		
Variable	(1)	(2)	(3)
VCAISkill	**-0.066**		
	(-0.31)		
UWAIExp		**-0.106**	
		(-0.23)	
AuditorAIExp			**-0.082**
			(-0.28)
Controls	Yes	Yes	Yes
Fixed Effects	Ind & Year	Ind & Year	Ind & Year
Observations	1,436	1,436	1,372
Pseudo R^2	11.0%	11.0%	12.0%