AI's Blind Spot: The Mispricing of Breakthrough Innovation

Shubo Kou* Zinuo Zhang†

Nankai University

July 2026

ABSTRACT

While AI models efficiently process standard technological disclosures, we show they systematically fail to recognize radical innovations, with severe consequences for asset pricing. Analyzing U.S. utility patents granted since 1976, we find a non-monotonic, reversed U-shaped relationship between machine-predicted patent grant probabilities (via XGBoost and LLM embeddings) and true economic quality. Crucially, AI models fail in the extreme upper tail, systematically assigning low quality scores to the most transformative breakthroughs. This AI underrecognition reflects genuine economic content rather than model noise: while machine-predicted scores fail to forecast firm performance, the breakthrough innovations missed by AI strongly predict future ROA improvements and higher acquisition target probability. Consequently, this mismatch creates severe market frictions, generating substantial long-term post-patent stock return drift (over 60 to 900 trading days) for firms holding AI-underrecognized breakthroughs, whereas average-scoring patents show no drift. Examining the ex-post shock of Generative AI (LLMs), we show that advanced models reduce information asymmetry for moderate-quality patents—erasing their return drift—yet remain blind to true breakthroughs. As a result, post-GPT return drift disappears for moderate innovations but dramatically intensifies for the unmapped radical breakthroughs. Our findings demonstrate both the power and limits of AI in financial markets, showing that machine learning progress can inadvertently deepen mispricing for creative disruptions.

Keywords: Artificial Intelligence; Patent Screening; Patent Value; Technological Innovation; Future Profitability; Corporate Acquisitions; Asset Pricing

*Email: koushubo@nankai.edu.cn; phone: 18801259487.

†Email: 2313210@mail.nankai.edu.cn; phone: 13339998118.

I. Introduction

Capital markets allocate capital by transforming costly signals into actionable asset prices (Grossman and Stiglitz, 1980; Bond et al., 2012). For corporate innovation, this price-discovery process is uniquely difficult: evaluating radical breakthroughs requires deciphering complex technical disclosures under vast uncertainty (Manso, 2011). The rise of artificial intelligence promises to resolve these information frictions through automated processing of complex technological data. However, if AI-based evaluation models fail precisely where information asymmetry is most severe—at the frontier of radical innovation—the technology may unintentionally introduce systemic valuation blind spots, altering both asset prices and corporate trajectories.

Patent disclosures present a uniquely demanding empirical setting to test this tension. Translating complex technical claims, legal boundaries, and prior art into economic valuation requires joint assessments of commercial viability, product-market competition, and downstream technological trajectories (Bena and Li, 2014; Celik et al., 2022). Crucially, corporate innovation value is sharply right-skewed: a small fraction of breakthrough inventions generates the vast majority of economic value and technological influence (Trajtenberg, 1990; Hall et al., 2000; Kogan et al., 2017). Consequently, average classification accuracy is insufficient to establish economic utility. An algorithm may reliably screen ordinary innovations while remaining structurally blind to the rare, upper-tail breakthroughs that drive firm value. This paper asks whether state-of-the-art AI evaluation models truly recognize these economically decisive innovations.

Two competing economic views frame our inquiry. Under the information-processing-cost view, artificial intelligence converts high-dimensional, unstructured disclosures into standardized, low-noise signals at a scale beyond human capacity (Mullainathan and Spiess, 2017; Brynjolfsson and Mitchell, 2017). If AI-based quality scores successfully distill fundamental economic potential, their predicted grant probabilities should align positively with post-grant market value, technological diffusion, and breakthrough orientation across the entire quality distribution. Conversely, under the technological-underrecognition view, data-driven models thrive on historical precedents and frequent comparables, making them inherently ill-equipped for paradigm-shifting innovations. Radical breakthroughs lack familiar textual anchors and depend on emerging markets or unestablished complementary assets. Thus, while AI algorithms effectively rank incremental innovations, they bend away from true upper-tail successes. This perspective aligns with theories in which cheap, coarse signals crowd out costly, in-depth fundamental analysis (Dugast and Foucault, 2018), evidence that alternative data shifts market focus toward short-term information (Dessaint et al., 2024), and findings that algorithmic screening improves average startup selection while systematically starving rare, highly disruptive ventures of capital (Bonelli,

2026).

Our inquiry connects to two active literature streams. The asset-pricing and innovation literature establishes that market reactions, citations, and patent text capture distinct economic dimensions of technological progress (Kogan et al., 2017; Kelly et al., 2021), with key insights linking innovation frictions to firm misvaluation and return predictability (Cohen et al., 2013; Hirshleifer et al., 2013, 2018; Bena and Garlappi, 2020; Shu et al., 2022). Simultaneously, a growing body of research shows that algorithmic screening efficiently extracts informative signals from patent disclosures and complements human examiner expertise (Yang, 2023; Zheng, 2025). What remains fundamentally unresolved, however, is whether AI grant-prediction models preserve the correct cross-sectional ordering for the most valuable inventions. We bridge this gap by examining the upper-tail quality margin, testing whether systematic disagreement between machine-predicted scores and realized technological benchmarks generates persistent capital-market mispricing and alters corporate trajectories.

To address this question, we construct a comprehensive dataset of U.S. utility patents granted since 1976 and evaluate two distinct machine-learning architectures. The first is an XGBoost model built on word2vec representations and application-level features. The second leverages large-language-model embeddings combined with a neural-network classifier, adapting the predictive-patentomics framework of Yang (2023). Both models generate estimated grant probabilities, where higher scores reflect stronger alignment with historically successful examination features. We evaluate these machine scores against three fundamental benchmarks: KPSS grant-window stock-market value, forward citations, and q_rank, an ex post text measure quantifying directional technological breakthroughs. Our empirical design proceeds in three steps: tracing full-distribution nonlinearities, estimating post-grant return dynamics via firm-grant-date regressions, and evaluating corporate operating and acquisition outcomes alongside calendar-time portfolio asset-pricing tests. These designs document rankings and conditional associations; they do not identify causal adoption effects.

Figure 1 establishes the paper's central empirical fact by plotting binned mean XGBoost and LLM grant probabilities across the full quantile distributions of KPSS private value, total forward citations, and directional breakthrough orientation (q_rank). For all three benchmarks, the relation is starkly non-monotonic: predicted grant scores rise from lower quantiles toward intermediate quality levels, but flatten and reverse sharply in the extreme upper tail. Across both model architectures—particularly for the LLM embeddings—the highest-value, most highly cited, and most directionally transformative patents systematically receive lower average grant-probability scores than ordinary, mid-tier inventions. Thus, while machine-learning models effectively extract signals relevant to routine technological quality, their cross-sectional ordering structurally collapses precisely where incremental economic and technological value is greatest.

This empirical pattern provides direct support for the technological-underrecognition mechanism. Supervised algorithms trained on historical examination outcomes inherently excel at recognizing familiar technological patterns with abundant precedents. Radical innovations, however, derive their value from novel technological combinations, intangible commercial opportunities, and prolonged diffusion timelines—features characterized by severe recombinant uncertainty and delayed recognition (Fleming, 2001; Singh and Fleming, 2010; Wang et al., 2017). As Cao et al. (2024) demonstrate, algorithmic performance degrades when information becomes unstructured, sparse, or non-standard. Consequently, machine-predicted scores do not merely exhibit mean-zero idiosyncratic noise around a positive relationship; their cross-sectional rankings systematically bend away from the creative disruptions that drive long-term economic growth. We call this pattern technological underrecognition.

[Insert Figure 1 here]

Focusing specifically on the upper tail sharpens rather than dilutes this full-distribution pattern. Restricting the sample to the top decile of each benchmark reveals a consistently downward-sloping relationship for both XGBoost and LLM scores: patents generating greater market value, broader technological diffusion, or deeper directional impact systematically receive lower AI grant probabilities. The q_rank benchmark provides an especially decisive test because both q_rank and our AI grant scores are derived directly from patent text; the divergence therefore cannot be dismissed as a mismatch between textual and non-textual metrics. The right-tail decline also remains robust across five-, ten-, and twenty-year citation horizons, reducing concerns about cohort-specific citation exposure. Collectively, these diagnostics show that technological underrecognition is a persistent structural property of economically vital innovations rather than an artifact of benchmark selection or horizon specification.

We next investigate whether this algorithmic underrecognition translates into persistent capital-market mispricing. At the firm-grant-date level, patent batches containing an upper-tail breakthrough invention generate economically large, progressively increasing post-grant abnormal returns, rising from roughly 0.5 percentage points over 60 trading days to 4.5–5.0 percentage points over a 900-trading-day horizon. Conditional on the presence of an upper-tail breakthrough, mean AI grant scores provide no incremental return-predictive power and yield significantly negative coefficients across multiple specifications. Joint-benchmark tests confirm that market-value and technological-impact signals are not absorbed by machine scores. These findings indicate that the economic content missed by AI models represents unpriced information, generating protracted post-patent return drift as financial markets slowly incorporate the value of creative disruptions. The estimates are descriptive and do not attribute a firm's return path to a single patent.

Calendar-time heterogeneity around the GPT-3 release provides a further test of the evolution of this friction. In the later period, post-patent return drift weakens for moderately high-quality patents but becomes markedly more pronounced for the smaller set of frontier breakthroughs that remain poorly ranked by AI scores. This contrast is consistent with generative AI reducing information-processing frictions for standard innovations with familiar textual precedents while remaining blind to true creative disruptions. As algorithmic adoption resolves information asymmetry for incremental advances, it can paradoxically isolate the unmapped upper-tail breakthroughs and deepen their capital-market mispricing. Because the cutoff also separates market regimes, patent cohorts, and data environments, we interpret this evidence as descriptive heterogeneity rather than a causal effect of GPT release or AI adoption.

Market returns, however, represent only one dimension through which technological underrecognition manifests. We show that this algorithmic blindness extends directly to real operating performance and the market for corporate control. Mean AI grant probabilities exhibit no positive predictability for one-, two-, or three-year future ROA changes, whereas upper-tail benchmark indicators forecast subsequent operating-profitability improvements: high KPSS value is associated with stronger second-year ROA gains, early citation leaders predict gains at shorter horizons, and the breakthrough-direction relation is weaker but directionally similar. A parallel contrast emerges in corporate-control transactions. Higher AI scores are negatively associated with the probability that a firm becomes an acquisition target, particularly over cumulative two- and three-year horizons, whereas firms holding AI-underrecognized breakthrough patents are more likely to be acquired. The convergence of return drift, operating outperformance, and acquisition-target incidence confirms that upper-tail underrecognition reflects genuine economic content, driving both capital-market mispricing and subsequent real-asset reallocation across firm boundaries.

Finally, we translate this underrecognition divergence into an investable trading strategy. Retrospective portfolios that exploit the mismatch—buying firms with high realized patent quality but low AI-predicted scores—earn monthly alphas of 0.258%–0.385% under the Fama–French three-factor model and 0.369%–0.528% under the five-factor model. The full-sample alpha is absorbed by the Hou, Mo, Xue, and Zhang (2021) q5 factor model, driven primarily by investment, profitability, and expected-growth exposures. This attenuation delineates the empirical boundary of the result: the capital-market mispricing induced by AI underrecognition is closely tied to fundamental risk factors, demonstrating that algorithmic evaluation limits reflect economically meaningful real-side frictions rather than ungrounded price noise.

Our paper makes three principal contributions. First, we advance the economics of artificial intelligence and automated decision-making by revealing a structural boundary of machine learning in high-stakes financial environments. Prior literature emphasizes that big data and machine learning reduce information-processing costs and improve screening accuracy across lending, venture capital,

and intellectual property (Mullainathan and Spiess, 2017; Brynjolfsson and Mitchell, 2017; Yang, 2023; Zheng, 2025). We shift the analytical focus from average classification accuracy to the cross-sectional sorting of the economically decisive right tail. Algorithmic efficiency at the mean coexists with structural failure at the frontier: machine-learning models effectively evaluate routine, incremental technological progress but experience technological underrecognition when confronted with radical breakthroughs.

Second, we contribute to innovation finance and asset pricing by identifying a technology-driven friction that generates protracted post-patent return drift. Extant literature establishes that innovation disclosures contain substantial unpriced information because of limited attention and processing complexity (Cohen et al., 2013; Hirshleifer et al., 2013, 2018; Kogan et al., 2017; Kelly et al., 2021). We show how algorithmic screening reshapes these dynamics. Because AI models assign lower quality scores to breakthrough innovations, firms holding these unmapped technologies experience persistent return drift over 60 to 900 trading days. By linking AI underrecognition to future ROA improvements, acquisition-target incidence, and calendar-time factor models (Hou et al., 2021), we show that this drift is not transient price noise; it reflects fundamental value eventually resolved through internal commercialization, operating profitability, and real-asset reallocation.

Third, we enrich the literature on information acquisition, alternative data, and financial-market efficiency. Classical theory posits that lower information-acquisition costs enhance price discovery and allocative efficiency (Grossman and Stiglitz, 1980; Bond et al., 2012). Recent theoretical and empirical work, however, suggests that cheap, standardized alternative data can crowd out fundamental analysis or shift market focus toward short-term, easily quantifiable information (Dugast and Foucault, 2018; Dessaint et al., 2024; Bonelli, 2026). We provide evidence of this mechanism in the context of advanced AI models and generative breakthroughs. New generations of AI reduce processing frictions for standard innovations and erase their return drift, yet remain structurally blind to rare creative disruptions. Scalable information processing can therefore accelerate the pricing of ordinary assets while deepening the isolation and mispricing of frontier innovations.

The remainder of the paper is organized as follows. Section II describes the data, AI-based grant-prediction models, and patent-benchmark construction. Section III investigates real economic outcomes, focusing on operating profitability and acquisition-target incidence. Section IV details the empirical evidence on full-distribution non-monotonicity and upper-tail technological underrecognition. Section V examines the financial-market consequences, including post-grant return drift, joint-benchmark dynamics, and temporal heterogeneity around the Generative AI period. Section VI presents calendar-time portfolio performance and factor-model asset-pricing tests. Section VII concludes. The appendix contains variable definitions, supplementary return regressions, non-overlapping-event robustness checks, portfolio characteristics, and raw holding-period returns.

II. Data, Patent Benchmarks, and Summary Statistics

2.1. U.S. Patent, Firm, and Capital-Market Data

Our empirical design links U.S. utility patents granted from January 1976 through December 2024 to the market and corporate outcomes through which technological underrecognition may become economically consequential. The merged data combine patent text, application characteristics, firm identifiers, stock returns, accounting variables, acquisition records, and factor returns. We use the USPTO–Compustat link in Kogan et al. (2017) to map patent numbers to GVKEYs and to obtain KPSS patent values and citations, and we merge these records with PatentsView claims, descriptions, CPC classifications, and application characteristics. CRSP supplies daily and monthly returns; Compustat supplies firm characteristics and ROA; strict Orbis target matches identify completed acquisitions; and the portfolio tests use the Fama–French and HXZ factor files (Fama and French, 1993, 2015; Hou et al., 2015).

The analysis is organized at four levels, each corresponding to one link in the paper's evidence chain. Patent-level observations reveal whether machine-predicted grant probabilities preserve the ordering of inventions in the upper tails of market value, technological diffusion, and breakthrough direction. Firm-grant-date observations combine all patents awarded to the same firm on the same day, relating the batch's mean grant score and upper-tail patent content to a single subsequent return path. Firm-year observations aggregate the same score and benchmark information to examine future ROA changes and acquisition-target incidence. Firm-month observations translate the high-benchmark, low-AI mismatch into factor-adjusted portfolio tests. This progression connects the patent-level ranking failure to capital-market prices, operating outcomes, corporate-control transactions, and investable returns without assigning a firm's subsequent path to any single patent.

The primary firm-grant-date outcomes are size-adjusted cumulative returns over 60, 300, 600, and 900 trading days; the annual sorting figure also reports 1,200 days when a complete window is available. The tables retain the label PEPD, but the event is a patent grant, and the text therefore refers to post-grant cumulative abnormal returns. At the annual level, the profitability outcome is the change in ROA from year $t+h-1$ to year $t+h$, for $h = 1,2,3$. We winsorize these annual ROA changes at the 1st and 99th percentiles within grant year. The acquisition indicators equal one when the patenting firm is strictly matched as a completed Orbis target within one, two, or three years.

Different outcomes impose different information horizons. In the corporate-outcome tests, citation quality is measured by citations received during the first five years after grant. Because incomplete 2020–2024 citation windows are recorded as low counts rather than missing values, we define the five-year-

citation benchmark only for the 1976–2019 grant cohorts; all other benchmarks retain their maximum available samples. Exact duplicate patent-number records are counted once in the firm-year aggregation. The portfolio analysis compounds CRSP returns to the monthly frequency and estimates time-series regressions for equal-weighted long-only validation portfolios. Effective sample endpoints follow the information required by each test: XGBoost scores extend through 2023, fixed-window citations and q_rank require subsequent patent information, future accounting outcomes require completed fiscal-year coverage, and longer return horizons end earlier. Baseline regressions retain all observations with the relevant score, benchmark, outcome, and lagged controls. Annual sorts and portfolio formation use the stated 2010–2024 window and retain portfolio months with at least five firms.

The empirical comparison deliberately separates model-implied scores from realized patent importance. KPSS captures private value revealed around grant, citations capture subsequent technological use, and q_rank captures whether a patent redirects later patent text. Because a patent can rank highly on one dimension without ranking highly on the others, consistent evidence across the three benchmarks is more informative than reliance on any single definition of quality. The AI variables are retrospectively constructed grant-probability scores based on patent and application information and, in the LLM specification, firm size. They provide the screening ranks whose economic ordering we test; they are not themselves treated as measures of realized economic quality. Throughout the paper, "technological underrecognition" refers to the upper-tail pattern in which patents of greater realized economic or technological importance receive lower machine-predicted grant probabilities.

2.2. Patent-Value Benchmarks and AI Grant-Probability Scores

The central test requires a sharp distinction between what an invention ultimately becomes and what an automated screening model recognizes from its grant-relevant inputs. The three benchmarks measure realized economic value, subsequent technological use, and ex post breakthrough direction. The two model-implied grant probabilities instead generate the rankings that an automated screening rule would use. Technological underrecognition is identified from the divergence between these two orderings, especially in the economically decisive right tail.

2.2.1. Realized and Ex Post Patent-Value Benchmarks

Forward citations measure subsequent technological use and diffusion (Trajtenberg, 1990; Hall et al., 2000). Total citations provide the broad, lifetime-oriented measure commonly used in the patent literature, whereas fixed windows place cohorts on a comparable exposure horizon. The five-year benchmark follows the citation-lag evidence in Hall et al. (2001): estimated forward-citation flows peak

at roughly three to five years for most technology fields. It therefore captures the most active early diffusion period. Ten- and twenty-year counts test whether the relation persists deeper into the citation life cycle.

KPSS, developed by Kogan et al. (2017), combines the firm's short-window stock-price response to a patent grant with its pre-event market capitalization. The measure captures private value revealed to equity investors around grant rather than the patent's eventual cash flows. Its event window overlaps with the post-grant return outcomes only on day $g+1$. The principal return evidence therefore comes from the growth in the estimated coefficient from 60 to 900 trading days and from its persistence when successive grants are prevented from reusing the same return segments in Appendix Table A2.

The third benchmark, q_rank, captures directional technological breakthroughs. Following the long-run text-measurement logic of Kelly et al. (2021), it uses 1,536-dimensional embeddings and within-CPC-year ranks to identify patents that are distant from prior work and toward which subsequent patent text moves. This dimension is economically distinct from value revealed at grant. Breakthrough inventions may require complementary assets, new product markets, and extended experimentation (Golder et al., 2009; O'Connor and Rice, 2013), and unusually novel ideas often face delayed recognition (Van Raan, 2004; Wang et al., 2017). Because q_rank uses subsequent patent text, it is an ex post benchmark for evaluating historical grant scores, not a signal that would have been observable at the grant date.

For each focal patent, we compute cosine similarity to patents granted during the preceding five years and the following five years within the relevant CPC group:

$$\mathrm{Sim}(P_i, P_j) = \frac{P_i \cdot P_j}{\|P_i\| \, \|P_j\|}. \tag{1}$$

We convert the raw similarities into within-CPC-year ranks and retain technology-year cells with at least 30 observations. A patent receives a higher novelty rank when it is less similar to preceding patents and a higher forward-direction rank when subsequent patents are more similar to it. q_rank is the product of these two rank components and is therefore high only when a patent both departs from prior work and helps orient subsequent innovation. Patents assigned to multiple CPC groups can generate multiple group-level values; the merged patent-level file retains the maximum q_rank across those assignments. Ranking within technology-year cells makes the measure comparable across fields and cohorts. The target cohorts span 1981–2018 so that both five-year windows can be formed from the 1976–2023 text universe. Embedding coverage and the forward-looking construction produce a shorter effective sample, reported separately in the relevant tables.

2.2.2. Construction and Interpretation of AI Grant-Probability Scores

We construct two model-implied grant probabilities. Grant prediction is a scalable screening task, and a probability-based rule would give greater priority to applications assigned higher probabilities. Throughout the paper, "AI score" refers to this predicted grant probability, not to realized patent value. Utility-patent examination screens novelty, usefulness, nonobviousness, and adequate disclosure, so a well-calibrated grant model should assign higher probabilities to applications with stronger grant-relevant attributes. Prior research treats predicted application success as an informative signal and shows that machine-learning grant models contain information related to citation-based quality and firm outcomes (Yang, 2023; Zheng, 2025). Separate evidence links patent quality and influence to citations, market value, firm value, and future returns (Trajtenberg, 1990; Hall et al., 2000; Kogan et al., 2017; Shu et al., 2022). Under the information-processing-cost view developed in the Introduction, the resulting score ordering should therefore remain positively aligned with private value, technological diffusion, and breakthrough direction. Whether that alignment survives in the upper tail is the paper's central empirical question.

Following the patent-screening design in Zheng (2025), the first model uses XGBoost to predict patent grant probability. We convert patent claims into 50-dimensional word2vec embeddings and combine these textual representations with application characteristics reported in Zheng (2025), including small-entity status, claim length and dependency structure, prior-art citations, originality measures, inventor location, and text-based similarity measures. For each focal application, we compute the average, maximum, and minimum cosine similarity to all prior patents and to the most recent 10, 100, and 1,000 prior patents. These variables summarize the application's proximity to existing technologies in the patent text space. The fitted XGBoost model produces one predicted grant probability for each application.

To train the model, we use the *pg_granted_pgpubs_crosswalk.tsv* file to identify whether an application is ultimately granted and PatentsView Detailed Description Text for word2vec processing. Because comprehensive application data begin in 2001, the training sample consists of applications filed during 2001–2003. The implemented hyperparameters are subsample $= 1$, column subsample $= 0.1$, minimum child weight $= 15$, maximum depth $= 5$, learning rate $= 0.01$, 100 estimators, and 10 early-stopping rounds. We then apply the fitted model retrospectively to granted patents from 1976 through 2023 to obtain Probability(XGBoost).

The second model uses text-embedding-ada-002 to map patent claims into 1,536-dimensional vectors. Combining these LLM embeddings with a neural classifier follows the predictive-patentomics approach of Yang (2023). Claims exceeding the model's 8,191-token input limit are first summarized

with GPT-5 under a fixed instruction that preserves substantive claim information while reducing length, consistent with the large-scale text-processing motivation in Breitung and Müller (2025). The resulting text is then embedded.

The prediction model uses: (i) the 1,536-dimensional embedding vectors; (ii) patent technological classifications based on CPC codes (Sections A–H and Y, along with their corresponding Classes), transformed into numerical sequences; and (iii) firm-level market capitalization in the patent year. Firm information is linked through GVKEY. Because firm size enters the score and is also related to returns, the return regressions control separately for lagged market capitalization. The resulting LLM probability combines textual, technological-classification, and firm-size information.

The LLM specification uses the same 2001–2003 application training period. The neural network has a 200–30–3–1 architecture, a 20% dropout rate, 100 training epochs, and a batch size of 500, matching the grant-acceptance architecture documented by Yang (2023). Its output, Probability(LLM), provides a second predicted grant-probability ranking based on a different text representation and model architecture. Comparing both rankings with the three realized benchmarks allows us to distinguish an architecture-specific error from a common failure to order frontier inventions.

2.3. Sample Summary Statistics and Cross-Measure Correlations

Table 1 describes 2,750,898 patents with an XGBoost score and 2,705,706 with an LLM score; the return tests use the subset linked to public firms with the required outcomes and controls. The two scores have similar means but markedly different dispersion. Probability(XGBoost) has a mean of 0.630 and a standard deviation of 0.113. Probability(LLM) has a mean of 0.626, a standard deviation of 0.201, and a substantially wider 10th-to-90th-percentile range. Evidence that appears under both scores therefore survives not only a different text representation and model architecture but also a materially different score distribution.

The benchmark distributions illustrate why average alignment is not the economically relevant test. KPSS has a median of 3.545 but a maximum above 6,000; total citations have a median of four but a maximum above 5,000; and five-year citations have a median of zero and a 90th percentile of four. Patent value and technological influence are therefore concentrated far beyond the typical observation. q_rank is less dominated by raw outliers because it is ranked within CPC-year cells, but it remains an ex post measure distinct from both grant-window market value and citations.

Panel B provides an initial indication that the machine rankings and realized benchmarks capture different information. The two AI scores correlate only 0.288 with one another, and each has small unconditional correlations with the three benchmarks. Most notably, the correlations of q_rank with

Probability(XGBoost) and Probability(LLM) are -0.080 and -0.189. These averages establish weak overall alignment, but they cannot distinguish mean-zero model noise from the systematic upper-tail reversal implied by technological underrecognition. Figures 1 and 2 therefore trace the relation across the benchmark distributions before the subsequent return tests examine whether the information missed by the models is also missed by capital markets.

[Insert Table 1 here]

III. Patent Benchmarks, AI Scores, and Subsequent Firm Outcomes

The summary evidence in Table 1 establishes weak unconditional alignment between the grant-probability scores and the realized patent benchmarks. This section asks whether the same score-benchmark divergence appears in subsequent firm outcomes before we turn to its distributional and capital-market dimensions. We examine two margins. Changes in operating profitability capture one possible path through internal commercialization, while acquisition-target incidence captures the transfer of innovative assets to firms with complementary technologies, capital, distribution, or organizational capabilities. Prior research connects innovative output to subsequent profitability, investment, returns, and corporate-control activity (Cohen et al., 2013; Hirshleifer et al., 2013, 2018; Bena and Garlappi, 2020; Phillips and Zhdanov, 2013; Bena and Li, 2014; Celik et al., 2022).

The comparison remains a measurement test rather than a causal channel test. If a grant-oriented score preserves the economic ordering of important inventions, higher firm-year mean scores should exhibit a stable positive relation with later firm outcomes and should absorb at least part of the corresponding benchmark association. Under technological underrecognition, upper-tail benchmark content may instead retain outcome-relevant information while the score supplies no stable positive ordering. Because five-year citations and q_rank are realized after grant, those benchmarks classify patent importance retrospectively; they do not create grant-date forecasts. The regressions therefore document conditional cross-sectional associations and do not identify effects of an AI score, a patent, or patent quality on firm outcomes.

3.1. Firm-Year Research Design and Outcome Construction

Because accounting and acquisition outcomes are observed annually, we translate the firm-grant-date return design to the PERMCO–grant-year level. For each firm-year, $AI_{j,t}$ is the raw mean XGBoost or LLM grant probability across all patents granted to the firm during year t. $HighQuality^{m}_{j,t}$ equals one

when at least one patent is strictly above the relevant 90th-percentile cutoff: 27.411 for KPSS, 5 citations received within five years after grant, or 0.267 for q_rank. The five-year-citation cutoff is calculated among the 1976–2019 grant cohorts with complete five-year exposure; because the rule is strict and citations are integer-valued, the indicator requires at least six citations. Raw patent count enters separately because firms with larger patent portfolios are mechanically more likely to contain an upper-tail patent. The regressions also include the eight lagged controls used in the return tests: log market capitalization, profitability, log book-to-market, book leverage, R&D intensity, capital expenditures, operating leverage, and cash.

The common specification is

$$Y_{j,t+h} = \alpha + \beta AI_{j,t} + \gamma HighQuality^{m}_{j,t} + \theta PatentCount_{j,t} + X'_{j,t-1}\delta + \lambda_t + \varepsilon_{j,t+h}, \tag{2}$$

where $Y_{j,t+h}$ is either a future annual change in ROA or a cumulative acquisition indicator, m denotes KPSS, five-year citations, or q_rank, and λ_t denotes grant-year fixed effects. Standard errors are clustered by PERMCO. Each score-benchmark pair uses its maximum available sample, so the five-year-citation sample ends in 2019, the q_rank sample reflects its own forward-information coverage, and the XGBoost and LLM samples can differ slightly. Grant-year fixed effects absorb common cohort conditions, including aggregate profitability and merger waves, while identification comes from differences among firms in the same grant year.

The design has three important limits. First, it includes neither firm fixed effects nor score-by-benchmark interactions, so the coefficients are not within-firm effects and do not test whether the score moderates the benchmark association. Second, the mean score averages all patents in a firm-year, whereas the benchmark indicator records whether any patent crosses an upper-tail threshold. Patent count controls for the most direct portfolio-size difference but not every nonlinear or compositional difference. Third, the tables report unadjusted two-sided significance levels across multiple scores, benchmarks, and horizons. Isolated marginal significance should therefore be interpreted cautiously; emphasis belongs on the direction, timing, and consistency of the score-benchmark contrast.

3.2. Future Changes in Operating Performance

Profitability is persistent but mean reverting, which makes future changes difficult to forecast (Fama and French, 2000). Innovation can nevertheless alter operating trajectories when inventions create products, lower costs, or support new investment opportunities (Hirshleifer et al., 2018). Table 2 therefore uses the annual change $\Delta ROA_{j,t+h} = ROA_{j,t+h} - ROA_{j,t+h-1}$ for $h = 1,2,3$. Each outcome is winsorized at the 1st and 99th percentiles within grant year. The columns represent three distinct annual changes, not

a cumulative three-year change. Because current ROA changes and their interaction with the score are not included, the specification also does not test whether AI-measured innovation attenuates profitability mean reversion.

[Insert Table 2 here]

The AI-score results are uniformly weak. None of the eighteen Mean Probability coefficients is statistically distinguishable from zero, and the signs do not form a stable positive sequence for either model. Several point estimates are negative, but the table does not establish a negative score-ROA relation. The narrower conclusion is that, conditional on the benchmark indicator, patent count, controls, and grant-year fixed effects, neither grant-probability score positively orders the next three annual changes in operating profitability.

The benchmark indicators display selective positive associations. A firm-year containing a top-decile KPSS patent is followed by a second-year ROA change that is 0.0081 under the XGBoost control and 0.0079 under the LLM control, approximately 0.8 percentage points in both cases. Five-year citations are positive in the first two annual changes: the XGBoost estimates of 0.0042 and 0.0043 are marginally significant, while the LLM estimate of 0.0047 is significant in the second year and the first-year estimate of 0.0036 is imprecise. The q_rank relation is weaker still; only the third-year XGBoost estimate is marginally significant. Thus, the ROA evidence is concentrated in particular benchmarks and horizons rather than uniform across the table.

This timing is economically plausible but not identified as a mechanism. KPSS records private value revealed around grant (Kogan et al., 2017); citations record later technological use (Trajtenberg, 1990; Hall et al., 2000); and q_rank records whether subsequent patent text moves toward the focal invention in the spirit of text-based innovation measures (Kelly et al., 2021). Five-year citations are realized over years $t+1$ through $t+5$ and therefore partly postdate every ROA outcome examined here. Citation flows also vary with field and cohort (Hall et al., 2001). Restricting the benchmark to completed 1976–2019 windows equalizes exposure, but it does not convert citations into information available at grant. Table 2 therefore provides retrospective evidence that some forms of realized patent importance covary with later operating trajectories; it does not establish a real-time forecasting strategy or a causal effect on profitability.

3.3. Subsequent Acquisition Outcomes

Acquisition provides a second outcome margin because innovative assets can be transferred to firms with complementary technologies, capital, distribution, or organizational capabilities. Takeover activity

can shape incentives to invest in R&D (Phillips and Zhdanov, 2013), and technological overlap and innovation synergies are central to innovative mergers (Bena and Li, 2014). At the same time, acquisition is an equilibrium match rather than a direct measure of standalone patent value, and information frictions can make its relation with innovation quality nonmonotonic (Celik et al., 2022).

Table 3 reports linear probability models in which the dependent variable equals one when the patenting firm is strictly matched as a completed Orbis target within one, two, or three years after the grant year. The outcomes are cumulative: when the longer window is observed, a firm acquired during the first year also counts as acquired within two and three years. The specification retains the firm-year score, benchmark indicator, patent count, lagged controls, grant-year fixed effects, and firm clustering used for ROA. Across benchmark samples, the one-year window contains 743 to 869 acquisitions and has an event rate of 3.23%–3.30%; the three-year window contains 2,036 to 2,241 acquisitions and has an event rate of 8.93%–9.04%.

[Insert Table 3 here]

The score pattern is clearer than in the ROA table. Every Mean Probability coefficient is negative and conventionally significant. For the LLM score, the one-year coefficients are -0.0460, -0.0557, and -0.0548 in the KPSS, five-year-citation, and q_rank samples. A 0.10 increase in the LLM score is therefore associated with a 0.46–0.56 percentage-point lower probability of acquisition within one year, compared with event rates of approximately 3.2%–3.3%. At two and three years, both score families retain negative coefficients in every benchmark sample. Coefficient magnitudes should not be compared directly across XGBoost and LLM because the score distributions have different dispersion. The common evidence is directional: grant-probability ranks do not positively order later acquisition-target incidence.

The benchmark coefficients move in the opposite direction with important differences across measures. KPSS is statistically positive in the two- and three-year windows, reaching 0.0193 under XGBoost and 0.0209 under LLM at three years. Five-year citations are positive and significant at every horizon, rising from 0.0083 and 0.0110 within one year to 0.0151 and 0.0203 within three years. In contrast, q_rank is generally imprecise; only the one-year LLM coefficient is statistically detectable. The acquisition evidence is therefore strongest for grant-window value and subsequent technological diffusion, not for every upper-tail benchmark. This pattern is consistent with buyers valuing technology, complementary assets, or growth options (Phillips and Zhdanov, 2013; Bena and Li, 2014), but the regressions do not identify which motive produces the match.

The acquisition outcome should be read narrowly. The table measures whether the patenting firm becomes a matched target, not whether it becomes an acquirer or whether a particular patent

motivates the transaction. A takeover depends on target willingness, acquirer capabilities, technological overlap, valuation, bargaining, financing, and the option to commercialize internally. The LLM score also combines patent text, technological classifications, and firm size, so its coefficient is not a pure textual assessment. Lagged market capitalization and the other controls absorb observed differences, but the absence of firm fixed effects leaves room for persistent firm heterogeneity. Strict Orbis matching reduces false target links at the cost of omitting transactions that cannot be matched with high confidence. The negative score coefficients and positive KPSS or citation coefficients are therefore descriptive sorting relations, not evidence that AI scores cause or prevent acquisitions.

3.4. Comparison of Market, Operating, and Acquisition Outcomes

Taken together, Tables 2 and 3 extend the score-benchmark comparison from patent-level correlations to subsequent firm outcomes, but with different strength across outcomes. The ROA evidence is selective: KPSS is associated with a larger second-year improvement, five-year citations with first- or second-year improvements, and q_rank only weakly with the third-year change. The acquisition evidence is more systematic for KPSS and five-year citations, while q_rank remains weak. Across both outcomes, the grant-probability scores provide no stable positive ordering: they are insignificant for annual ROA changes and negative for acquisition-target incidence.

This asymmetry is consistent with limited transportability from the grant-prediction task to later firm outcomes (Yang, 2023; Zheng, 2025). Features that help classify examination outcomes need not be the features associated with commercialization inside the firm or matching in the acquisition market. The result does not require every benchmark to predict every outcome at every horizon. KPSS, five-year citations, and q_rank capture distinct information sets and stages of innovation, and only KPSS is measured around grant. The narrower common lesson is that some realized measures of upper-tail patent importance contain subsequent firm-outcome information that the mean grant score does not absorb.

The corporate-outcome results provide complementary evidence for the distributional and post-grant return tests that follow. Stock prices can respond to distant expected cash flows before current ROA improves, innovative firms can invest aggressively while reporting low current profitability, and acquisitions depend on strategic complementarity and bargaining as well as standalone value. Combined with firm-year aggregation, post-grant benchmark timing, multiple comparisons, and the absence of firm fixed effects, these features rule out a causal or real-time forecasting interpretation. What the section establishes is a conditional measurement divergence: under the same control structure, upper-tail KPSS and citation classifications are associated with selected subsequent corporate outcomes, whereas the AI grant-probability scores are not positively associated with either outcome family.

IV. Distributional Relations Between Patent Benchmarks and AI Scores

Figure 1 in the Introduction establishes the paper's full-distribution fact: both grant-probability scores rise over parts of the ordinary patent distribution but lose alignment with all three benchmarks in the economically important right tail. This section sharpens that fact along two margins. Figure 2 asks whether the decline continues among patents already located in the top decile of a benchmark. Figure 3 replaces total citations with fixed post-grant citation windows, reducing mechanical differences in citation exposure across grant cohorts. These are descriptive ranking tests. They establish where the machine and benchmark orderings diverge; the subsequent section examines whether the information omitted from the machine rankings is also associated with post-grant returns.

Figure 2 focuses directly on ordinal performance within the right tail. If the scores preserve information about economically important inventions, a patent with a still higher benchmark value should not receive a systematically lower grant probability merely because it lies farther into the top decile. Each panel therefore restricts the available benchmark-score sample to its own top decile and then traces the score relation within that subsample. These panel-specific samples accommodate differences in benchmark and score coverage. The return regressions in Section V use indicators constructed from the corresponding full patent samples rather than from the figure bins.

[Insert Figure 2 here]

The fitted relation slopes downward in all six panels of Figure 2. Within the KPSS top decile, both Probability(XGBoost) and Probability(LLM) decline as grant-window market value increases. Within the citation top decile, both scores decline as later patents rely more heavily on the focal invention. Within the q_rank top decile, they decline as patents become more novel relative to prior work and more influential for the direction of subsequent patent text. The result is therefore stronger than a plateau at high benchmark values: conditional on entering the top decile, moving farther into the tail is associated with a progressively lower machine-predicted grant probability.

The q_rank comparison is particularly informative, but its interpretation requires care. q_rank is constructed entirely from patent-text embeddings, whereas the two grant scores combine claim-text representations with application characteristics and, for Probability(LLM), technological classifications and firm size. The negative relation therefore cannot be dismissed as arising only because a text-based model is compared with a nontextual outcome. At the same time, it does not imply that the models fail to process patent text generally. Rather, information extracted for a grant-oriented prediction does not preserve the ordering of textual innovations that later redirect technological search, the long-run object

emphasized by Kelly et al. (2021). The same upper-tail decline across two model architectures and three economically distinct benchmarks is thus the paper's principal patent-level evidence of technological underrecognition. Figure 3 next tests whether the citation component of this pattern is mechanically driven by unequal exposure to future citations.

[Insert Figure 3 here]

The fixed-window design follows the citation-lag evidence in Hall et al. (2001). Raw citation totals combine technological influence with cohort age, changes in patenting and citation intensity, field-specific practices, and right truncation. Hall, Jaffe, and Trajtenberg show that forward-citation flows peak at roughly three to five years for most technology fields. The five-year window therefore captures early technological diffusion while retaining relatively recent grant cohorts; the ten- and twenty-year windows trace progressively longer diffusion among older cohorts with completed exposure periods.

The fixed-window evidence preserves the upper-tail result. In the five-year panels, both grant-probability scores decline across the higher citation bins. The ten- and twenty-year relations are more hump-shaped at low and intermediate citation counts, consistent with the scores retaining some ordering ability away from the frontier, but both turn sharply downward among the most highly cited patents. The reversal observed with total citations is therefore not explained solely by older patents having had more time to accumulate citations. We use five-year citations as the regression benchmark because this horizon spans the peak early-diffusion period while preserving broader cohort coverage.

Together, Figures 1–3 reveal a non-monotonic ranking: the grant scores retain information for ordinary patents but bend away from all three benchmarks in the upper tail. This establishes cross-sectional technological underrecognition, not economic value or mispricing.

V. Patent Grants and Post-Grant Returns

The preceding section establishes a patent-level ranking failure. This section asks whether the information contained in upper-tail patents is also associated with the subsequent returns of the firms that own them, after controlling for the machine-predicted grant score. Patent grants are public events, but translating technical claims into products, licensing opportunities, competitive threats, and follow-on innovation can take time, especially when commercialization depends on complementary assets or later technological diffusion. We therefore trace size-adjusted returns over progressively longer horizons after each grant.

The return design is retrospective and descriptive. KPSS is measured around grant, whereas citations

and q_rank use post-grant information and are not available to investors at the grant date. The latter benchmarks therefore classify patents by importance realized ex post; they do not define implementable real-time trading signals. A return difference that accumulates with the horizon is consistent with slow incorporation of innovation-related information, but it does not by itself identify a causal effect of a patent, an AI score, or AI adoption. The tables retain the label PEPD, while the event is a patent grant and the outcome is a post-grant cumulative abnormal return.

Following Bernard and Thomas (1989), daily abnormal returns are defined as:

$$AR_{jt} = R_{jt} - R_{pt}, \tag{3}$$

where AR_{jt} denotes the abnormal return for firm j on day t; R_{jt} is the raw return of firm j on day t; and R_{pt} is the equally weighted average return on day t of the benchmark portfolio to which firm j belongs at the beginning of the calendar year. For a patent grant date g, the PEPD measure over horizon h is the cumulative abnormal return after the grant:

$$PEPD_{j,g}^{(h)} = \sum_{s=1}^{h} AR_{j,g+s}, \qquad h \in \{60,300,600,900,1200\}. \tag{4}$$

The outcome sums daily abnormal returns rather than compounding buy-and-hold returns. At the beginning of each calendar year, firms are assigned to market-cap deciles, and R_{pt} is the equal-weighted daily return of the corresponding size portfolio from the Kenneth R. French Data Library. The horizons range from 60 to 1,200 trading days and are expressed in percentage points. Their profile traces whether the benchmark-return association is concentrated near grant or continues to accumulate over the longer intervals emphasized in the innovation-return literature (Cohen et al., 2013; Hirshleifer et al., 2013; Shu et al., 2022). Baseline regressions report 60, 300, 600, and 900 days; the annual sorts additionally report 1,200 days when a complete window is available.

The regression design translates the patent-level comparison to the firm-grant-date level at which stock returns are observed. A firm may receive several patents on the same day. We therefore average grant probabilities across that day's patent batch and define the benchmark indicator to equal one when at least one patent in the batch exceeds the corresponding full patent-sample 90th percentile. This pairing compares the batch's mean grant-oriented screening rank with the presence of upper-tail patent content without assigning the firm's single return path to any one patent. Patent count enters separately because larger batches are mechanically more likely to contain an upper-tail patent. Following Bena and Garlappi (2020), the controls include lagged market capitalization, profitability, book-to-market, leverage, R&D intensity, capital expenditures, operating leverage, and cash. Grant-year fixed effects absorb common

annual variation, and standard errors are clustered by firm.

The baseline retains all eligible firm-grant dates. Long return windows can encompass later grants by the same firm, so Appendix Tables A2–A4 repeat the analysis on horizon-specific samples that require more than h trading days between retained events. For KPSS, the grant-value window overlaps the return outcome on at most day $g+1$. The non-overlapping-event estimates and the continued association at longer horizons show that the main pattern is not confined to this shared day. Because observations and cohort coverage decline with the horizon, coefficient growth across columns should be interpreted as a horizon profile rather than as a fixed-sample decomposition. Detailed control definitions appear in Appendix Table A1.

The baseline specification takes the following form:

$$PEPD_{j,g}^{(h)} = \alpha + \beta AI_{j,g} + \gamma Breakthrough_{j,g} + X'_{j,g-1}\delta + \lambda_{\tau(g)} + \varepsilon_{j,g}^{(h)}, \tag{5}$$

where $PEPD_{j,g}^{(h)}$ cumulates abnormal returns from trading day $g+1$ through $g+h$, $AI_{j,g}$ is the firm-date mean grant probability, $Breakthrough_{j,g}$ identifies a batch containing an upper-tail patent, $X_{j,g-1}$ contains lagged firm controls, and $\lambda_{\tau(g)}$ denotes grant-year fixed effects. The mean score summarizes the batch's position in a grant-oriented screening model; it is not treated as realized patent quality. The coefficient γ measures the conditional association between upper-tail benchmark content and subsequent returns, while β measures the conditional association between the mean grant score and those returns.

Under the information-processing view, a score that preserves the ordering of economically important inventions should have a stable positive β and should absorb at least part of the benchmark-return association. The return counterpart to technological underrecognition is instead a positive γ together with no stable positive β, interpreted jointly with the within-tail decline in Figure 2. A negative β indicates a negative conditional score-return association over the full firm-date sample; it is not, by itself, an upper-tail estimate. Coefficient magnitudes on the XGBoost and LLM scores are also not directly comparable because the two probability distributions have different dispersion.

5.1. Baseline Firm-Date Post-Grant Return Regressions

Table 4 reports the baseline XGBoost regressions over 60, 300, 600, and 900 trading days. Panels A–C enter upper-tail indicators for KPSS, total citations, and q_rank, respectively. Every indicator uses the corresponding full patent-sample 90th percentile. All specifications include patent count, lagged firm controls, grant-year fixed effects, and firm-clustered standard errors.

The benchmark-score contrast is sharpest for KPSS. A batch containing an upper-tail KPSS patent

is associated with 0.548 percentage points of abnormal return over 60 trading days, increasing to 1.872, 3.238, and 4.467 points over 300, 600, and 900 days. Although the estimation sample changes with the available horizon, the profile shows that the association is not confined to the short grant window. It also remains positive when retained events are separated by more than the outcome horizon in Appendix Table A2. Conditional on KPSS status and the firm controls, the XGBoost coefficient is negative and significant at 300, 600, and 900 days. Upper-tail grant-window value therefore retains a positive return association that the batch's mean XGBoost probability does not absorb.

Total citations and q_rank produce the same qualitative separation with different timing. The upper-tail total-citation indicator is positive at every horizon, while the XGBoost coefficient is negative in all four columns and significant at 300 and 900 days. Because citations are realized after grant, these estimates are retrospective: patents that subsequently become widely used are associated with higher returns even after the grant score is controlled for. The q_rank indicator is positive at 300 and 600 days, whereas the XGBoost coefficient is statistically indistinguishable from zero throughout that panel. The text-direction specifications therefore show an absence of positive incremental score information, not a statistically significant negative score-return relation. Across the three panels, the benchmark coefficients are generally stronger at medium and long horizons, consistent with gradual incorporation of innovation-related information (Cohen et al., 2013; Hirshleifer et al., 2013; Shu et al., 2022).

[Insert Table 4 here]

Table 5 repeats the three comparisons using Probability(LLM). This replication changes the text representation and model architecture, adds technological classifications and firm size, and produces a substantially more dispersed score. Despite those differences, the benchmark-score separation closely matches the XGBoost results. The upper-tail KPSS coefficient is positive and significant at every horizon, increasing from 0.577 percentage points over 60 days to 4.980 points over 900 days. The mean LLM-score coefficient is negative at every horizon and significant through 600 days. Total citations show the same contrast: the benchmark coefficient is positive at all four horizons, while the score coefficient is significantly negative through 600 days.

Panel C again locates the positive return association in the benchmark rather than in the grant probability. q_rank is positive at 300 and 600 days, while the LLM-score coefficient is statistically indistinguishable from zero throughout. As in the XGBoost panel, this is evidence of no positive incremental score association, not a significant negative reversal. Agreement across the two model constructions matters because it shows that the conditional divergence is not specific to one text representation, architecture, or score distribution. It does not imply that the two score coefficients have comparable economic

magnitudes, because their dispersions differ materially.

[Insert Table 5 here]

Table 6 replaces total citations with citations received during the first five years after grant. The common window reduces cohort-age differences and covers the interval in which forward-citation flows typically peak (Hall et al., 2001). For both scores, the upper-tail five-year-citation indicator is positive and significant at 60, 300, and 600 trading days. Its coefficient increases from 0.338 to 2.346 percentage points between 60 and 600 days in the XGBoost specifications and from 0.317 to 2.226 points in the LLM specifications, before becoming smaller and statistically imprecise at 900 days. Patents that experience unusually high early technological use therefore retain a positive subsequent-return association after the firm-date mean grant score is controlled for.

The score point estimates move in the opposite direction. Probability(XGBoost) is negative in all four columns and significant at 300 and 900 days; Probability(LLM) is negative throughout and significant through 600 days. The fixed-window design therefore reproduces the conditional divergence without relying on unequal lifetime citation exposure. Because the citation indicator is known only after five years, the result validates the economic content of the retrospective benchmark rather than a real-time citation strategy.

[Insert Table 6 here]

5.2. Summary of the Baseline Return Evidence

The baseline results connect the patent-level ranking failure to three distinct return patterns. Upper-tail KPSS status is positively associated with returns from 60 through 900 days; total and five-year citations are positive at most or all reported horizons; and q_rank is positive mainly at 300 and 600 days. Each benchmark coefficient is estimated conditional on the batch's mean grant probability. The score therefore does not subsume the return association of patents that are valuable around grant, widely used afterward, or direction-changing in subsequent text.

The Mean Probability coefficients provide the complementary comparison. If a grant-oriented score preserved the economic ordering of upper-tail inventions, its conditional coefficient should be stably positive. Instead, score coefficients are frequently negative in the KPSS and citation specifications and statistically weak throughout the q_rank specifications. The evidence is therefore asymmetric: realized benchmark importance retains positive return associations, while neither model supplies a stable positive incremental score-return relation.

The timing differs across benchmarks. KPSS has the strongest increasing horizon profile, citations remain informative under a common five-year exposure window, and q_rank is concentrated at intermediate horizons. These differences are consistent with the benchmarks capturing distinct stages of innovation, but the regressions do not identify the mechanism behind each horizon profile. Non-overlapping-event estimates in Appendix Tables A2–A4 preserve the principal KPSS and fixed-window citation results. Overall, the baseline establishes conditional return divergence, not that the AI scores caused the return paths or that the ex post benchmarks were tradable at grant.

5.3. Joint-Benchmark Return Specifications

The three benchmarks can overlap. A direction-changing invention may attract later citations, and either dimension of technological importance may also be reflected in KPSS grant-window value. Table 7 therefore enters two upper-tail indicators jointly. Panel A uses Probability(XGBoost), and Panel B uses Probability(LLM). The first block combines KPSS with q_rank, while the second combines five-year citations with q_rank. These specifications ask which benchmark associations survive conditional on another measure of realized importance; they do not isolate independent causal effects. All columns retain the baseline controls, fixed effects, and clustering.

The specification for both panels of Table 7 is:

$$PEPD_{j,g}^{(h)} = \alpha + \beta AI_{j,g} + \gamma_1 Breakthrough_{j,g}^{m} + \gamma_2 Breakthrough_{j,g}^{\text{q_rank}} + X'_{j,g-1}\delta + \lambda_{\tau(g)} + \varepsilon_{j,g}^{(h)}. \quad (6)$$

Here $Breakthrough_{j,g}^{m}$ denotes either the KPSS or five-year-citation indicator, and $Breakthrough_{j,g}^{\text{q_rank}}$ denotes the q_rank indicator. In Panel A, KPSS remains positive at 60, 300, and 600 days, with coefficients of 0.455, 1.380, and 2.470; q_rank remains positive at 300 and 600 days, with coefficients of 0.970 and 1.898. When five-year citations replace KPSS, the citation coefficients remain positive at all horizons—0.446, 1.756, 2.872, and 1.986—while q_rank becomes statistically imprecise. The attenuation is consistent with overlap between technological use and the forward-direction component of q_rank, whereas KPSS captures value revealed around grant.

Panel B produces the same allocation of benchmark associations with Probability(LLM). KPSS remains positive through 600 days when q_rank is included, and q_rank remains positive at 300 and 600 days. With five-year citations, the citation indicator remains positive at every horizon while q_rank again attenuates. Neither AI score has a stable positive coefficient after the benchmark pairs enter. Most score point estimates are negative at 300 and 600 days, but they are statistically imprecise in these joint

specifications. The joint tests therefore show that the return divergence is not driven by one benchmark alone: distinct or overlapping dimensions of realized patent importance survive together, whereas the grant-score ordering adds no detectable positive return association.

[Insert Table 7 here]

5.4. Annual Low-AI Sorts and Upper-Tail Return Differences

Figure 5 provides a nonparametric view of the score-benchmark mismatch. Within each issue year beginning in 2010, low AI is the bottom 30% of the average percentile rank of Probability(XGBoost) and Probability(LLM). Within this low-score group, we compare patents above and below the same-year 90th percentile of q_rank, KPSS, or five-year citations. The plotted outcome is the difference in mean post-grant cumulative abnormal returns; the light line reports the annual difference and the dark line its three-year moving average. Annual thresholds absorb broad shifts in the score and benchmark distributions and therefore differ from the full-sample cutoffs used in the regressions.

The moving averages are positive in most medium- and long-horizon panels. Thus, even among patents assigned low grant-probability ranks, upper-tail benchmark status separates the subsequent return paths of the associated firms, with many differences accumulating over several hundred trading days. This horizon profile is consistent with the delayed valuation of innovative output documented by Cohen et al. (2013), Hirshleifer et al. (2013), and Shu et al. (2022). The timing is not uniform, however. The q_rank, KPSS, and five-year-citation series turn at different points, and the available issue years differ across return horizons because each horizon requires a complete subsequent return window.

Benchmark-specific timing is not surprising. KPSS records value revealed around grant, citations record later technological use, and q_rank records a later change in technological direction. The displayed AI release dates are visual markers only; they also coincide with changes in discount rates, technology valuations, patent cohorts, data coverage, and the post-pandemic allocation of capital. The figure therefore documents calendar-time variation in retrospective sorts, not an event-study effect attributable to a particular model release.

[Insert Figure 5 here]

Figure 5 complements the controlled firm-date regressions with transparent within-year comparisons. The next subsection formalizes one calendar split around the paper's GPT-3 cutoff, while retaining the same descriptive interpretation. Shorter horizons reach the periods around the most recent release markers; longer horizons necessarily end in earlier issue years.

5.5. Descriptive Calendar-Time Heterogeneity Around the GPT-3 Cutoff

We finally ask whether the conditional return associations differ across two calendar periods. The paper uses June 11, 2020 as its GPT-3 cutoff and interacts the grant score and benchmark indicator with a later-period dummy. KPSS, five-year citations, and q_rank enter in separate column groups, and the outcomes cover 60, 300, and 600 trading days. The cutoff is a descriptive calendar marker, not a measure of investor or examiner adoption, an exogenous shock, or a treatment date. The exercise asks only whether the benchmark-score contrast is different in the later data.

The estimating equation is:

$$\begin{aligned} PEPD_{j,g}^{(h)} = \alpha + \beta AI_{j,g} + \gamma Breakthrough_{j,g}^{m} + \theta Post_g \\ + \phi(AI_{j,g} \times Post_g) + \psi(Breakthrough_{j,g}^{m} \times Post_g) + X_{j,g-1}'\delta + \lambda_{\tau(g)} + \varepsilon_{j,g}^{(h)}. \end{aligned} \tag{7}$$

Here $Post_g$ equals one after the cutoff, so $\beta + \phi$ is the later-period grant-score slope and $\gamma + \psi$ is the later-period benchmark slope. Panel A of Table 8 shows that every XGBoost interaction is negative. At 300 and 600 days, the interactions are -15.797 and -31.676 with KPSS, -15.587 and -31.404 with five-year citations, and -12.543 and -27.104 with q_rank. In the KPSS columns, these estimates imply later-period grant-score slopes of -16.456 and -30.935. Benchmark interactions generally move in the opposite direction. KPSS interactions of 5.564 and 9.983 imply later-period slopes of 5.845 and 10.418; the corresponding later-period slopes are 2.980 and 4.346 for five-year citations and 2.961 and 5.831 for q_rank. Not every benchmark interaction is statistically significant, particularly for citations at 600 days and q_rank at 60 days.

Panel B yields the same directional split with Probability(LLM). At 300 and 600 days, the score interactions are -7.710 and -13.714 with KPSS, -8.692 and -15.555 with citations, and -6.753 and -11.054 with q_rank. KPSS benchmark interactions are positive at all three horizons, while q_rank interactions are positive at 300 and 600 days. The citation interaction is statistically detectable at 60 days but imprecise at the longer horizons. Thus, within each model's own scale, later-period score slopes become more negative while several benchmark slopes become more positive. The magnitudes should not be compared directly across XGBoost and LLM because their score distributions differ.

Figure 4 provides an exploratory candidate-tail diagnostic. The earlier-period benchmark tail is fixed at the annual upper 20% of each patent benchmark. For the later period, the exercise considers upper-tail thresholds of 7.5%, 10%, and 12.5% and, separately for each benchmark, retains the candidate producing the largest absolute implied later-period PEPD600 coefficient. The selected tails are 7.5% for

KPSS and 12.5% for q_rank and five-year citations, averaging 10.8%. The 37.5%–62.5% contraction in tail width is therefore imposed by the candidate design rather than estimated from the data.

Within those outcome-selected post-period tails, the absolute implied PEPD600 coefficient rises from 0.834 to 5.886 for KPSS, from 0.532 to 1.952 for q_rank, and from 2.559 to 3.032 for five-year citations. The corresponding harder ratios are 605.5%, 266.6%, and 18.5%. Because the post threshold is selected in sample to maximize absolute severity within the candidate band, these values are descriptive and should not be interpreted as unbiased treatment estimates or formal evidence that a model release caused mispricing to intensify. The figure visualizes a pattern consistent with a narrower residual tail and a larger return association; the original fixed upper-20%/upper-10% placebo and separate-subsample bootstrap diagnostics remain in Appendix Figure A1.

[Insert Figure 4 here]

The later-period interactions and Figure 4 are consistent with the upper-tail divergence documented elsewhere, but the cutoff also separates market regimes, patent cohorts, and data environments. The information-acquisition mechanisms in Dugast and Foucault (2018) and Dessaint et al. (2024) offer possible interpretations, yet the design cannot distinguish them from other calendar-time changes. We therefore interpret the evidence as descriptive heterogeneity, not as an effect of GPT-3 release timing or AI adoption.

[Insert Table 8 here]

VI. Retrospective Portfolio Validation and Factor Exposures

The final main-text exercise aggregates the high-benchmark, low-score classification into firm-month return series. The purpose is validation, not strategy design: we ask whether firms that are retrospectively classified as holding upper-tail patents poorly ranked by the grant models exhibit distinctive historical returns and factor exposures. Factor adjustment is essential because innovative firms differ systematically in size, valuation, profitability, investment, and expected growth (Hirshleifer et al., 2013; Bena and Garlappi, 2020).

For each holding month, the historical classification uses patents granted during the prior 12 months. A firm is included when at least one patent lies in the same-issue-year top decile of q_rank, KPSS, or five-year citations and in the bottom 30% of the corresponding XGBoost or embedding-based grant score.

Portfolios are equal weighted, held for one month, and rebalanced monthly. None is an implementable real-time portfolio as constructed. Five-year citations and q_rank use information realized after grant; KPSS incorporates the grant-window stock-price response; and the same-issue-year thresholds are calculated with the completed annual distribution. The exercise therefore evaluates the historical return paths and common-factor exposures of the classification, not returns available to an investor using only contemporaneous information.

6.1. Full-Sample Retrospective Alphas and Factor Exposures

Table 9 reports time-series regressions over a nominal 2010–2024 window, retaining months in which the relevant portfolio contains at least five firms. Because benchmark coverage differs, the effective samples contain 119 months for q_rank, 179 for KPSS, and 131 for five-year citations. Panels A and B use the Fama and HXZ factor models to ask how much of each retrospective portfolio's average excess return is captured by common factor exposures:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p' F_t + \varepsilon_{p,t},$$

where $R_{p,t}$ is the equal-weighted portfolio return, $R_{f,t}$ is the risk-free rate, and F_t contains contemporaneous monthly factors. The intercept α_p is monthly. The reported annual alpha is the linear quantity $12\alpha_p$, not a compounded annual return, and it has the same t-statistic as the monthly estimate. Unlike the firm-date regressions, each column follows one portfolio return series through time and describes which common-factor realizations covary with its average return.

Panel A evaluates the portfolios against the Fama three- and five-factor models (Fama and French, 1993, 2015). Three-factor monthly alphas are positive in all six columns, ranging from 0.258% to 0.385%, and are conventionally significant at the 10% level or better. Five-factor alphas are also positive and significant in all columns, ranging from 0.369% to 0.528% per month. Similar signs across the XGBoost and embedding-based screens show that the historical classification is not unique to one score construction. Because the portfolios are formed with ex post information, however, these intercepts validate a sorting relation; they are not attainable strategy returns.

The Fama loadings reveal substantial characteristic imbalance. Market betas lie between 0.993 and 1.144. HML loadings are negative throughout and are most pronounced for KPSS, while q_rank and five-year-citation portfolios load positively on SMB and KPSS portfolios have size loadings near zero or negative. In the five-factor regressions, RMW and CMA loadings are negative in every column, although several q_rank and citation estimates are imprecise. Adding those factors raises rather than reduces alpha because low profitability and aggressive investment lower the return predicted by the Fama five-factor

model. These patterns are consistent with innovative firms financing current investment in anticipation of future opportunities (Cohen et al., 2013; Hirshleifer et al., 2013; Shu et al., 2022), but the loadings themselves do not distinguish risk compensation from mispricing.

Panel B evaluates the same returns with HXZ q5, which combines market, size, investment, profitability, and expected-growth factors (Hou et al., 2015, 2021). Monthly intercepts fall to 0.043%–0.132%, and none is statistically significant. The corresponding linear annual quantities range from 0.52% to 1.58%. Thus, the full-sample positive Fama intercepts are not robust to this alternative factor benchmark. The table cannot determine whether HXZ captures risk, characteristics associated with expected returns, or the economic channel behind the classification; it establishes factor-model sensitivity.

[Insert Table 9 here]

The HXZ loadings clarify the empirical source of that sensitivity. IA loadings range from -0.160 to -0.300 and ROE loadings from -0.184 to -0.350; all are statistically significant. Expected-growth loadings are positive in every column but differ in precision: they are large and significant for five-year citations, marginal for KPSS, and statistically weak for q_rank. Size loadings are positive and significant for q_rank and citations but near zero for KPSS. The retrospective classification therefore selects firms with systematically different investment, profitability, growth, and size profiles (Bena and Garlappi, 2020).

Panels A and B establish a clear boundary. Standard Fama factors leave positive intercepts for the high-benchmark, low-score classifications, whereas HXZ q5 absorbs those full-sample intercepts statistically. This weakens any claim of factor-robust abnormal performance without changing the patent-level ranking divergence in Figures 1–3. The portfolio evidence supports economic relevance through its association with firms' return characteristics, but it does not by itself distinguish mispricing from compensation associated with investment, profitability, and expected growth.

6.2. Later-Period Descriptive Portfolio Evidence

Figure 6 plots cumulative factor-model residuals for a later calendar window beginning in September 2018. The date defines a descriptive subsample and is not an AI-release or adoption treatment. Rows correspond to q_rank, KPSS, and five-year citations; columns compare XGBoost and LLM classifications. Panel A uses residuals from the Fama three- and five-factor models, and Panel B uses HXZ q5. Effective histories differ sharply: q_rank contributes 16 holding months, KPSS 76, and five-year citations 28. The paths therefore illustrate when model residuals accumulate within each available sample, but their endpoints are not directly comparable across benchmarks.

The q_rank row shows the strongest factor-model sensitivity and the shortest history. The three-factor paths end near 1.3% for XGBoost and 2.0% for LLM after giving back earlier gains, whereas the five-factor and HXZ paths end near 10%–12%. With only 16 observations, these differences can be driven by a small number of months and should not be interpreted as persistent portfolio performance. They mainly demonstrate that the residual assigned to the same returns changes materially with the factor benchmark.

KPSS supplies the longest descriptive path. Over 76 months, the three-factor residuals end near 18.6% and 25.3%, and the five-factor residuals near 26.3% and 37.6%. Under HXZ q5, however, the XGBoost path rises and then reverses to 7.3%, while the LLM path ends at 15.2%. The repeated advances and drawdowns show that the accumulation is not a single grant-window jump. They also show that its endpoint is sensitive to investment-, profitability-, and expected-growth factors.

The five-year-citation histories contain 28 months. Their Fama residuals finish at approximately 8% for XGBoost and 13% for LLM, while both HXZ paths end near 12%. Positive residual accumulation appears under both score classifications, consistent with delayed incorporation when value becomes visible through technological diffusion (Cohen et al., 2013; Hirshleifer et al., 2013; Shu et al., 2022). The short and saw-tooth paths nevertheless make the magnitude sensitive to the selected calendar window. Across all rows, Figure 6 is best read as a visualization of factor-model and sample-window dependence rather than evidence of an ex ante trading opportunity.

[Insert Figure 6 here]

[Insert Table 10 here]

Table 10 summarizes the same later calendar window. Under the Fama three-factor model, only the LLM-KPSS intercept is marginally significant; the other five intercepts are imprecise. Under the five-factor model, all six intercepts are positive and conventionally significant, ranging from 0.280% to 0.717% per month. As in the full sample, negative profitability and investment loadings cause the five-factor model to assign lower expected returns and therefore leave larger positive intercepts.

HXZ q5 produces heterogeneous results. KPSS intercepts are 0.104% and 0.200% and statistically insignificant. Five-year-citation intercepts are 0.415% and 0.421% and significant for both score classifications. The LLM-q_rank intercept is 0.697% and significant, while the XGBoost estimate is not. These later-period estimates are selected from short and unequal samples and are reported across eighteen factor-model-by-portfolio combinations with unadjusted significance levels. They should therefore be treated as descriptive subsample evidence rather than proof of a new factor-robust anomaly.

Sample length determines how much weight the estimates can bear. KPSS has 76 valid months, compared with 28 for five-year citations and only 16 for q_rank; a six-lag Newey–West correction is especially demanding in the shortest series. The later window is therefore most informative about the time path and factor sensitivity of KPSS. For citations and q_rank, the positive HXZ intercepts are suggestive but vulnerable to endpoint choice, individual months, and multiple comparisons. Read together, the figure and table reinforce the section's main boundary: full-sample portfolio intercepts disappear under HXZ q5, while a few later-period retrospective classifications retain positive residuals in much shorter samples.

VII. Conclusion

This paper documents upper-tail underrecognition in two grant-prediction scores. Higher predicted grant probabilities should indicate stronger grant-relevant patent quality and, through that channel, greater economic value and technological influence. Yet both the XGBoost and embedding-based scores turn downward near the upper ends of KPSS market value, forward citations, and q_rank, and they continue to decline within the top decile of every benchmark. Thus, among patents with high ex post economic importance, greater market value, wider subsequent use, and greater directional change are associated with lower average predicted grant probabilities.

At firm grant dates, batches containing upper-tail KPSS or citation patents are followed by higher post-grant cumulative abnormal returns across multiple horizons, while the q_rank indicator is positive mainly over 300 and 600 days. These benchmark associations remain after controlling for the mean grant score. The score coefficient provides no stable positive incremental return information: it is negative in many KPSS and citation specifications and statistically insignificant in the q_rank specifications. Joint regressions show that grant-window value and breakthrough direction can retain separate return associations, while the overlap between citations and q_rank is consistent with a shared dimension of future technological diffusion. Annual low-AI sorts exhibit large turning points around selected AI release dates, and the post-2017 regressions show more negative score slopes alongside stronger benchmark slopes. These calendar splits document substantial temporal heterogeneity, but they do not identify its source.

The firm-year outcomes extend this distinction beyond returns. We find no evidence that either mean grant score positively predicts annual changes in ROA over the next three years. By contrast, several transformative-innovation indicators have economically sensible positive associations with future outcomes: KPSS is positive in the second-year ROA change, five-year citations are positive in the

first two annual ROA changes, and KPSS and five-year citations are positively associated with later acquisition incidence. The acquisition tests are especially informative because the AI scores move in the opposite direction: the LLM score is negatively associated with acquisition within one year across all three benchmarks, and both score families are negatively associated with cumulative acquisition incidence over two and three years. These estimates use the full lagged-control structure of the return regressions and grant-year fixed effects, but no firm fixed effects. Omitting grant-year fixed effects leaves all AI-coefficient signs and the AI multiplicity conclusions unchanged. The estimates therefore document conditional cross-sectional relations, not causal effects of AI or patent quality. The ex post timing of citations and q_rank, the target-only acquisition measure, and the equilibrium nature of acquisition further limit a real-time or policy interpretation.

The firm-month portfolios extend the same classification into asset pricing. Firms with high-benchmark, low-AI patents exhibit positive Fama three- and five-factor alphas. The full-sample alphas become small and insignificant under HXZ q5, whose investment, profitability, and expected-growth factors account for much of the portfolios' average return. This result limits the asset-pricing interpretation: the retrospective portfolios do not deliver abnormal performance relative to q5 in the full sample, even though the patent-level score-benchmark divergence remains.

The broader lesson is that success at predicting patent approval need not extend to ranking granted patents by subsequent economic importance. Although higher grant probabilities should identify stronger patent quality and stronger quality is associated with economic value, the grant-prediction rankings diverge from the ex post ordering of the small set of patents with the greatest market value, technological diffusion, or influence on technological direction. For financial institutions using such scores to prioritize innovation, this upper-tail divergence is the central underrecognition problem.

References

Bena, J. and Garlappi, L. (2020). Corporate innovation and returns. *The Review of Corporate Finance Studies*, 9(2):340–383.

Bena, J. and Li, K. (2014). Corporate innovations and mergers and acquisitions. *The Journal of Finance*, 69(5):1923–1960.

Bernard, V. L. and Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research*, 27:1–36.

Bond, P., Edmans, A., and Goldstein, I. (2012). The real effects of financial markets. *The Journal of*

Finance, 67(1):339–360.

Bonelli, M. (2026). Data-driven investors. *The Review of Financial Studies*, 39(7):1909–1969.

Breitung, C. and Müller, S. (2025). Global business networks. *Journal of Financial Economics*, 166:104007.

Brynjolfsson, E. and Mitchell, T. (2017). What can machine learning do? workforce implications. *Science*, 358(6370):1530–1534.

Cao, S., Jiang, W., Wang, J., and Yang, B. (2024). From man vs. machine to man + machine: The art and AI of stock analyses. *Journal of Financial Economics*, 160:103910.

Celik, M. A., Tian, X., and Wang, W. (2022). Acquiring innovation under information frictions. *The Review of Financial Studies*, 35(10):4474–4517.

Cohen, L., Diether, K. B., and Malloy, C. J. (2013). Misvaluing innovation. *The Review of Financial Studies*, 26(3):635–666.

Dessaint, O., Foucault, T., and Frésard, L. (2024). Does alternative data improve financial forecasting? the horizon effect. *The Journal of Finance*, 79(3):2237–2287.

Dugast, J. and Foucault, T. (2018). Data abundance and asset price informativeness. *Journal of Financial Economics*, 130(2):367–391.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2000). Forecasting profitability and earnings. *The journal of business*, 73(2):161–175.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132.

Golder, P. N., Shacham, R., and Mitra, D. (2009). Innovations' origins: When, by whom, and how are radical innovations developed? *Marketing Science*, 28(1):166–179.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2000). Market value and patent citations: A first look. National Bureau of Economic Research Working Paper No. 7741.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. Working Paper 8498, National Bureau of Economic Research.

Hirshleifer, D., Hsu, P.-H., and Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, 107(3):632–654.

Hirshleifer, D., Hsu, P.-H., and Li, D. (2018). Innovative originality, profitability, and stock returns. *The Review of Financial Studies*, 31(7):2553–2605.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance*, 25(1):1–41.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring technological innovation over the long run. *American Economic Review: Insights*, 3(3):303–320.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132(2):665–712.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, 66(5):1823–1860.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

O'Connor, G. C. and Rice, M. P. (2013). New market creation for breakthrough innovations: Enabling and constraining mechanisms. *Journal of Product Innovation Management*, 30(2):209–227.

Phillips, G. M. and Zhdanov, A. (2013). R&d and the incentives from merger and acquisition activity. *The Review of Financial Studies*, 26(1):34–78.

Shu, T., Tian, X., and Zhan, X. (2022). Patent quality, firm value, and investor underreaction: Evidence from patent examiner busyness. *Journal of Financial Economics*, 143(3):1043–1069.

Singh, J. and Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management Science*, 56(1):41–56.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The RAND Journal of Economics*, 21(2):172–187.

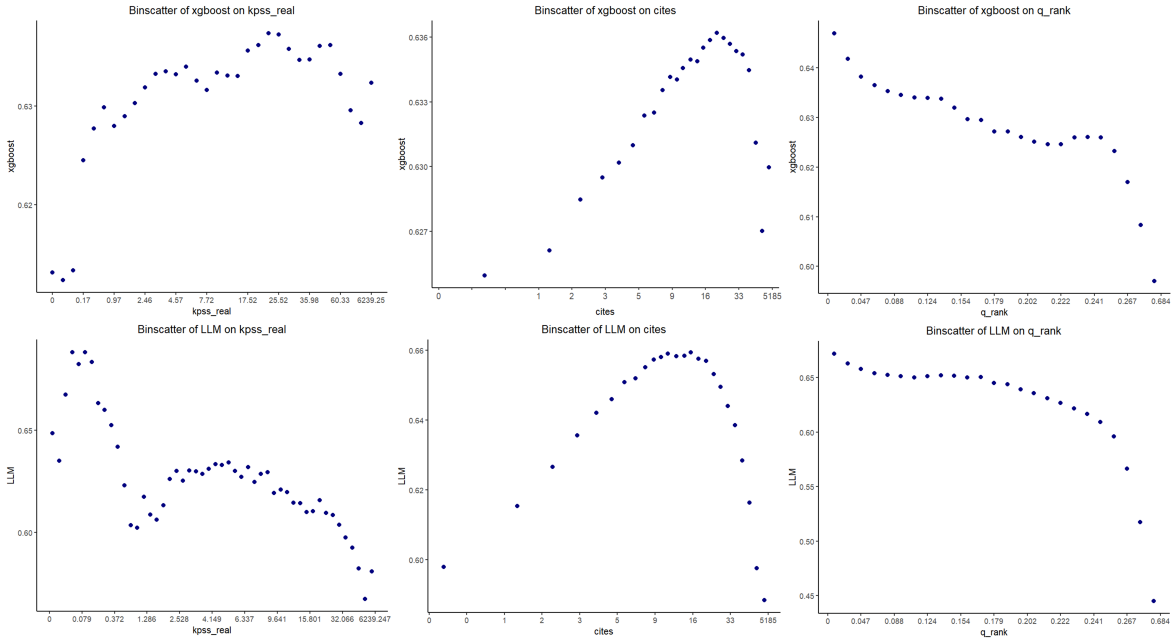
Van Raan, A. F. (2004). Sleeping beauties in science. *Scientometrics*, 59(3):467–472.

Wang, J., Veugelers, R., and Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436.

Yang, S. (2023). Understanding innovation quality and success with large-language models. *Available at SSRN 4482536*.

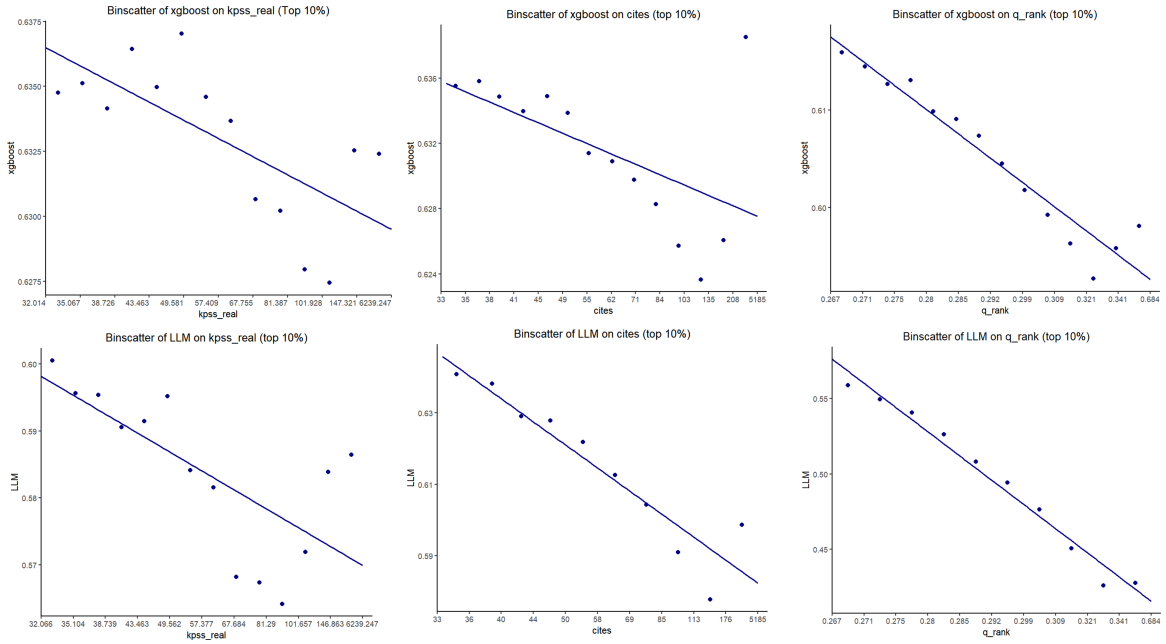
Zheng, X. (2025). How can innovation screening be improved? a machine learning analysis with economic consequences for firm performance. *Journal of Financial and Quantitative Analysis*, 60(8):3722–3752.

Figure 1. AI Grant-Probability Scores Across Patent Benchmarks



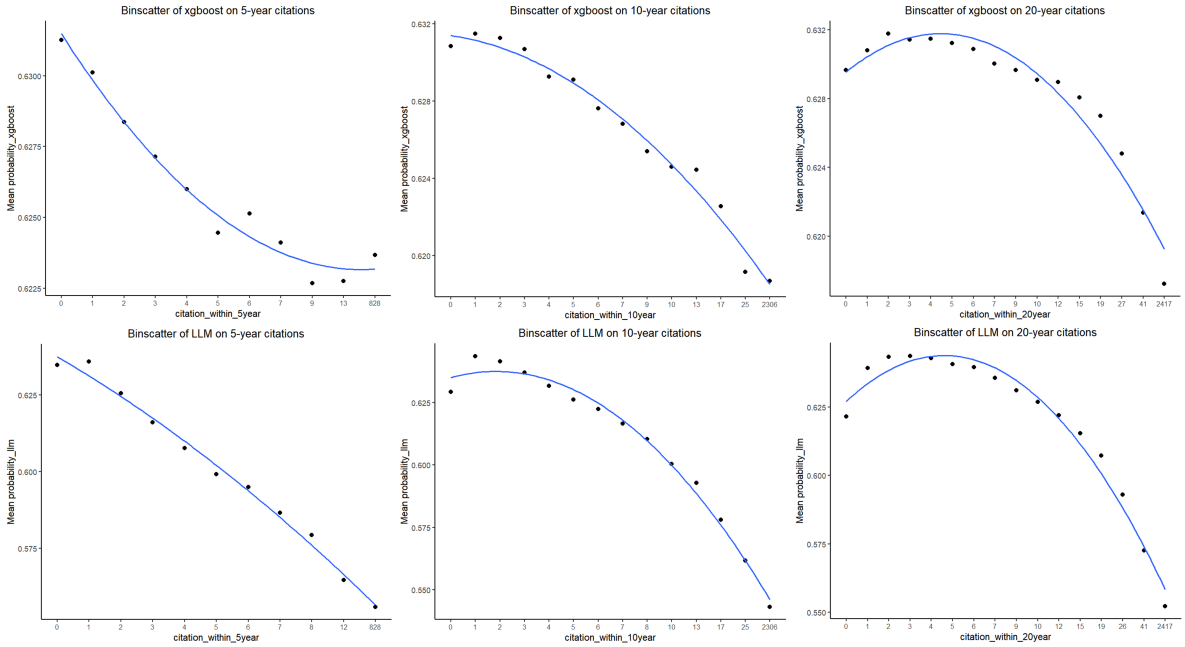
Notes: This figure reports binned means of model-implied grant probabilities across KPSS grant-window value, total forward citations, and the ex post text-based rank q_rank. Observations are ordered on a quantile-rank horizontal coordinate; selected tick labels report the corresponding raw benchmark values, so horizontal distances should not be read as equal raw-value increments. The upper row uses XGBoost probabilities and the lower row uses LLM-embedding-based probabilities. The plots are unconditional and descriptive.

Figure 2. AI Grant-Probability Scores Within Benchmark Top Deciles



Notes: Each panel restricts its available benchmark-score sample to patents at or above the 90th percentile of KPSS grant-window value, total forward citations, or q_rank. Points are binned means plotted on a within-subsample quantile-rank coordinate; selected tick labels show the corresponding raw benchmark values. The fitted lines are descriptive. Thresholds and effective samples can differ across panels because coverage differs by benchmark and AI score.

Figure 3. Grant-Probability Scores and Fixed-Window Forward Citations



Notes: This figure relates model-implied grant probabilities to citations received within five, ten, and twenty years after grant. Zero-citation patents form a separate group; positive citation counts are divided into rank-based bins and plotted as bin means. Horizontal positions indicate bin order, while selected tick labels report the corresponding citation counts. Only patents with the relevant completed citation window and nonmissing score are included, so samples differ across panels and longer windows end earlier. Citation counts are ex post benchmarks and are not available at grant. All relations are descriptive.

Figure 4. GPT-3 Cutoff and Metric-Specific Scope-versus-Severity

Panel A. Candidate tail widths in the earlier and later calendar periods

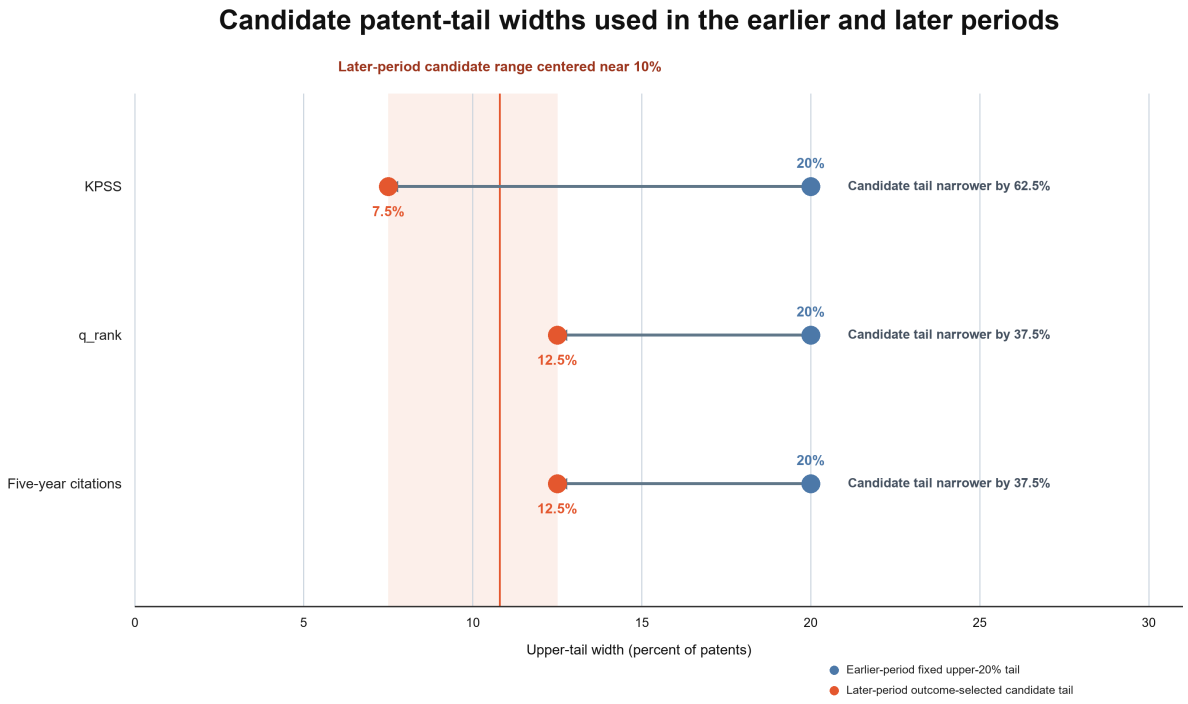


Figure 4. GPT-3 Cutoff and Metric-Specific Scope-versus-Severity (continued)

Panel B. Outcome-selected pooled-interaction coefficients by calendar period

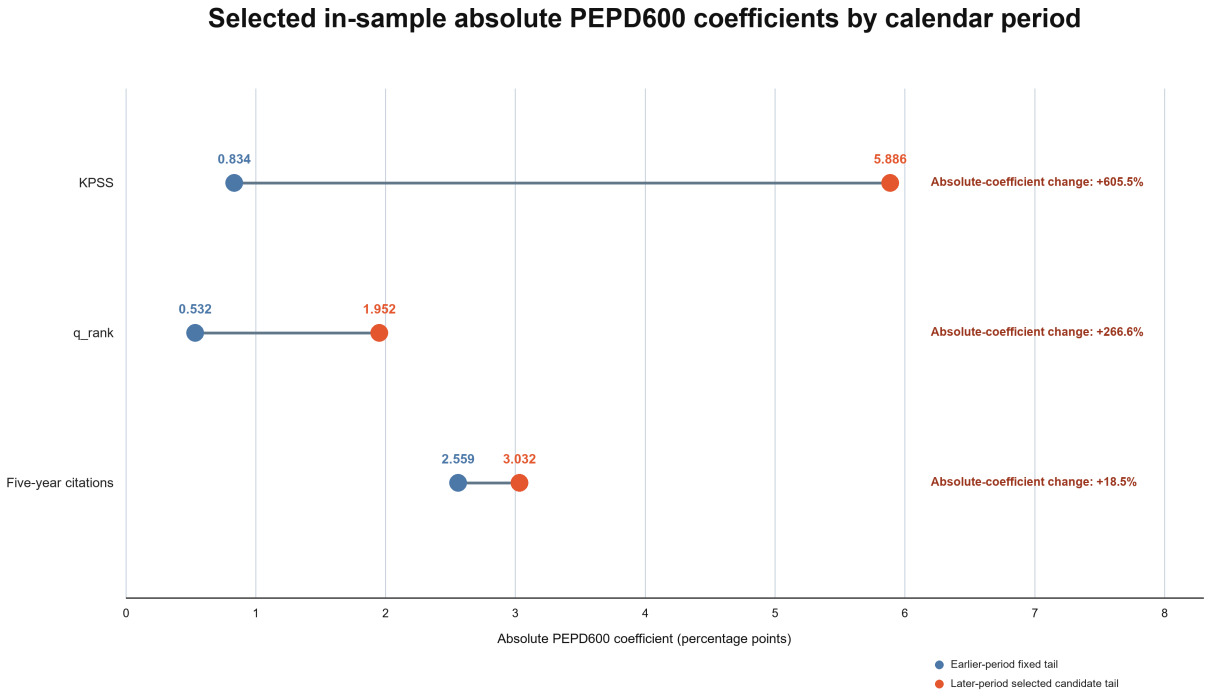
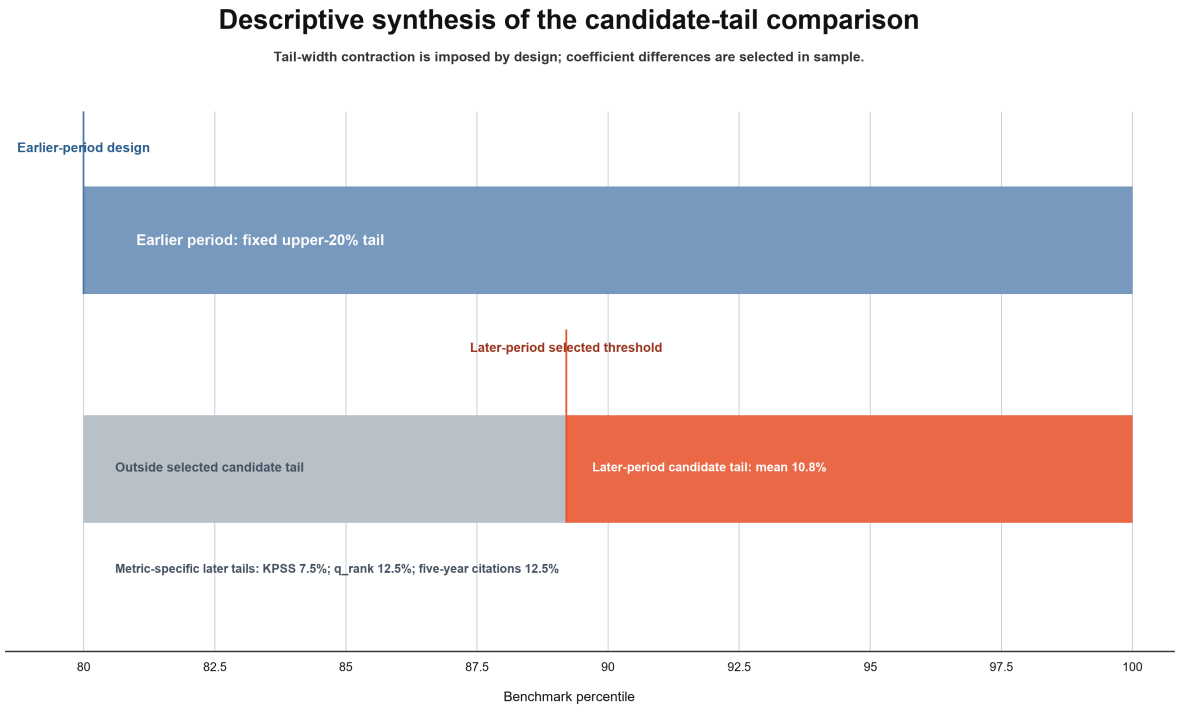


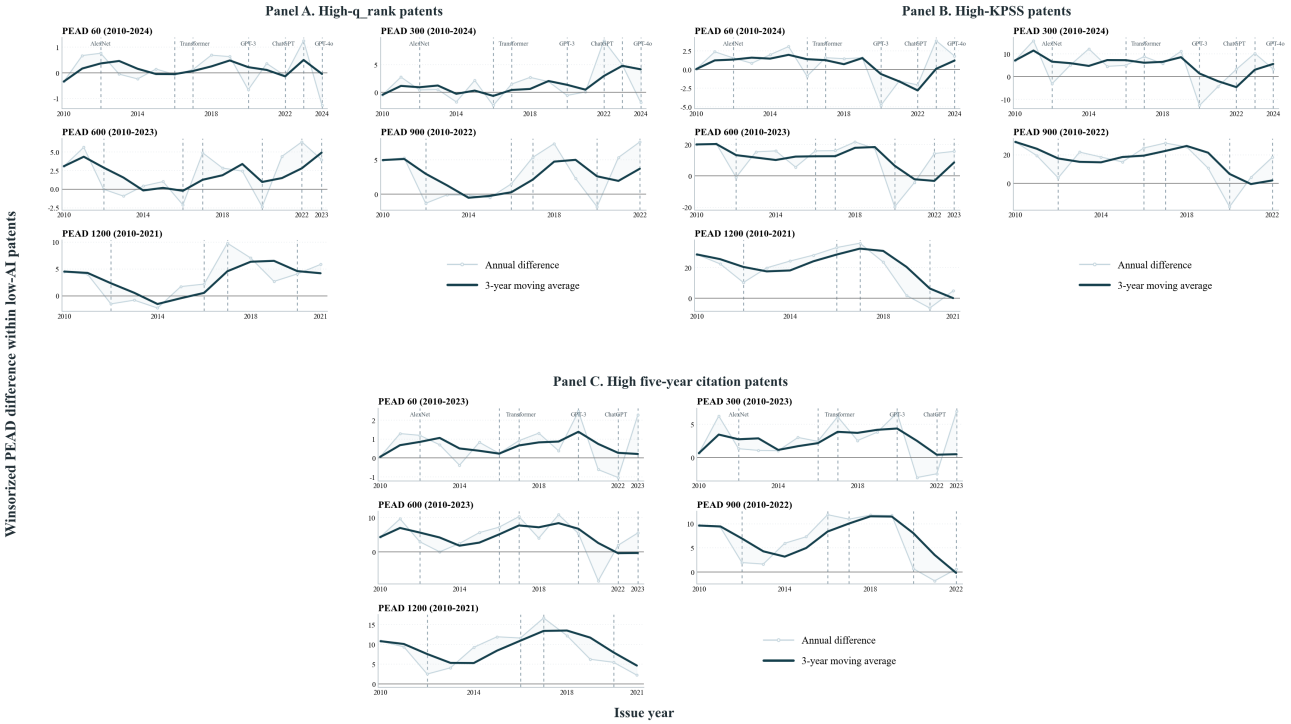
Figure 4. GPT-3 Cutoff and Metric-Specific Scope-versus-Severity (continued)

Panel C. Descriptive synthesis of the candidate-tail comparison



Notes: This exploratory figure compares a broad earlier-period benchmark tail with a smaller later-period candidate tail around the paper's GPT-3 cutoff. The sample begins in 2010, and the cutoff is June 11, 2020. For KPSS, q_rank, and five-year forward citations, the earlier-period breakthrough indicator uses a strict same-issue-year upper-20% patent threshold. Candidate later-period thresholds are restricted to the annual upper 7.5%, 10%, and 12.5%. Because every candidate is narrower than the earlier-period upper-20% range, the displayed contraction is imposed by design rather than estimated. For each benchmark, the retained later-period threshold is the candidate producing the largest absolute implied later-period PEPD600 coefficient in the pooled-interaction specification. This outcome-dependent selection chooses upper 7.5% for KPSS and upper 12.5% for both q_rank and five-year citations; their average later-period tail width is 10.8%. Panel A displays the resulting patent-level candidate-tail comparison. Because returns are observed at the firm-grant-date level, the indicator equals one when at least one patent in the day's batch exceeds the applicable same-issue-year cutoff; realized firm-date incidence can therefore exceed the stated patent-level tail width. Panel B follows the pooled-interaction logic of Table 8: PEPD600 is regressed on the firm-date mean LLM grant probability, the period-specific breakthrough indicator, Post, both Post interactions, patent count, lagged log market capitalization, profitability, log book-to-market, book leverage, R&D intensity, capital expenditures, operating leverage, cash, and issue-year fixed effects. PEPD600 is winsorized at the 1st and 99th percentiles within issue year, and standard errors are clustered by firm. If γ is the earlier-period breakthrough coefficient and ψ its Post interaction, the absolute coefficient is $|\gamma|$ before the cutoff and $|\gamma+\psi|$ afterward. The reported percentage change is $100\times(|\gamma+\psi|/|\gamma|-1)$. Absolute PEPD600 coefficients rise from 0.834 to 5.886 for KPSS, from 0.532 to 1.952 for q_rank, and from 2.559 to 3.032 for five-year citations, giving changes of 605.5%, 266.6%, and 18.5%, respectively. Panel C summarizes the selected in-sample pattern: a smaller candidate patent tail accompanies a larger absolute return coefficient. These are descriptive cutoff diagnostics rather than estimates of an AI-adoption treatment. The original fixed upper-20%/upper-10% placebo and separate-subsample bootstrap diagnostics are retained in Appendix Figure A1.

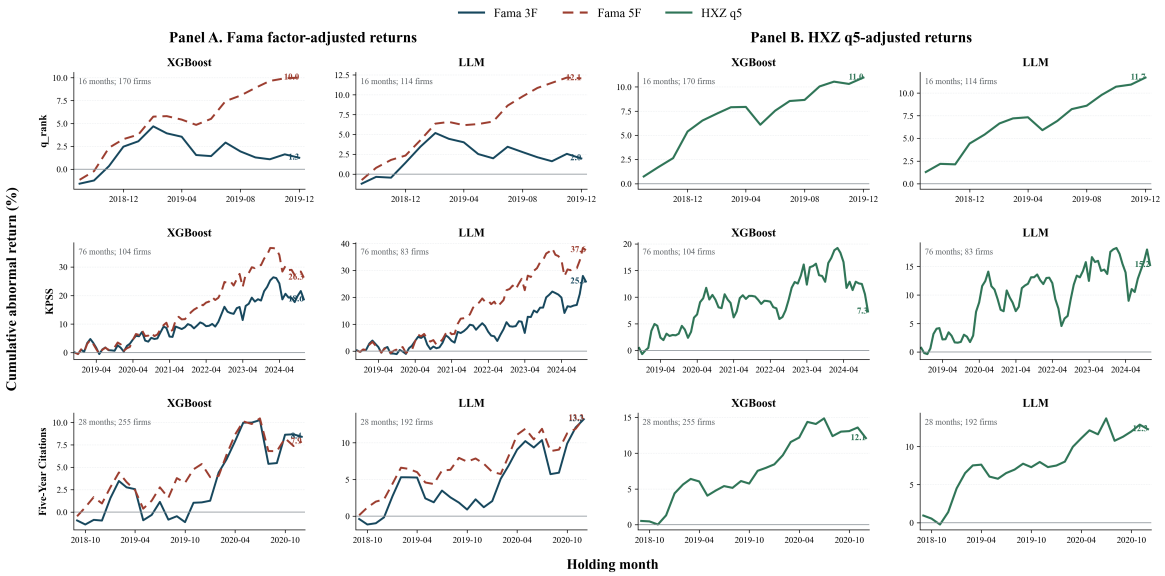
Figure 5. Annual Post-Grant Return Differences Within Low-AI Patent Sorts



Low AI is the bottom 30% of the within-year average percentile rank of XGBoost and LLM scores. Difference = High objective quality minus non-high objective quality within low-AI patents. Longer horizons stop at their last full-window issue year.

Notes: Low AI is the bottom 30% of the within-year average percentile rank of the XGBoost and LLM-embedding grant-probability scores. Separately for each issue year, a benchmark-top-decile patent is at or above the 90th percentile of q_rank, KPSS grant-window value, or five-year forward citations. Each series is the mean post-grant cumulative abnormal return of benchmark-top-decile patents minus the mean for the remaining patents within the low-AI subset. Returns are winsorized within issue year at the 1st and 99th percentiles. Subplot endpoints differ because benchmark coverage and completed return windows differ. The dark line is a three-year moving average and the light line is the annual difference. Vertical markers denote selected AI release dates and organize a descriptive comparison across calendar periods; they do not measure investor or examiner adoption. Citations and q_rank use post-grant information.

Figure 6. Later-Period Retrospective Portfolios: Cumulative Factor Residuals



Notes: This figure plots cumulative monthly factor-model residuals for equal-weighted retrospective portfolios in a calendar window beginning in September 2018. The start date defines a descriptive subsample and is not an AI-release or adoption treatment. A firm is classified into a portfolio when a patent issued in the prior 12 months is in the same-issue-year top decile of q_rank, KPSS, or five-year citations and in the bottom 30% of the corresponding XGBoost or LLM grant-probability distribution. Panel A uses Fama three- and five-factor residuals; Panel B uses HXZ q5. Effective holding months differ because portfolio eligibility and factor coverage differ. Five-year citations and q_rank use post-grant information, KPSS uses the grant-window market response, and the annual thresholds use completed issue-year distributions. The portfolios are therefore not implementable in real time. The paths cumulate regression residuals and should not be read as compounded portfolio wealth.

Table 1. Summary Statistics and Correlations

Notes: This table reports summary statistics and pairwise Pearson correlations for the main variables. The AI variables are model-implied grant probabilities. KPSS is a grant-window market-value measure, citations are ex post technological-use measures, and q_rank is an ex post text-based breakthrough-direction rank. Post-grant cumulative abnormal returns, denoted PEPD in the table, are measured in percentage points. Sample sizes vary across variables because coverage differs.

Panel A: Descriptive Statistics

Variable	Mean	Std. Dev.	Median	p10	p90	Min	Max
XGBoost Probability	0.630	0.113	0.643	0.472	0.765	0.139	0.928
LLM Probability	0.626	0.201	0.681	0.322	0.852	0.015	1.000
KPSS Patent Value	12.092	37.231	3.545	0.111	27.411	0.000	6,239.247
Forward Citations	13.952	53.247	4.000	0.000	29.000	0.000	5,185.000
5-Year Forward Citations	1.669	8.691	0.000	0.000	4.000	0.000	828.000
Breakthrough Rank Index q_rank	0.169	0.082	0.179	0.047	0.267	0.000	0.684
PEPD 60	0.022	14.577	-0.086	-15.359	15.205	-298.590	3,881.081
PEPD 300	0.451	31.863	-0.203	-34.033	34.848	-662.879	3,812.649
PEPD 600	0.603	45.073	-0.452	-47.767	49.751	-2,156.322	3,788.389
PEPD 900	3.501	49.669	2.489	-51.596	59.233	-851.864	3,664.951

Panel B: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)
(1) XGBoost Probability	1.000					
(2) LLM Probability	0.288	1.000				
(3) KPSS Patent Value	0.006	-0.054	1.000			
(4) Forward Citations	0.011	0.004	0.034	1.000		
(5) 5-Year Citations	-0.003	-0.031	0.026	0.495	1.000	
(6) Breakthrough Rank Index q_rank	-0.080	-0.189	0.008	0.027	0.019	1.000

Table 2. AI Grant-Probability Scores, Patent Benchmarks, and Future Changes in ROA

Notes: The dependent variable in column $t+h$ is the annual change $ROA_{t+h} - ROA_{t+h-1}$; the three outcomes are annual rather than cumulative changes. Coefficients are in ROA units, so 0.01 equals one percentage point. Annual ROA changes are winsorized at the 1st and 99th percentiles within grant year. The observation is a PERMCO–grant-year. Mean Probability is the raw firm-year mean XGBoost or LLM grant probability. Breakthrough Innovation equals one when at least one patent is strictly above the relevant 90th-percentile cutoff: 27.411 for KPSS, 5 citations received within five years after grant, and 0.267 for q_rank. The five-year-citation cutoff is computed using the 1976–2019 grant cohorts with a complete five-year exposure window; the strict rule therefore requires at least six citations. Exact duplicate patent-number records are counted once. All regressions include raw patent count, lagged log market capitalization, profitability, log book-to-market, book leverage, R&D, CapEx, operating leverage, cash, and grant-year fixed effects; control coefficients are omitted. Patent Count coefficients are rescaled to represent a 100-patent increase; regressions use the raw count. Firm fixed effects are not included. Standard errors clustered by PERMCO are in parentheses. Five-year citations and q_rank are ex post measures. The coefficients are conditional cross-sectional associations, not causal effects or grant-date forecasts. *, **, and *** denote unadjusted $p < 0.10$, $p < 0.05$, and $p < 0.01$.

	XGBoost			LLM		
	$t+1$	$t+2$	$t+3$	$t+1$	$t+2$	$t+3$
Panel A: KPSS Grant-Window Value						
Mean Probability	-0.0058	-0.0226	-0.0180	-0.0107	-0.0067	0.0018
	(0.0142)	(0.0146)	(0.0138)	(0.0076)	(0.0082)	(0.0075)
Breakthrough Innovation	0.0027	0.0081***	-0.0008	0.0025	0.0079***	-0.0005
	(0.0023)	(0.0021)	(0.0021)	(0.0023)	(0.0021)	(0.0021)
Patent Count (100 patents)	-0.00026*	-0.00060***	-0.00036***	-0.00029**	-0.00057***	-0.00032***
	(0.00013)	(0.00015)	(0.00010)	(0.00013)	(0.00014)	(0.00010)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	42,283	39,009	35,997	42,510	39,216	36,198
Adj. R-squared	0.021	0.018	0.019	0.021	0.018	0.019
Panel B: Five-Year Forward Citations						
Mean Probability	0.0063	-0.0142	-0.0139	-0.0076	0.0035	0.0048
	(0.0145)	(0.0148)	(0.0139)	(0.0080)	(0.0086)	(0.0078)
Breakthrough Innovation	0.0042*	0.0043*	0.0008	0.0036	0.0047**	0.0015
	(0.0023)	(0.0023)	(0.0021)	(0.0023)	(0.0023)	(0.0021)
Patent Count (100 patents)	-0.00042**	-0.00065***	-0.00036***	-0.00045**	-0.00061***	-0.00033***
	(0.00018)	(0.00018)	(0.00011)	(0.00018)	(0.00018)	(0.00011)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	38,453	36,177	34,046	38,687	36,392	34,252
Adj. R-squared	0.022	0.019	0.020	0.022	0.019	0.020
Panel C: Forward Text Breakthrough Rank (q_rank)						
Mean Probability	0.0041	-0.0190	-0.0244	-0.0079	-0.0049	0.0010
	(0.0154)	(0.0159)	(0.0151)	(0.0082)	(0.0089)	(0.0081)
Breakthrough Innovation	0.0024	0.0001	0.0044*	0.0017	0.0003	0.0037
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
Patent Count (100 patents)	-0.00032**	-0.00056***	-0.00027***	-0.00034**	-0.00054***	-0.00023**
	(0.00015)	(0.00015)	(0.00010)	(0.00015)	(0.00015)	(0.00010)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37,465	34,876	32,336	37,465	34,876	32,336
Adj. R-squared	0.021	0.018	0.019	0.021	0.018	0.019

Table 3. AI Grant-Probability Scores, Patent Benchmarks, and Acquisition Within One to Three Years

Notes: This table reports linear probability models. Panels A, B, and C correspond to $h = 1,2,3$, respectively. For horizon h, the dependent variable equals one when the patenting firm is strictly matched as a completed Orbis target within h years after the patent grant year; the outcomes are cumulative across horizons and measure targets only. The observation is a PERMCO–grant-year. Mean Probability is the raw firm-year mean grant score, and Breakthrough Innovation indicates at least one patent strictly above the relevant 90th-percentile cutoff: 27.411 for KPSS, 5 citations received within five years after grant, and 0.267 for q_rank. The five-year-citation cutoff and sample use 1976–2019 grant cohorts with a complete five-year exposure window; the strict rule requires at least six citations. Exact duplicate patent-number records are counted once. Five-year citations and q_rank are ex post measures. All regressions include raw patent count, the same eight lagged firm controls listed in Table 2, and grant-year fixed effects; control coefficients are omitted. Patent Count coefficients are rescaled to represent a 100-patent increase; regressions use the raw count. Firm fixed effects and score-by-quality interactions are not included. Standard errors clustered by PERMCO are in parentheses. Coefficients are conditional associations, not causal acquisition effects; score magnitudes are not directly comparable across XGBoost and LLM. *, **, and *** denote unadjusted $p < 0.10$, $p < 0.05$, and $p < 0.01$.

	KPSS		Five-Year Citations		q_rank	
	XGBoost	LLM	XGBoost	LLM	XGBoost	LLM
Panel A: Acquired Within One Year						
Mean Probability	-0.0273*	-0.0460***	-0.0306*	-0.0557***	-0.0296*	-0.0548***
	(0.0144)	(0.0077)	(0.0158)	(0.0086)	(0.0151)	(0.0082)
Breakthrough Innovation	0.0016	0.0024	0.0083***	0.0110***	0.0018	0.0055**
	(0.0030)	(0.0030)	(0.0027)	(0.0027)	(0.0025)	(0.0026)
Patent Count (100 patents)	-0.00049**	-0.00060***	-0.00043**	-0.00050***	-0.00041**	-0.00045***
	(0.00020)	(0.00020)	(0.00020)	(0.00018)	(0.00018)	(0.00017)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26,789	26,761	22,557	22,537	25,376	25,376
Acquisition Events	869	869	743	743	819	819
Event Rate	3.24%	3.25%	3.29%	3.30%	3.23%	3.23%
Adj. R-squared	0.006	0.008	0.007	0.009	0.006	0.008
Panel B: Acquired Within Two Years						
Mean Probability	-0.0551**	-0.0786***	-0.0614**	-0.0906***	-0.0552**	-0.0854***
	(0.0234)	(0.0132)	(0.0248)	(0.0142)	(0.0239)	(0.0136)
Breakthrough Innovation	0.0102**	0.0115**	0.0131***	0.0170***	0.0010	0.0046
	(0.0049)	(0.0050)	(0.0040)	(0.0039)	(0.0039)	(0.0039)
Patent Count (100 patents)	-0.00115***	-0.00132***	-0.00092**	-0.00102***	-0.00100***	-0.00106***
	(0.00038)	(0.00038)	(0.00037)	(0.00035)	(0.00036)	(0.00034)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,961	25,933	22,557	22,537	25,085	25,085
Acquisition Events	1,586	1,584	1,398	1,396	1,527	1,527
Event Rate	6.11%	6.11%	6.20%	6.19%	6.09%	6.09%
Adj. R-squared	0.011	0.014	0.012	0.015	0.012	0.014
Panel C: Acquired Within Three Years						
Mean Probability	-0.0880***	-0.1131***	-0.0873***	-0.1252***	-0.0788**	-0.1180***
	(0.0318)	(0.0183)	(0.0329)	(0.0192)	(0.0322)	(0.0186)
Breakthrough Innovation	0.0193***	0.0209***	0.0151***	0.0203***	0.0021	0.0056
	(0.0067)	(0.0067)	(0.0052)	(0.0051)	(0.0052)	(0.0051)
Patent Count (100 patents)	-0.00178***	-0.00201***	-0.00144***	-0.00157***	-0.00155***	-0.00163***
	(0.00058)	(0.00058)	(0.00054)	(0.00051)	(0.00055)	(0.00051)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,088	25,060	22,557	22,537	24,644	24,644
Acquisition Events	2,241	2,237	2,040	2,036	2,200	2,200
Event Rate	8.93%	8.93%	9.04%	9.03%	8.93%	8.93%
Adj. R-squared	0.016	0.019	0.016	0.020	0.016	0.019

Table 4. XGBoost Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns

Notes: This table reports the baseline firm-grant-date regressions using the mean XGBoost grant probability. Panels A–C use upper-tail indicators for KPSS grant-window value, total forward citations, and q_rank, respectively. Benchmark indicators use full patent-sample 90th-percentile thresholds. PEPD is the post-grant cumulative size-adjusted return in percentage points. All specifications include patent count, the full set of lagged firm controls, and grant-year fixed effects; control-variable coefficients are omitted for brevity and reported in Appendix Table A5. Standard errors are clustered by firm. Total citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. The regressions are descriptive. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A: KPSS Grant-Window Value

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	-0.246	-3.625**	-5.046*	-9.704*
	(0.400)	(1.534)	(2.811)	(4.963)
Breakthrough Innovation (KPSS)	0.548***	1.872***	3.238***	4.467**
	(0.130)	(0.562)	(1.044)	(1.746)
Patent Count	0.009*	0.038	0.067	0.038
	(0.006)	(0.027)	(0.052)	(0.075)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	541285	536564	523817	305012
Adj. R-squared	0.016	0.067	0.088	0.081

Panel B: Total Forward Citations

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	-0.179	-3.389**	-4.597	-9.279*
	(0.402)	(1.542)	(2.826)	(4.956)
Breakthrough Innovation (Cites)	0.470***	1.665***	3.287***	1.873**
	(0.075)	(0.263)	(0.494)	(0.898)
Patent Count	0.001	0.009	0.014	-0.006
	(0.005)	(0.025)	(0.047)	(0.069)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	541285	536564	523817	305012
Adj. R-squared	0.016	0.067	0.088	0.080

Panel C: Forward Text Breakthrough Rank (q_rank)

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	0.355	-1.253	-1.509	-5.278
	(0.420)	(1.578)	(2.914)	(5.370)
Breakthrough Innovation (q_rank)	0.111	0.884**	1.743**	-0.645
	(0.096)	(0.370)	(0.755)	(1.347)
Patent Count	0.001	-0.001	-0.001	-0.029
	(0.005)	(0.023)	(0.044)	(0.062)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table 5. LLM-Embedding Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns

Notes: This table repeats Table 4 using the firm-date mean LLM grant probability. Panels A–C use upper-tail indicators for KPSS grant-window value, total forward citations, and q_rank, respectively. All specifications include patent count, the full set of lagged firm controls, and grant-year fixed effects; control-variable coefficients are omitted for brevity and reported in Appendix Table A6. Standard errors are clustered by firm. Total citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. The regressions are descriptive. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A: KPSS Grant-Window Value

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.685**	-3.073***	-5.412**	-2.434
	(0.281)	(1.182)	(2.171)	(3.621)
Breakthrough Innovation (KPSS)	0.577***	1.994***	3.439***	4.980***
	(0.130)	(0.561)	(1.041)	(1.748)
Patent Count	0.009	0.040	0.070	0.053
	(0.006)	(0.028)	(0.052)	(0.077)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	545790	541061	528257	309485
Adj. R-squared	0.016	0.068	0.089	0.081

Panel B: Total Forward Citations

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.582**	-2.707**	-4.680**	-2.184
	(0.287)	(1.199)	(2.198)	(3.663)
Breakthrough Innovation (Cites)	0.448***	1.601***	3.153***	1.921**
	(0.075)	(0.260)	(0.487)	(0.880)
Patent Count	0.001	0.011	0.017	0.006
	(0.005)	(0.025)	(0.047)	(0.070)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	545790	541061	528257	309485
Adj. R-squared	0.016	0.068	0.089	0.080

Panel C: Forward Text Breakthrough Rank (q_rank)

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.358	-1.753	-2.824	1.847
	(0.291)	(1.207)	(2.230)	(3.784)
Breakthrough Innovation (q_rank)	0.067	0.756**	1.524**	-0.302
	(0.094)	(0.354)	(0.727)	(1.299)
Patent Count	0.000	-0.001	-0.001	-0.026
	(0.005)	(0.023)	(0.044)	(0.062)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table 6. Grant-Probability Scores, Five-Year Citations, and Post-Grant Returns

Notes: This table reports firm-grant-date regressions that use five-year forward citations as the upper-tail patent benchmark. Panel A uses the firm-date mean XGBoost grant probability, and Panel B uses the firm-date mean LLM grant probability. Breakthrough Innovation equals one when the firm-date batch contains at least one patent above the full patent-sample 90th percentile of five-year forward citations. PEPD is the post-grant cumulative size-adjusted return in percentage points. All specifications include patent count, the full set of lagged firm controls, and grant-year fixed effects. The control-variable coefficients are omitted from this main-text presentation and reported in Appendix Table A7, which also defines the controls. Standard errors are clustered by firm. Five-year citations use post-grant information and are not observable at grant; the regressions are retrospective and descriptive. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A: XGBoost

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	-0.131	-3.168**	-4.269	-9.118*
	(0.401)	(1.536)	(2.819)	(4.954)
Breakthrough Innovation (5-Year Cites)	0.338***	1.394***	2.346***	1.255
	(0.083)	(0.309)	(0.584)	(0.912)
Patent Count	0.001	0.007	0.016	-0.006
	(0.005)	(0.025)	(0.047)	(0.068)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	541285	536564	523817	305012
Adj. R-squared	0.016	0.067	0.088	0.080

Panel B: LLM

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.590**	-2.666**	-4.736**	-2.131
	(0.287)	(1.195)	(2.193)	(3.666)
Breakthrough Innovation (5-Year Cites)	0.317***	1.336***	2.226***	1.390
	(0.083)	(0.304)	(0.575)	(0.900)
Patent Count	0.001	0.010	0.018	0.005
	(0.005)	(0.025)	(0.047)	(0.069)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	545790	541061	528257	309485
Adj. R-squared	0.016	0.068	0.089	0.080

Table 7. Grant-Probability Scores with Multiple Patent Benchmarks

Notes: This table reports specifications that enter two upper-tail patent benchmarks jointly. Panel A uses the firm-date mean XGBoost grant probability, and Panel B uses the firm-date mean LLM grant probability. Within each panel, the first block jointly includes KPSS and q_rank indicators, while the second block jointly includes five-year-citation and q_rank indicators. Each indicator equals one when the firm-date batch contains at least one patent above the relevant full patent-sample 90th percentile. PEPD is the post-grant cumulative size-adjusted return in percentage points. All regressions include patent count, the full set of lagged firm controls, and grant-year fixed effects. Control-variable coefficients are omitted and reported in Appendix Tables A8 and A9. Standard errors are clustered by firm. Five-year citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. The regressions are descriptive. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A: XGBoost

KPSS and q_rank *specification*

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	0.344	-1.284	-1.563	-5.406
	(0.418)	(1.569)	(2.895)	(5.377)
Breakthrough Innovation (KPSS)	0.455***	1.380**	2.470**	2.250
	(0.135)	(0.580)	(1.069)	(1.759)
Breakthrough Innovation (q_rank)	0.139	0.970***	1.898**	-0.452
	(0.097)	(0.375)	(0.762)	(1.362)
Patent Count	0.004	0.008	0.016	-0.015
	(0.005)	(0.024)	(0.046)	(0.066)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Five-Year Citations and q_rank *specification*

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	0.453	-0.864	-0.870	-4.959
	(0.419)	(1.573)	(2.907)	(5.374)
Breakthrough Innovation (5-Year Cites)	0.446***	1.756***	2.872***	1.986**
	(0.087)	(0.309)	(0.577)	(0.902)
Breakthrough Innovation (q_rank)	0.022	0.536	1.175	-0.997
	(0.095)	(0.352)	(0.717)	(1.305)
Patent Count	-0.003	-0.016	-0.026	-0.043
	(0.005)	(0.024)	(0.045)	(0.061)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table 7. Grant-Probability Scores with Multiple Patent Benchmarks (continued)

Notes: Panel B repeats Panel A using the firm-date mean LLM grant probability. Both blocks include patent count, lagged firm controls, and grant-year fixed effects. The omitted control-variable coefficients and complete estimates appear in Appendix Table A9. Standard errors are clustered by firm.

Panel B: LLM

KPSS and q_rank *specification*

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.342	-1.703	-2.731	1.906
	(0.287)	(1.192)	(2.202)	(3.761)
Breakthrough Innovation (KPSS)	0.453***	1.363**	2.443**	2.238
	(0.135)	(0.578)	(1.065)	(1.759)
Breakthrough Innovation (q_rank)	0.097	0.847**	1.688**	-0.100
	(0.094)	(0.358)	(0.731)	(1.316)
Patent Count	0.003	0.008	0.015	-0.013
	(0.005)	(0.024)	(0.046)	(0.066)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Five-Year Citations and q_rank *specification*

	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.235	-1.265	-2.029	2.228
	(0.292)	(1.203)	(2.225)	(3.784)
Breakthrough Innovation (5-Year Cites)	0.432***	1.716***	2.803***	2.092**
	(0.088)	(0.307)	(0.573)	(0.897)
Breakthrough Innovation (q_rank)	-0.010	0.451	1.025	-0.641
	(0.093)	(0.341)	(0.697)	(1.263)
Patent Count	-0.004	-0.016	-0.026	-0.041
	(0.005)	(0.023)	(0.045)	(0.061)
Firm Controls	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table 8. Grant-Probability Scores Around the GPT-3 Cutoff

Notes: Panels A and B report the XGBoost and LLM specifications, respectively. Post equals one for grant dates on or after the paper's June 11, 2020 GPT-3 cutoff. The cutoff describes calendar-time heterogeneity and is not an AI-adoption treatment or exogenous shock. The three blocks use KPSS, five-year forward citations, and q_rank as separate upper-tail patent benchmarks. Each benchmark indicator is defined using its full patent-sample 90th percentile and is entered with its Post interaction. All regressions include patent count, the full set of lagged firm controls, and grant-year fixed effects; control-variable coefficients are omitted and reported in Appendix Tables A10 and A11. Standard errors are clustered by firm. Citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panels A and B: XGBoost and LLM

Benchmark: KPSS

	Panel A: XGBoost			Panel B: LLM		
	PEPD60	PEPD300	PEPD600	PEPD60	PEPD300	PEPD600
Mean Probability	0.403	-0.659	0.741	-0.278	-1.145	-2.038
	(0.421)	(1.525)	(2.779)	(0.295)	(1.168)	(2.162)
Breakthrough Innovation	0.233*	0.281	0.435	0.272*	0.423	0.698
	(0.139)	(0.592)	(1.090)	(0.139)	(0.594)	(1.093)
Mean Probability × Post	-3.509***	-15.797***	-31.676***	-1.638***	-7.710***	-13.714***
	(0.999)	(3.873)	(7.348)	(0.602)	(2.448)	(4.646)
Breakthrough Innovation × Post	1.116***	5.564***	9.983***	1.067***	5.421***	9.636***
	(0.241)	(1.030)	(2.063)	(0.239)	(1.025)	(2.058)
Post	-1.832**	7.704***	20.602***	-3.056***	2.414	8.829***
	(0.790)	(2.732)	(4.976)	(0.611)	(1.905)	(3.216)
Patent Count	0.009*	0.037	0.067	0.009*	0.040	0.069
	(0.006)	(0.027)	(0.051)	(0.006)	(0.027)	(0.051)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	541285	536564	523817	545790	541061	528257
Adj. R-squared	0.016	0.068	0.089	0.016	0.068	0.090

Benchmark: Five-Year Forward Citations

	Panel A: XGBoost			Panel B: LLM		
	PEPD60	PEPD300	PEPD600	PEPD60	PEPD300	PEPD600
Mean Probability	0.443	-0.513	1.012	-0.182	-0.671	-1.198
	(0.417)	(1.496)	(2.726)	(0.294)	(1.143)	(2.107)
Breakthrough Innovation	0.279***	1.241***	2.181***	0.270***	1.258***	2.180***
	(0.086)	(0.317)	(0.604)	(0.085)	(0.310)	(0.588)
Mean Probability × Post	-3.427***	-15.587***	-31.404***	-1.809***	-8.692***	-15.555***
	(1.029)	(4.020)	(7.579)	(0.638)	(2.626)	(4.947)
Breakthrough Innovation × Post	0.633**	1.739*	2.165	0.643**	1.694	2.107
	(0.250)	(1.037)	(1.827)	(0.252)	(1.039)	(1.832)
Post	-1.884**	7.949***	21.380***	-2.958***	3.398*	10.893***
	(0.815)	(2.858)	(5.214)	(0.634)	(2.037)	(3.493)
Patent Count	0.000	0.004	0.010	-0.000	0.005	0.011
	(0.005)	(0.024)	(0.046)	(0.005)	(0.024)	(0.046)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	541285	536564	523817	545790	541061	528257
Adj. R-squared	0.016	0.067	0.089	0.016	0.068	0.089

Benchmark: q_rank

	Panel A: XGBoost			Panel B: LLM		
	PEPD60	PEPD300	PEPD600	PEPD60	PEPD300	PEPD600
Mean Probability	0.639	0.124	1.500	-0.204	-0.712	-1.166
	(0.442)	(1.581)	(2.882)	(0.305)	(1.202)	(2.213)
Breakthrough Innovation	0.079	0.656*	1.306*	0.038	0.591	1.150
	(0.101)	(0.378)	(0.774)	(0.099)	(0.364)	(0.749)
Mean Probability × Post	-2.559**	-12.543***	-27.104***	-1.005	-6.753**	-11.054**
	(1.172)	(4.490)	(7.912)	(0.695)	(2.710)	(4.845)
Breakthrough Innovation × Post	0.341	2.305**	4.525**	0.372	2.310**	4.640**
	(0.269)	(1.135)	(2.078)	(0.268)	(1.123)	(2.060)
Post	-2.325***	6.066*	18.347***	-3.351***	2.221	7.782**
	(0.886)	(3.160)	(5.417)	(0.648)	(2.063)	(3.406)
Patent Count	-0.001	-0.008	-0.015	-0.001	-0.009	-0.016
	(0.005)	(0.022)	(0.042)	(0.005)	(0.022)	(0.042)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Grant-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	464880	460629	451224	464880	460629	451224
Adj. R-squared	0.017	0.070	0.091	0.017	0.070	0.091

Table 9. Retrospective Long-Only Portfolio Alphas: Fama and HXZ Factor Models, 2010–2024

Notes: This table reports time-series regressions of equal-weighted portfolio excess returns. Panel A uses the Fama three- and five-factor models, and Panel B uses HXZ q5. Portfolios are rebalanced monthly using patents granted in the prior 12 months and include firms with at least one patent in the objective-quality top decile and the corresponding grant-probability score bottom 30%. Each column uses months in which the portfolio contains at least five firms. Monthly alpha is in percentage points; annual alpha is the linear quantity 12 times monthly alpha, not a compounded return, and has the same t-statistic. Factor loadings are in levels. Newey–West t-statistics with lag 6 appear in parentheses. Five-year citations and q_rank use post-grant information, KPSS uses the grant-window market response, and the annual cutoffs require completed issue-year distributions. All six portfolios are retrospective rather than implementable trading strategies. *, **, and *** denote unadjusted significance at the 10%, 5%, and 1% levels.

	q_rank		KPSS		Five-Year Citations	
	XGBoost	LLM	XGBoost	LLM	XGBoost	LLM
Panel A1: Fama Three-Factor Model						
Monthly α	0.276**	0.263*	0.286***	0.258**	0.359***	0.385***
	(1.99)	(1.81)	(3.07)	(2.49)	(2.67)	(2.61)
Annual α	3.312**	3.154*	3.428***	3.091**	4.308***	4.616***
Mkt-RF	1.144***	1.109***	1.048***	1.033***	1.134***	1.089***
	(31.91)	(31.32)	(43.60)	(45.01)	(32.47)	(33.21)
SMB	0.403***	0.467***	0.082	-0.020	0.533***	0.401***
	(3.54)	(3.94)	(1.25)	(-0.32)	(5.59)	(4.05)
HML	-0.186**	-0.200**	-0.346***	-0.479***	-0.243***	-0.256***
	(-2.03)	(-2.25)	(-8.53)	(-10.49)	(-2.81)	(-2.64)
Panel A2: Fama Five-Factor Model						
Monthly α	0.369**	0.398***	0.406***	0.428***	0.473***	0.528***
	(2.39)	(2.67)	(4.33)	(4.30)	(3.42)	(3.61)
Annual α	4.423**	4.775***	4.871***	5.139***	5.674***	6.338***
Mkt-RF	1.109***	1.060***	1.019***	0.993***	1.098***	1.046***
	(23.39)	(22.01)	(42.86)	(38.47)	(30.50)	(33.24)
SMB	0.326***	0.353***	-0.050	-0.210**	0.447***	0.297***
	(2.70)	(2.62)	(-0.67)	(-2.57)	(4.41)	(2.97)
HML	-0.215	-0.260	-0.359***	-0.484***	-0.306**	-0.313**
	(-1.48)	(-1.60)	(-6.55)	(-7.29)	(-2.53)	(-2.56)
RMW	-0.275	-0.413*	-0.397***	-0.553***	-0.456***	-0.554***
	(-1.31)	(-1.69)	(-4.85)	(-4.77)	(-3.09)	(-3.54)
CMA	-0.223	-0.260	-0.199**	-0.270**	-0.317	-0.365
	(-0.92)	(-0.94)	(-2.04)	(-2.37)	(-1.45)	(-1.57)
Panel B: HXZ q5 Model						
Monthly α	0.051	0.053	0.052	0.043	0.110	0.132
	(0.39)	(0.34)	(0.60)	(0.44)	(0.89)	(0.91)
Annual α	0.609	0.634	0.618	0.521	1.322	1.582
MKT	1.085***	1.043***	1.031***	1.030***	1.063***	1.026***
	(23.67)	(20.72)	(46.74)	(42.67)	(33.16)	(26.73)
ME	0.272***	0.308***	0.056	-0.008	0.409***	0.378***
	(3.80)	(3.57)	(1.07)	(-0.14)	(7.28)	(5.98)
IA	-0.259**	-0.300***	-0.160***	-0.277***	-0.199**	-0.234**
	(-2.54)	(-2.74)	(-2.68)	(-3.76)	(-2.16)	(-2.45)
ROE	-0.250***	-0.212**	-0.184***	-0.202**	-0.326***	-0.350***
	(-2.72)	(-2.23)	(-2.83)	(-2.56)	(-4.70)	(-5.62)
EG	0.121	0.098	0.146*	0.184*	0.296***	0.338***
	(1.64)	(1.07)	(1.84)	(1.71)	(3.07)	(3.39)
Months	119	119	179	179	131	131
Median firms	187	190	122	87	262	201

Table 10. Later-Period Retrospective Portfolio Alphas

Notes: This table repeats the portfolio alpha tests using holding months beginning in September 2018, matching the descriptive later-period window in Figure 6; the start date is not an AI-adoption treatment. Portfolios are equal weighted and rebalanced monthly using patents granted in the prior 12 months. A firm is classified into a portfolio if it has at least one patent in the objective-quality top decile and the corresponding grant-probability score bottom 30%. Each regression uses months in which the portfolio contains at least five firms. Monthly alpha is in percentage points; annual alpha is the linear quantity 12 times monthly alpha, not a compounded return, and has the same t-statistic. Newey–West t-statistics use lag 6 and appear in parentheses. The effective samples contain 16 months for q_rank, 76 for KPSS, and 28 for five-year citations. Post-grant information, grant-window returns, and completed annual cutoffs make all portfolios retrospective and nonimplementable in real time. *, **, and *** denote unadjusted significance at the 10%, 5%, and 1% levels.

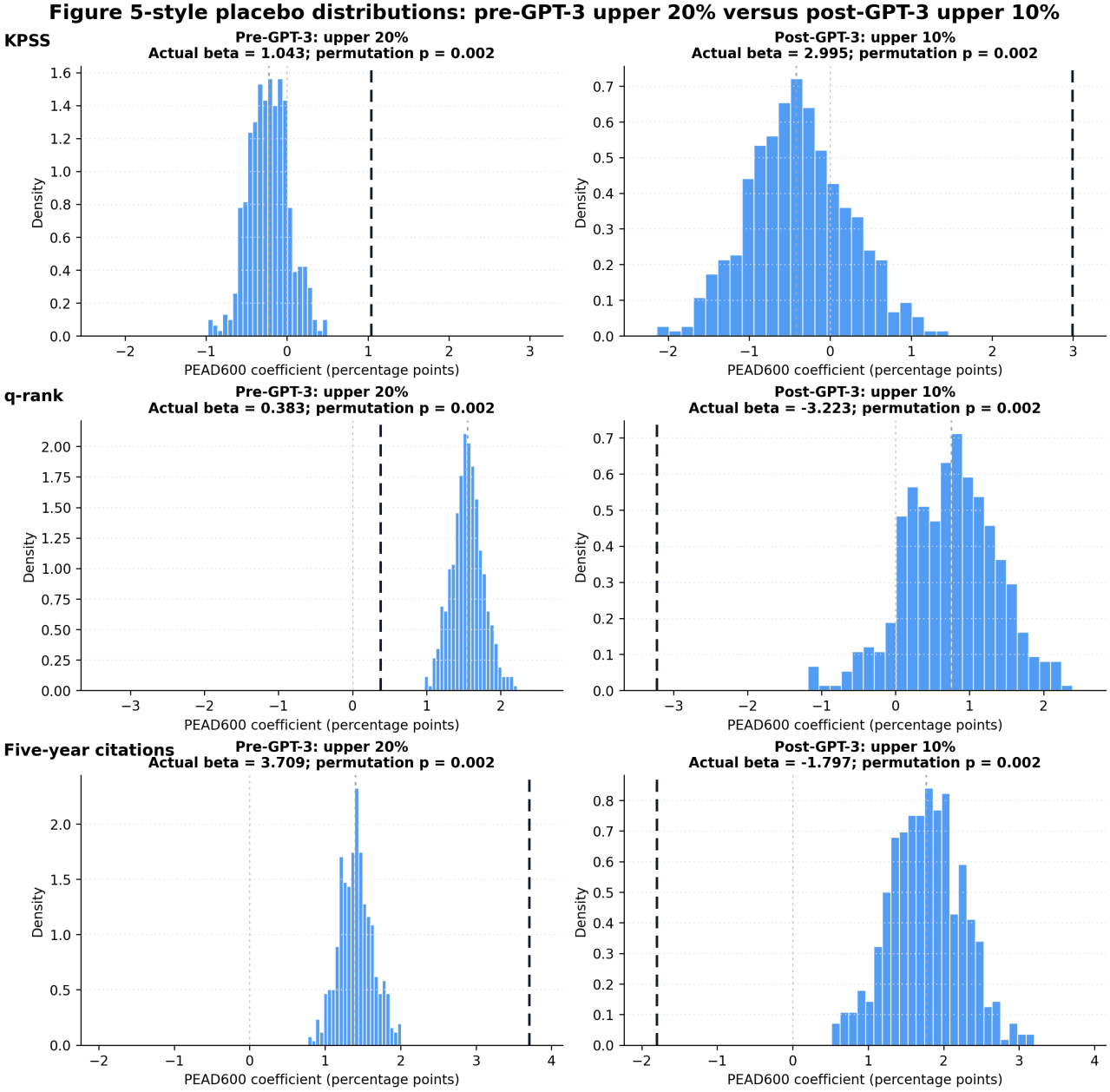
	q_rank		KPSS		Five-Year Citations	
	XGBoost	LLM	XGBoost	LLM	XGBoost	LLM
Fama 3F monthly α	0.084	0.129	0.238	0.319*	0.303	0.458
	(0.23)	(0.35)	(1.47)	(1.68)	(1.05)	(1.19)
Fama 5F monthly α	0.602***	0.717***	0.318*	0.432**	0.280*	0.455**
	(3.02)	(3.05)	(1.85)	(2.47)	(1.84)	(2.36)
HXZ q5 monthly α	0.656	0.697***	0.104	0.200	0.415**	0.421**
	(1.65)	(2.59)	(0.65)	(1.07)	(2.17)	(2.44)
Fama 3F annual α	1.012	1.546	2.858	3.832*	3.630	5.496
Fama 5F annual α	7.228***	8.608***	3.818*	5.189**	3.357*	5.459**
HXZ q5 annual α	7.872	8.366***	1.242	2.400	4.976**	5.050**
Months	16	16	76	76	28	28
Median firms	170	114	104	83	255	192

Appendix

Original GPT-3 Cutoff Diagnostics

Figure A1. Original GPT-3 Cutoff Placebo and Bootstrap Diagnostics

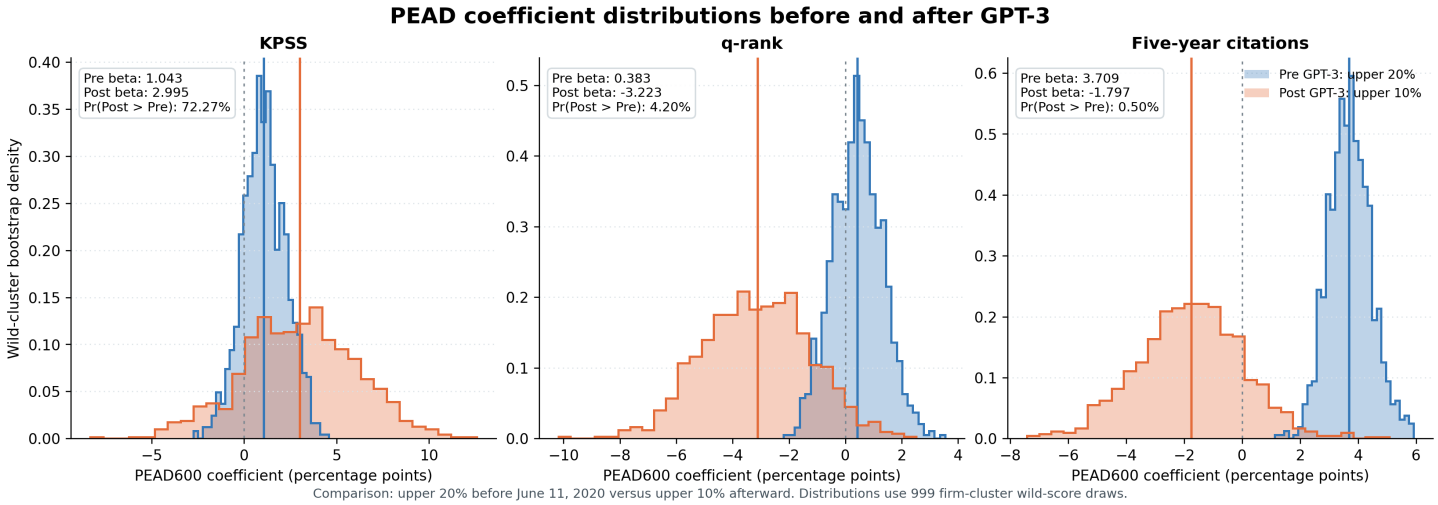
Panel A. Original Figure 5-style placebo distributions



Notes: high-quality labels are shuffled within issue-year x patent-count strata. Dashed black lines are actual coefficients. PEAD is winsorized at 1%/99% by issue year; regressions include the LLM score, firm controls, and issue-year fixed effects.

Figure A1. Original GPT-3 Cutoff Placebo and Bootstrap Diagnostics (continued)

Panel B. Original pre-upper-20% versus post-upper-10% bootstrap distributions



Notes: This appendix figure preserves the original fixed-cutoff diagnostics formerly reported as Panels A and B of Figure 4. The cutoff date is June 11, 2020. The pre-period indicator identifies firm-grant dates containing at least one patent above the same-issue-year 80th percentile of KPSS, q_rank, or five-year forward citations; the post-period indicator analogously uses the same-issue-year 90th percentile. PEPD600 is winsorized at the 1st and 99th percentiles within issue year. The underlying regressions include the firm-date mean LLM grant probability, the full set of lagged firm controls, and issue-year fixed effects. In Panel A, high-quality labels are permuted within issue-year-by-patent-count strata; the histograms are the resulting placebo coefficient distributions, the dashed black lines are the actual coefficients, and the displayed probabilities are calculated from the shuffled-label distributions. Panel B plots firm-cluster wild-score bootstrap distributions based on 999 draws for the separately estimated pre- and post-cutoff coefficients and reports the bootstrap probability that the post coefficient exceeds the pre coefficient. The appendix panels retain the original common 20%/10% tail definition and separate-subsample regressions. Main-text Figure 4 fixes the pre-period patent tail at 20%, restricts the post-period patent tail to 7.5%–12.5%, and uses a Table 8-style pooled-interaction specification to compare implied pre- and post-period pricing severity.

Variable Definitions

Table A1. Definitions of Firm-Level Control Variables

Notes: This table defines the firm-level control variables used in the empirical analysis. Accounting variables are obtained from Compustat Fundamentals Annual and are lagged by one year to mitigate look-ahead bias. Market equity is measured using Compustat in annual specifications and CRSP Monthly Stock File variables when monthly market capitalization is required.

Variable	**Definition**
ln(Market capitalization)	Natural logarithm of market capitalization of a firm's common equity. In annual regressions, market capitalization is measured using Compustat fiscal year-end price times common shares outstanding, $prcc_f \times csho$. In regressions requiring CRSP market equity, market capitalization is measured using the CRSP Monthly Stock File as the absolute value of stock price times shares outstanding, $\lvert prc \rvert \times shrout$.
Profitability	Earnings before interest, taxes, depreciation, and amortization, measured as Compustat item $oibdp$, scaled by total assets, Compustat item at.
$\ln(B/M)$	Natural logarithm of the book-to-market ratio. Book equity is measured using the book value of common equity, Compustat item ceq, scaled by the market value of common equity, measured as $prcc_f \times csho$.
Cash	Cash and short-term investments, Compustat item che, scaled by total assets, Compustat item at.
Book leverage	Total long-term debt plus debt in current liabilities, Compustat items $dltt + dlc$, scaled by total assets, Compustat item at.
R&D intensity	Research and development expenses, Compustat item xrd, scaled by sales, Compustat item $sale$.
CapEx	Capital expenditures, Compustat item $capx$, scaled by sales, Compustat item $sale$.
Operating leverage	Cost of goods sold plus selling, general, and administrative expenses, Compustat items $cogs + xsga$, scaled by total assets, Compustat item at.

Table A2. PEPD Regressions Using Non-Overlapping Patent-Grant Events: KPSS Patent Value

Notes: This table compares the baseline PEPD regressions with horizon-specific non-overlapping-event regressions. Within each firm, patent-grant dates are ordered on the CRSP trading calendar. The first event is retained, and a subsequent event is retained only when it occurs more than h trading days after the previously retained event, where h is the PEPD horizon. Thus, retained observations for the same firm do not reuse the same post-grant return segment. PEPD is measured in percentage points. All regressions include patent count, lagged firm controls, and grant-year fixed effects. Standard errors, reported in parentheses, are clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	PEPD60		PEPD300		PEPD600		PEPD900	
	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap
Mean Probability (XGBoost)	-0.246	-0.360	-3.625**	-5.157*	-5.046*	-8.221	-9.704*	-9.788
	(0.400)	(0.696)	(1.534)	(2.698)	(2.811)	(5.472)	(4.963)	(9.962)
Breakthrough Innovation (XGBoost specification)	0.548***	0.903***	1.872***	5.141***	3.238***	9.321***	4.467**	3.955
	(0.130)	(0.213)	(0.562)	(0.938)	(1.044)	(1.775)	(1.746)	(3.146)
Mean Probability (LLM)	-0.685**	-0.618	-3.073***	-4.947***	-5.412**	-15.323***	-2.434	-6.804
	(0.281)	(0.415)	(1.182)	(1.786)	(2.171)	(3.522)	(3.621)	(5.890)
Breakthrough Innovation (LLM specification)	0.577***	0.965***	1.994***	5.513***	3.439***	9.098***	4.980***	4.769
	(0.130)	(0.213)	(0.561)	(0.935)	(1.041)	(1.745)	(1.748)	(3.132)
Observations (XGBoost)	541285	106920	536564	33596	523817	18813	305012	6621
Observations (LLM)	545790	108065	541061	34104	528257	18927	309485	6704

Table A3. PEPD Regressions Using Non-Overlapping Patent-Grant Events: Five-Year Forward Citations

Notes: This table repeats the non-overlapping-event design in Table A2 using five-year forward citations to define upper-tail patents. PEPD is measured in percentage points. All regressions include patent count, lagged firm controls, and grant-year fixed effects. Standard errors, reported in parentheses, are clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	PEPD60		PEPD300		PEPD600		PEPD900	
	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap
Mean Probability (XGBoost)	-0.131	-0.275	-3.168**	-4.546*	-4.269	-7.410	-9.118*	-9.375
	(0.401)	(0.698)	(1.536)	(2.700)	(2.819)	(5.484)	(4.954)	(9.975)
Breakthrough Innovation (XGBoost specification)	0.338***	0.456**	1.394***	3.839***	2.346***	4.435***	1.255	3.749
	(0.083)	(0.186)	(0.309)	(0.806)	(0.584)	(1.500)	(0.912)	(3.073)
Mean Probability (LLM)	-0.590**	-0.519	-2.666**	-4.041**	-4.736**	-14.515***	-2.131	-6.115
	(0.287)	(0.422)	(1.195)	(1.795)	(2.193)	(3.564)	(3.666)	(5.934)
Breakthrough Innovation (LLM specification)	0.317***	0.458**	1.336***	3.712***	2.226***	3.637**	1.390	3.696
	(0.083)	(0.188)	(0.304)	(0.802)	(0.575)	(1.503)	(0.900)	(3.087)
Observations (XGBoost)	541285	106920	536564	33596	523817	18813	305012	6621
Observations (LLM)	545790	108065	541061	34104	528257	18927	309485	6704

Table A4. PEPD Regressions Using Non-Overlapping Patent-Grant Events: Breakthrough Rank Index q_rank

Notes: This table repeats the non-overlapping-event design in Table A2 using the breakthrough rank index q_rank to define upper-tail patents. PEPD is measured in percentage points. All regressions include patent count, lagged firm controls, and grant-year fixed effects. Standard errors, reported in parentheses, are clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	PEPD60		PEPD300		PEPD600		PEPD900	
	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap	Full	Non-overlap
Mean Probability (XGBoost)	0.355	0.361	-1.253	-3.574	-1.509	-4.038	-5.278	-12.122
	(0.420)	(0.764)	(1.578)	(2.920)	(2.914)	(5.978)	(5.370)	(11.437)
Breakthrough Innovation (XGBoost specification)	0.111	0.250	0.884**	0.944	1.743**	4.308**	-0.645	0.319
	(0.096)	(0.210)	(0.370)	(0.973)	(0.755)	(1.799)	(1.347)	(3.019)
Mean Probability (LLM)	-0.358	-0.258	-1.753	-4.862**	-2.824	-13.160***	1.847	-2.286
	(0.291)	(0.460)	(1.207)	(1.993)	(2.230)	(3.952)	(3.784)	(6.733)
Breakthrough Innovation (LLM specification)	0.067	0.209	0.756**	0.481	1.524**	2.751	-0.302	0.428
	(0.094)	(0.214)	(0.354)	(0.999)	(0.727)	(1.852)	(1.299)	(3.185)
Observations (XGBoost)	464880	91005	460629	28537	451224	16078	240656	5021
Observations (LLM)	464880	91005	460629	28537	451224	16078	240656	5021

Table A5. Full Results: XGBoost Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns

Panel A: KPSS Grant-Window Value and Total Forward Citations

Notes: This table estimates $PEPD_{j,g}^{(h)} = \alpha + \beta AI_{j,g} + \gamma HighBenchmark_{j,g} + X'_{j,g-1}\delta + \lambda_{\tau(g)} + \varepsilon_{j,g}^{(h)}$. Columns (1)–(4) use KPSS grant-window value and columns (5)–(8) use total forward citations. An indicator equals one when the firm-date contains at least one patent above the full patent-sample 90th percentile of the corresponding benchmark. $AI_{j,g}$ is the firm-date mean XGBoost grant probability. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. Total citations use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	KPSS				Cites			
	PEPD60	PEPD300	PEPD600	PEPD900	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	-0.246	-3.625**	-5.046*	-9.704*	-0.179	-3.389**	-4.597	-9.279*
	(0.400)	(1.534)	(2.811)	(4.963)	(0.402)	(1.542)	(2.826)	(4.956)
Breakthrough Innovation	0.548***	1.872***	3.238***	4.467**	0.470***	1.665***	3.287***	1.873**
	(0.130)	(0.562)	(1.044)	(1.746)	(0.075)	(0.263)	(0.494)	(0.898)
Patent Count	0.009*	0.038	0.067	0.038	0.001	0.009	0.014	-0.006
	(0.006)	(0.027)	(0.052)	(0.075)	(0.005)	(0.025)	(0.047)	(0.069)
Log(Market Cap)	-0.384***	-1.696***	-3.337***	-2.539***	-0.351***	-1.585***	-3.159***	-2.186***
	(0.032)	(0.137)	(0.257)	(0.462)	(0.030)	(0.132)	(0.252)	(0.458)
Profitability	1.089**	6.463***	9.004***	18.397***	1.093**	6.480***	9.073***	18.022***
	(0.475)	(1.892)	(3.244)	(4.981)	(0.477)	(1.907)	(3.268)	(4.950)
Log(B/M)	0.141*	0.842***	1.926***	0.098	0.118	0.764**	1.798***	-0.105
	(0.074)	(0.315)	(0.584)	(0.984)	(0.073)	(0.313)	(0.583)	(0.976)
Book Leverage	0.297	3.709**	6.480**	12.337**	0.315	3.780**	6.667**	12.382**
	(0.378)	(1.601)	(3.051)	(5.224)	(0.381)	(1.611)	(3.065)	(5.245)
R&D	-0.001	0.005	-0.001	-0.006	-0.001	0.005	-0.001	-0.006
	(0.002)	(0.007)	(0.009)	(0.010)	(0.002)	(0.007)	(0.009)	(0.010)
CapEx	0.006	-0.012	0.011	0.023	0.006	-0.013	0.011	0.023
	(0.007)	(0.030)	(0.032)	(0.035)	(0.007)	(0.030)	(0.032)	(0.035)
Operating Leverage	0.360**	1.684**	2.624**	6.191***	0.370**	1.716**	2.680**	6.348***
	(0.150)	(0.668)	(1.298)	(2.280)	(0.152)	(0.675)	(1.312)	(2.299)
Cash	1.791***	7.656***	15.763***	14.891**	1.761***	7.544***	15.515***	14.943**
	(0.391)	(1.715)	(3.343)	(5.823)	(0.395)	(1.733)	(3.380)	(5.895)
Observations	541285	536564	523817	305012	541285	536564	523817	305012
Adj. R-squared	0.016	0.067	0.088	0.081	0.016	0.067	0.088	0.080

Table A5. Full Results: XGBoost Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns (continued)

Panel B: Forward Text Breakthrough Rank (q_rank)

Notes: The q_rank indicator equals one when a firm-date contains at least one patent above the full patent-sample 90th percentile of the forward-looking text-based rank. The AI variable is the firm-date mean XGBoost grant probability. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. q_rank uses post-grant text and is not contemporaneously observable. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) PEPD60	(2) PEPD300	(3) PEPD600	(4) PEPD900
Mean Probability (XGBoost)	0.355	-1.253	-1.509	-5.278
	(0.420)	(1.578)	(2.914)	(5.370)
Breakthrough Innovation (q_rank)	0.111	0.884**	1.743**	-0.645
	(0.096)	(0.370)	(0.755)	(1.347)
Patent Count	0.001	-0.001	-0.001	-0.029
	(0.005)	(0.023)	(0.044)	(0.062)
Log(Market Cap)	-0.354***	-1.624***	-3.238***	-1.795***
	(0.033)	(0.142)	(0.271)	(0.504)
Profitability	1.050**	6.346***	9.683***	20.648***
	(0.507)	(2.070)	(3.425)	(5.046)
Log(B/M)	0.208***	1.129***	2.107***	-0.557
	(0.080)	(0.344)	(0.648)	(1.132)
Book Leverage	0.224	3.738**	6.212*	10.212*
	(0.403)	(1.701)	(3.206)	(5.581)
R&D	-0.001	0.005	-0.003	-0.011
	(0.002)	(0.008)	(0.010)	(0.010)
CapEx	0.008	-0.011	0.024	0.046
	(0.007)	(0.033)	(0.034)	(0.033)
Operating Leverage	0.418**	1.944***	2.835**	6.465**
	(0.168)	(0.746)	(1.443)	(2.628)
Cash	1.966***	8.490***	16.224***	15.431**
	(0.414)	(1.775)	(3.461)	(5.998)
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table A6. Full Results: LLM-Embedding Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns

Panel A: KPSS Grant-Window Value and Total Forward Citations

Notes: This table repeats the specification in Table 4 using the firm-date mean LLM grant probability. Benchmark indicators use the full patent-sample 90th percentile of KPSS grant-window value or total forward citations. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. Total citations use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	KPSS				Cites			
	PEPD60	PEPD300	PEPD600	PEPD900	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.685**	-3.073***	-5.412**	-2.434	-0.582**	-2.707**	-4.680**	-2.184
	(0.281)	(1.182)	(2.171)	(3.621)	(0.287)	(1.199)	(2.198)	(3.663)
Breakthrough Innovation	0.577***	1.994***	3.439***	4.980***	0.448***	1.601***	3.153***	1.921**
	(0.130)	(0.561)	(1.041)	(1.748)	(0.075)	(0.260)	(0.487)	(0.880)
Patent Count	0.009	0.040	0.070	0.053	0.001	0.011	0.017	0.006
	(0.006)	(0.028)	(0.052)	(0.077)	(0.005)	(0.025)	(0.047)	(0.070)
Log(Market Cap)	-0.397***	-1.775***	-3.452***	-2.833***	-0.362***	-1.653***	-3.254***	-2.441***
	(0.031)	(0.136)	(0.255)	(0.457)	(0.030)	(0.131)	(0.250)	(0.453)
Profitability	1.096**	6.646***	9.170***	19.226***	1.100**	6.666***	9.246***	18.812***
	(0.473)	(1.887)	(3.239)	(4.999)	(0.475)	(1.901)	(3.263)	(4.965)
Log(B/M)	0.146**	0.865***	1.951***	0.086	0.121*	0.776**	1.803***	-0.143
	(0.074)	(0.315)	(0.581)	(0.981)	(0.073)	(0.312)	(0.580)	(0.974)
Book Leverage	0.381	4.013**	7.091**	12.693**	0.384	4.029**	7.170**	12.712**
	(0.379)	(1.612)	(3.071)	(5.238)	(0.381)	(1.619)	(3.080)	(5.257)
R&D	-0.001	0.005	-0.001	-0.007	-0.001	0.005	-0.000	-0.007
	(0.002)	(0.007)	(0.009)	(0.010)	(0.002)	(0.007)	(0.009)	(0.010)
CapEx	0.006	-0.013	0.010	0.024	0.006	-0.013	0.010	0.024
	(0.007)	(0.030)	(0.032)	(0.035)	(0.007)	(0.030)	(0.032)	(0.035)
Operating Leverage	0.362**	1.698**	2.617**	5.842***	0.370**	1.727**	2.668**	6.010***
	(0.149)	(0.663)	(1.288)	(2.261)	(0.151)	(0.672)	(1.304)	(2.283)
Cash	1.687***	7.376***	15.144***	14.486**	1.670***	7.313***	14.998***	14.550**
	(0.389)	(1.703)	(3.316)	(5.834)	(0.394)	(1.724)	(3.357)	(5.907)
Observations	545790	541061	528257	309485	545790	541061	528257	309485
Adj. R-squared	0.016	0.068	0.089	0.081	0.016	0.068	0.089	0.080

Table A6. Full Results: LLM-Embedding Grant-Probability Scores, Patent Benchmarks, and Post-Grant Returns (continued)

Panel B: Forward Text Breakthrough Rank (q_rank)

Notes: This table repeats Table 4 using the firm-date mean LLM grant probability. The q_rank indicator uses the full patent-sample 90th percentile. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. q_rank uses post-grant text and is not contemporaneously observable. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) PEPD60	(2) PEPD300	(3) PEPD600	(4) PEPD900
Mean Probability (LLM)	-0.358	-1.753	-2.824	1.847
	(0.291)	(1.207)	(2.230)	(3.784)
Breakthrough Innovation (q_rank)	0.067	0.756**	1.524**	-0.302
	(0.094)	(0.354)	(0.727)	(1.299)
Patent Count	0.000	-0.001	-0.001	-0.026
	(0.005)	(0.023)	(0.044)	(0.062)
Log(Market Cap)	-0.351***	-1.621***	-3.232***	-1.821***
	(0.033)	(0.142)	(0.271)	(0.506)
Profitability	1.017**	6.292***	9.576***	21.005***
	(0.507)	(2.068)	(3.418)	(5.047)
Log(B/M)	0.209***	1.147***	2.134***	-0.586
	(0.080)	(0.344)	(0.648)	(1.131)
Book Leverage	0.281	3.930**	6.534**	9.737*
	(0.408)	(1.720)	(3.238)	(5.623)
R&D	-0.001	0.005	-0.003	-0.012
	(0.002)	(0.008)	(0.010)	(0.010)
CapEx	0.008	-0.012	0.023	0.048
	(0.007)	(0.033)	(0.034)	(0.033)
Operating Leverage	0.420**	1.979***	2.887**	6.404**
	(0.169)	(0.746)	(1.444)	(2.649)
Cash	1.894***	8.329***	15.928***	15.985***
	(0.412)	(1.767)	(3.444)	(6.030)
Observations	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074

Table A7. Full Results: Grant-Probability Scores, Five-Year Citations, and Post-Grant Returns

Notes: The five-year citation indicator equals one when a firm-date contains at least one patent above the full patent-sample 90th percentile. Columns (1)–(4) use the firm-date mean XGBoost grant probability and columns (5)–(8) use the firm-date mean LLM grant probability. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. Five-year citations use post-grant information and are not contemporaneously observable. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	XGBoost				LLM			
	PEPD60	PEPD300	PEPD600	PEPD900	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability	-0.131	-3.168**	-4.269	-9.118*	-0.590**	-2.666**	-4.736**	-2.131
	(0.401)	(1.536)	(2.819)	(4.954)	(0.287)	(1.195)	(2.193)	(3.666)
Breakthrough Innovation (5-Year Cites)	0.338***	1.394***	2.346***	1.255	0.317***	1.336***	2.226***	1.390
	(0.083)	(0.309)	(0.584)	(0.912)	(0.083)	(0.304)	(0.575)	(0.900)
Patent Count	0.001	0.007	0.016	-0.006	0.001	0.010	0.018	0.005
	(0.005)	(0.025)	(0.047)	(0.068)	(0.005)	(0.025)	(0.047)	(0.069)
Log(Market Cap)	-0.352***	-1.596***	-3.162***	-2.183***	-0.362***	-1.663***	-3.256***	-2.441***
	(0.030)	(0.132)	(0.254)	(0.460)	(0.030)	(0.131)	(0.252)	(0.455)
Profitability	1.067**	6.391***	8.883***	17.962***	1.073**	6.574***	9.049***	18.751***
	(0.477)	(1.905)	(3.264)	(4.952)	(0.475)	(1.900)	(3.259)	(4.968)
Log(B/M)	0.107	0.722**	1.714***	-0.134	0.109	0.734**	1.722***	-0.174
	(0.073)	(0.313)	(0.583)	(0.978)	(0.073)	(0.312)	(0.580)	(0.975)
Book Leverage	0.275	3.656**	6.387**	12.233**	0.347	3.910**	6.914**	12.566**
	(0.380)	(1.608)	(3.060)	(5.238)	(0.381)	(1.616)	(3.076)	(5.251)
R&D	-0.001	0.005	-0.001	-0.006	-0.001	0.005	-0.001	-0.007
	(0.002)	(0.007)	(0.009)	(0.010)	(0.002)	(0.007)	(0.009)	(0.010)
CapEx	0.006	-0.012	0.011	0.023	0.006	-0.013	0.010	0.024
	(0.007)	(0.030)	(0.032)	(0.035)	(0.007)	(0.030)	(0.032)	(0.035)
Operating Leverage	0.364**	1.694**	2.644**	6.344***	0.365**	1.704**	2.632**	6.006***
	(0.152)	(0.675)	(1.312)	(2.302)	(0.151)	(0.672)	(1.304)	(2.285)
Cash	1.727***	7.372***	15.281***	14.804**	1.637***	7.144***	14.757***	14.384**
	(0.396)	(1.736)	(3.390)	(5.913)	(0.394)	(1.727)	(3.366)	(5.923)
Observations	541285	536564	523817	305012	545790	541061	528257	309485
Adj. R-squared	0.016	0.067	0.088	0.080	0.016	0.068	0.089	0.080

Table A8. Full Results: XGBoost Screening Scores with Multiple Patent Benchmarks

Notes: Columns (1)–(4) include full-sample upper-tail indicators for KPSS and q_rank; columns (5)–(8) include full-sample upper-tail indicators for five-year citations and q_rank. Each indicator equals one when a firm-date contains at least one patent above the corresponding full patent-sample 90th percentile. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects; standard errors are clustered by firm. Five-year citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	KPSS + q_rank				5-Year Cites + q_rank			
	PEPD60	PEPD300	PEPD600	PEPD900	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (XGBoost)	0.344	-1.284	-1.563	-5.406	0.453	-0.864	-0.870	-4.959
	(0.418)	(1.569)	(2.895)	(5.377)	(0.419)	(1.573)	(2.907)	(5.374)
Breakthrough Innovation (KPSS)	0.455***	1.380**	2.470**	2.250				
	(0.135)	(0.580)	(1.069)	(1.759)				
Breakthrough Innovation (5-Year Cites)					0.446***	1.756***	2.872***	1.986**
					(0.087)	(0.309)	(0.577)	(0.902)
Breakthrough Innovation (q_rank)	0.139	0.970***	1.898**	-0.452	0.022	0.536	1.175	-0.997
	(0.097)	(0.375)	(0.762)	(1.362)	(0.095)	(0.352)	(0.717)	(1.305)
Patent Count	0.004	0.008	0.016	-0.015	-0.003	-0.016	-0.026	-0.043
	(0.005)	(0.024)	(0.046)	(0.066)	(0.005)	(0.024)	(0.045)	(0.061)
Log(Market Cap)	-0.393***	-1.746***	-3.456***	-2.017***	-0.371***	-1.695***	-3.354***	-1.871***
	(0.035)	(0.147)	(0.276)	(0.513)	(0.033)	(0.142)	(0.273)	(0.505)
Profitability	1.075**	6.425***	9.823***	20.929***	1.057**	6.376***	9.736***	20.747***
	(0.506)	(2.065)	(3.416)	(5.078)	(0.509)	(2.083)	(3.443)	(5.042)
Log(B/M)	0.234***	1.210***	2.253***	-0.456	0.208***	1.128***	2.104***	-0.540
	(0.081)	(0.347)	(0.650)	(1.139)	(0.080)	(0.344)	(0.648)	(1.129)
Book Leverage	0.272	3.887**	6.483**	10.355*	0.268	3.915**	6.508**	10.423*
	(0.404)	(1.705)	(3.212)	(5.584)	(0.405)	(1.708)	(3.217)	(5.585)
R&D	-0.001	0.005	-0.003	-0.011	-0.001	0.005	-0.003	-0.011
	(0.002)	(0.008)	(0.010)	(0.010)	(0.002)	(0.008)	(0.010)	(0.010)
CapEx	0.008	-0.012	0.024	0.046	0.008	-0.012	0.023	0.046
	(0.007)	(0.033)	(0.034)	(0.033)	(0.007)	(0.033)	(0.034)	(0.033)
Operating Leverage	0.408**	1.915***	2.783*	6.368**	0.412**	1.923***	2.804*	6.485**
	(0.166)	(0.736)	(1.423)	(2.599)	(0.167)	(0.742)	(1.435)	(2.616)
Cash	1.934***	8.392***	16.048***	15.225**	1.852***	8.039***	15.482***	14.945**
	(0.408)	(1.754)	(3.419)	(5.935)	(0.412)	(1.772)	(3.460)	(6.014)
Observations	464880	460629	451224	240656	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074	0.016	0.070	0.091	0.074

Table A9. Full Results: LLM Screening Scores with Multiple Patent Benchmarks

Notes: This table repeats Table 7 using the firm-date mean LLM grant probability. Columns (1)–(4) include full-sample upper-tail indicators for KPSS and q_rank; columns (5)–(8) include full-sample upper-tail indicators for five-year citations and q_rank. PEPD is the post-grant cumulative size-adjusted return in percentage points. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects; standard errors are clustered by firm. Five-year citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	KPSS + q_rank				5-Year Cites + q_rank			
	PEPD60	PEPD300	PEPD600	PEPD900	PEPD60	PEPD300	PEPD600	PEPD900
Mean Probability (LLM)	-0.342	-1.703	-2.731	1.906	-0.235	-1.265	-2.029	2.228
	(0.287)	(1.192)	(2.202)	(3.761)	(0.292)	(1.203)	(2.225)	(3.784)
Breakthrough Innovation (KPSS)	0.453***	1.363**	2.443**	2.238				
	(0.135)	(0.578)	(1.065)	(1.759)				
Breakthrough Innovation (5-Year Cites)					0.432***	1.716***	2.803***	2.092**
					(0.088)	(0.307)	(0.573)	(0.897)
Breakthrough Innovation (q_rank)	0.097	0.847**	1.688**	-0.100	-0.010	0.451	1.025	-0.641
	(0.094)	(0.358)	(0.731)	(1.316)	(0.093)	(0.341)	(0.697)	(1.263)
Patent Count	0.003	0.008	0.015	-0.013	-0.004	-0.016	-0.026	-0.041
	(0.005)	(0.024)	(0.046)	(0.066)	(0.005)	(0.023)	(0.045)	(0.061)
Log(Market Cap)	-0.391***	-1.741***	-3.449***	-2.043***	-0.369***	-1.691***	-3.347***	-1.902***
	(0.035)	(0.147)	(0.276)	(0.514)	(0.033)	(0.143)	(0.274)	(0.506)
Profitability	1.044**	6.373***	9.721***	21.294***	1.028**	6.335***	9.650***	21.122***
	(0.506)	(2.063)	(3.410)	(5.079)	(0.508)	(2.081)	(3.436)	(5.041)
Log(B/M)	0.236***	1.226***	2.278***	-0.486	0.208***	1.141***	2.122***	-0.573
	(0.081)	(0.347)	(0.651)	(1.138)	(0.080)	(0.344)	(0.648)	(1.128)
Book Leverage	0.327	4.069**	6.789**	9.866*	0.311	4.050**	6.738**	9.914*
	(0.409)	(1.724)	(3.245)	(5.625)	(0.410)	(1.724)	(3.245)	(5.625)
R&D	-0.001	0.005	-0.003	-0.012	-0.001	0.005	-0.003	-0.012
	(0.002)	(0.008)	(0.010)	(0.010)	(0.002)	(0.008)	(0.010)	(0.010)
CapEx	0.008	-0.012	0.023	0.047	0.008	-0.012	0.023	0.048
	(0.007)	(0.033)	(0.034)	(0.033)	(0.007)	(0.033)	(0.034)	(0.033)
Operating Leverage	0.410**	1.950***	2.836**	6.306**	0.412**	1.949***	2.841**	6.412**
	(0.166)	(0.737)	(1.424)	(2.619)	(0.168)	(0.743)	(1.437)	(2.637)
Cash	1.865***	8.241***	15.771***	15.795***	1.794***	7.931***	15.274***	15.505**
	(0.407)	(1.747)	(3.405)	(5.970)	(0.411)	(1.766)	(3.449)	(6.047)
Observations	464880	460629	451224	240656	464880	460629	451224	240656
Adj. R-squared	0.016	0.070	0.091	0.074	0.016	0.070	0.091	0.074

Table A10. Full Results: XGBoost Screening Scores Around the GPT-3 Cutoff

Notes: This table reports descriptive slope heterogeneity around the paper's June 11, 2020 GPT-3 cutoff. Post equals one for grant dates on or after the cutoff; it is not a measure of investor AI adoption or an exogenous treatment. Columns (1)–(3) use KPSS, columns (4)–(6) use five-year citations, and columns (7)–(9) use q_rank. Each benchmark indicator is defined using its full patent-sample 90th percentile and is entered separately with its Post interaction. The AI variable is the firm-date mean XGBoost grant probability. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. Citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) KPSS-PEPD60	(2) KPSS-PEPD300	(3) KPSS-PEPD600	(4) Cite5-PEPD60	(5) Cite5-PEPD300	(6) Cite5-PEPD600	(7) q_rank-PEPD60	(8) q_rank-PEPD300	(9) q_rank-PEPD600
Mean Probability (XGBoost)	0.403	-0.659	0.741	0.443	-0.513	1.012	0.639	0.124	1.500
	(0.421)	(1.525)	(2.779)	(0.417)	(1.496)	(2.726)	(0.442)	(1.581)	(2.882)
Breakthrough Innovation (KPSS)	0.233*	0.281	0.435						
	(0.139)	(0.592)	(1.090)						
Breakthrough Innovation (5-Year Cites)				0.279***	1.241***	2.181***			
				(0.086)	(0.317)	(0.604)			
Breakthrough Innovation (q_rank)							0.079	0.656*	1.306*
							(0.101)	(0.378)	(0.774)
Mean Probability (XGBoost) × Post	-3.509***	-15.797***	-31.676***	-3.427***	-15.587***	-31.404***	-2.559**	-12.543***	-27.104***
	(0.999)	(3.873)	(7.348)	(1.029)	(4.020)	(7.579)	(1.172)	(4.490)	(7.912)
Breakthrough Innovation (KPSS) × Post	1.116***	5.564***	9.983***						
	(0.241)	(1.030)	(2.063)						
Breakthrough Innovation (5-Year Cites) × Post				0.633**	1.739*	2.165			
				(0.250)	(1.037)	(1.827)			
Breakthrough Innovation (q_rank) × Post							0.341	2.305**	4.525**
							(0.269)	(1.135)	(2.078)
Post	-1.832**	7.704***	20.602***	-1.884**	7.949***	21.380***	-2.325***	6.066*	18.347***
	(0.790)	(2.732)	(4.976)	(0.815)	(2.858)	(5.214)	(0.886)	(3.160)	(5.417)
Patent Count	0.009*	0.037	0.067	0.000	0.004	0.010	-0.001	-0.008	-0.015
	(0.006)	(0.027)	(0.051)	(0.005)	(0.024)	(0.046)	(0.005)	(0.022)	(0.042)
Log(Market Cap)	-0.381***	-1.683***	-3.315***	-0.351***	-1.595***	-3.163***	-0.352***	-1.616***	-3.222***
	(0.032)	(0.137)	(0.257)	(0.030)	(0.132)	(0.254)	(0.032)	(0.141)	(0.271)
Profitability	1.081**	6.433***	8.966***	1.068**	6.389***	8.874***	1.033**	6.259***	9.511***
	(0.474)	(1.872)	(3.211)	(0.477)	(1.904)	(3.259)	(0.506)	(2.056)	(3.398)
Log(B/M)	0.162**	0.944***	2.114***	0.118	0.767**	1.796***	0.211***	1.147***	2.144***
	(0.074)	(0.315)	(0.582)	(0.073)	(0.313)	(0.581)	(0.080)	(0.343)	(0.646)
Book Leverage	0.285	3.640**	6.341**	0.272	3.649**	6.388**	0.217	3.685**	6.100*
	(0.374)	(1.583)	(3.018)	(0.379)	(1.601)	(3.045)	(0.404)	(1.696)	(3.187)
R&D	-0.001	0.005	-0.000	-0.001	0.005	-0.001	-0.001	0.005	-0.003
	(0.002)	(0.007)	(0.009)	(0.002)	(0.007)	(0.009)	(0.002)	(0.008)	(0.010)
CapEx	0.006	-0.013	0.010	0.006	-0.013	0.010	0.008	-0.012	0.023
	(0.007)	(0.030)	(0.033)	(0.007)	(0.030)	(0.032)	(0.007)	(0.033)	(0.034)
Operating Leverage	0.368**	1.708**	2.660**	0.369**	1.705**	2.659**	0.424**	1.963***	2.865**
	(0.150)	(0.670)	(1.303)	(0.151)	(0.672)	(1.308)	(0.168)	(0.743)	(1.437)
Cash	1.810***	7.736***	15.896***	1.728***	7.392***	15.332***	1.968***	8.487***	16.207***
	(0.389)	(1.704)	(3.325)	(0.394)	(1.727)	(3.374)	(0.413)	(1.770)	(3.447)
Observations	541285	536564	523817	541285	536564	523817	464880	460629	451224
Adj. R-squared	0.016	0.068	0.089	0.016	0.067	0.089	0.017	0.070	0.091

Table A11. Full Results: LLM Screening Scores Around the GPT-3 Cutoff

Notes: This table repeats Table 8 using the firm-date mean LLM grant probability. Post equals one for grant dates on or after the paper's June 11, 2020 GPT-3 cutoff; it is not a measure of investor AI adoption or an exogenous treatment. Columns (1)–(3) use KPSS, columns (4)–(6) use five-year citations, and columns (7)–(9) use q_rank. Benchmark indicators use full patent-sample 90th-percentile thresholds. Regressions include patent count, lagged firm controls, and grant-year fixed effects, but not firm fixed effects. Standard errors are clustered by firm. Citations and q_rank use post-grant information, and the KPSS window may overlap PEPD on day $g+1$. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) KPSS-PEPD60	(2) KPSS-PEPD300	(3) KPSS-PEPD600	(4) Cite5-PEPD60	(5) Cite5-PEPD300	(6) Cite5-PEPD600	(7) q_rank-PEPD60	(8) q_rank-PEPD300	(9) q_rank-PEPD600
Mean Probability (LLM)	-0.278	-1.145	-2.038	-0.182	-0.671	-1.198	-0.204	-0.712	-1.166
	(0.295)	(1.168)	(2.162)	(0.294)	(1.143)	(2.107)	(0.305)	(1.202)	(2.213)
Breakthrough Innovation (KPSS)	0.272*	0.423	0.698						
	(0.139)	(0.594)	(1.093)						
Breakthrough Innovation (5-Year Cites)				0.270***	1.258***	2.180***			
				(0.085)	(0.310)	(0.588)			
Breakthrough Innovation (q_rank)							0.038	0.591	1.150
							(0.099)	(0.364)	(0.749)
Mean Probability (LLM) × Post	-1.638***	-7.710***	-13.714***	-1.809***	-8.692***	-15.555***	-1.005	-6.753**	-11.054**
	(0.602)	(2.448)	(4.646)	(0.638)	(2.626)	(4.947)	(0.695)	(2.710)	(4.845)
Breakthrough Innovation (KPSS) × Post	1.067***	5.421***	9.636***						
	(0.239)	(1.025)	(2.058)						
Breakthrough Innovation (5-Year Cites) × Post				0.643**	1.694	2.107			
				(0.252)	(1.039)	(1.832)			
Breakthrough Innovation (q_rank) × Post							0.372	2.310**	4.640**
							(0.268)	(1.123)	(2.060)
Post	-3.056***	2.414	8.829***	-2.958***	3.398*	10.893***	-3.351***	2.221	7.782**
	(0.611)	(1.905)	(3.216)	(0.634)	(2.037)	(3.493)	(0.648)	(2.063)	(3.406)
Patent Count	0.009*	0.040	0.069	-0.000	0.005	0.011	-0.001	-0.009	-0.016
	(0.006)	(0.027)	(0.051)	(0.005)	(0.024)	(0.046)	(0.005)	(0.022)	(0.042)
Log(Market Cap)	-0.395***	-1.762***	-3.431***	-0.363***	-1.668***	-3.266***	-0.350***	-1.617***	-3.223***
	(0.031)	(0.137)	(0.256)	(0.030)	(0.132)	(0.252)	(0.033)	(0.141)	(0.271)
Profitability	1.102**	6.687***	9.260***	1.090**	6.652***	9.186***	1.007**	6.248***	9.482***
	(0.472)	(1.869)	(3.216)	(0.475)	(1.901)	(3.262)	(0.506)	(2.057)	(3.402)
Log(B/M)	0.167**	0.964***	2.124***	0.122*	0.785**	1.806***	0.212***	1.168***	2.167***
	(0.074)	(0.315)	(0.582)	(0.073)	(0.313)	(0.580)	(0.080)	(0.344)	(0.648)
Book Leverage	0.341	3.814**	6.724**	0.318	3.782**	6.693**	0.268	3.838**	6.361**
	(0.377)	(1.598)	(3.047)	(0.380)	(1.612)	(3.067)	(0.408)	(1.714)	(3.222)
R&D	-0.001	0.005	-0.000	-0.001	0.005	-0.000	-0.001	0.005	-0.003
	(0.002)	(0.007)	(0.009)	(0.002)	(0.007)	(0.009)	(0.002)	(0.008)	(0.010)
CapEx	0.006	-0.013	0.009	0.006	-0.013	0.010	0.008	-0.012	0.022
	(0.007)	(0.030)	(0.032)	(0.007)	(0.030)	(0.032)	(0.007)	(0.033)	(0.034)
Operating Leverage	0.367**	1.711**	2.632**	0.368**	1.708**	2.632**	0.425**	1.994***	2.908**
	(0.149)	(0.667)	(1.296)	(0.150)	(0.670)	(1.302)	(0.168)	(0.744)	(1.439)
Cash	1.702***	7.445***	15.240***	1.630***	7.136***	14.740***	1.893***	8.313***	15.879***
	(0.387)	(1.696)	(3.304)	(0.393)	(1.721)	(3.357)	(0.412)	(1.763)	(3.433)
Observations	545790	541061	528257	545790	541061	528257	464880	460629	451224
Adj. R-squared	0.016	0.068	0.090	0.016	0.068	0.089	0.017	0.070	0.091

Characteristics and Factor Exposures of Retrospective Portfolios

Table A12 compares the retrospective portfolios with all patenting firm-months. Portfolio firms are larger and substantially more patent intensive: average log market capitalization ranges from 8.63 to 10.56, versus 7.42 in the benchmark, and average patent counts range from about 201 to 317, versus 79. Recent returns, momentum, and profitability also differ. The classification therefore concentrates in a nonrandom subset of patent owners with characteristics that can themselves predict returns. These imbalances motivate the factor regressions and caution against attributing raw portfolio performance to patent underrecognition alone.

Table A12. Portfolio Characteristics of High-Quality Low-AI Firms

Notes: This table reports winsorized characteristics of the patenting firm-month sample and the six high objective-quality low-AI portfolios over the broader 2010–2024 portfolio-formation window. A firm enters a portfolio in month t if it has at least one patent issued in the prior 12 months that is in the top decile of the relevant objective-quality measure and in the bottom 30% of the AI-score distribution. Characteristics are winsorized at the 1st and 99th percentiles, first averaged across firms within each month, and then averaged over months, matching the equal-weighted portfolio construction. REV, MOM, and Turnover are reported in percentage points. Turnover is monthly trading volume divided by shares outstanding.

	All patenting firms	q_rank XGBoost	q_rank LLM	KPSS XGBoost	KPSS LLM	Five-Year Citations XGBoost	Five-Year Citations LLM
Log(Market Cap)	7.415	8.845	8.627	10.490	10.561	8.651	8.660
Log(B/M)	-1.155	-1.114	-1.132	-1.647	-1.745	-1.204	-1.212
Profitability	-0.003	0.107	0.098	0.130	0.134	0.079	0.090
Book Leverage	0.188	0.177	0.161	0.241	0.224	0.174	0.162
R&D	2.626	0.365	0.189	0.301	0.160	0.824	0.635
CapEx	0.118	0.061	0.056	0.067	0.055	0.074	0.066
Operating Leverage	0.718	0.645	0.650	0.580	0.564	0.647	0.626
Cash	0.316	0.268	0.286	0.232	0.256	0.294	0.298
Patent Count	78.768	316.796	305.553	200.836	229.323	254.382	281.510
REV	0.861	1.177	1.149	1.339	1.310	1.403	1.376
MOM	12.827	15.461	15.329	21.493	21.444	16.896	16.114
Turnover	22.936	23.471	24.030	21.401	22.133	23.765	24.401
Months	180	119	119	179	179	131	131
Median firms	948	187	191	122	87	262	202
Firm-months	169711	22142	21582	21219	15432	34084	26397

Raw Returns Across Holding Horizons

Table A13 reports equal-weighted raw buy-and-hold returns over one, three, six, and twelve months. In the 2010–2024 window, one-month returns range from 1.25% to 1.53%, and twelve-month returns range from 14.00% to 18.81%. In the later period, twelve-month returns are 13.32% and 14.97% for the q_rank classifications and 32.05% and 34.27% for the five-year-citation classifications. The increase with the horizon partly reflects mechanical compounding and changing holding-period coverage.

Because membership uses completed annual thresholds, grant-window returns, or post-grant benchmark information, these are historical validation returns rather than implementable strategy returns. They connect the patent-level classification to firm performance, but the full-sample factor tests show that positive Fama intercepts become small and statistically insignificant under HXZ q5.

Table A13. Forward Buy-and-Hold Returns for High-Quality Low-AI Firms

Notes: This table reports equal-weighted forward buy-and-hold returns for high objective-quality low-AI portfolios. Panel A uses the broader 2010–2024 portfolio-formation window. Panel B restricts the sample to post-Transformer holding months beginning in 2018-09. A firm enters a portfolio in month t if it has at least one patent issued in the prior 12 months that is in the objective-quality top decile and AI-score bottom 30%. Returns are cumulative raw stock returns over the next 1, 3, 6, or 12 months and are reported in percentage points. Parentheses report Newey-West t-statistics, using lag $\max(6,h)$ for horizon h. The months column reports valid monthly observations for the 1/3/6/12-month horizons; counts differ across columns because patent-quality measures have different coverage and portfolios require at least five firms.

Objective quality	AI score	1-month	3-month	6-month	12-month	Months	Median firms
Panel A. Broader 2010–2024 portfolio window							
q_rank	XGBoost	1.294***	3.570***	6.778***	14.004***	119/119/119/119	187
		(3.68)	(3.74)	(3.86)	(4.57)		
q_rank	LLM	1.250***	3.544***	6.775***	14.197***	119/119/119/119	190
		(3.52)	(3.77)	(4.05)	(4.71)		
KPSS	XGBoost	1.262***	3.682***	7.384***	15.665***	179/179/179/179	122
		(4.11)	(4.33)	(4.74)	(5.06)		
KPSS	LLM	1.258***	3.649***	7.229***	15.261***	179/179/179/179	87
		(3.90)	(4.09)	(4.58)	(4.89)		
Five-Year Citations	XGBoost	1.528***	4.499***	9.048***	18.807***	131/131/131/131	262
		(4.07)	(4.25)	(4.50)	(4.43)		
Five-Year Citations	LLM	1.520***	4.494***	8.942***	18.572***	131/131/131/131	201
		(4.12)	(4.31)	(4.52)	(4.27)		
Panel B. Post-Transformer holding months beginning 2018-09							
q_rank	XGBoost	0.911	2.564*	4.109**	13.320***	16/16/16/16	170
		(0.97)	(1.67)	(2.14)	(4.56)		
q_rank	LLM	0.819	2.572*	4.432***	14.965***	16/16/16/16	114
		(0.88)	(1.72)	(3.96)	(3.51)		
KPSS	XGBoost	1.202**	3.751**	7.562***	16.418***	76/76/76/76	104
		(2.11)	(2.38)	(2.61)	(2.75)		
KPSS	LLM	1.288**	3.947**	7.738**	16.669***	76/76/76/76	83
		(2.04)	(2.26)	(2.51)	(2.62)		
Five-Year Citations	XGBoost	1.747*	6.241**	14.074***	32.052***	28/28/28/28	255
		(1.69)	(2.22)	(2.68)	(3.12)		
Five-Year Citations	LLM	1.827*	6.491**	14.754***	34.266***	28/28/28/28	192
		(1.77)	(2.26)	(2.69)	(2.82)		