Taking Stock: Do Equity-Based Managerial Incentives Drive Corporate Innovation?

Mark A. Chen* Erol Oytun Ercan† Yaswanth Pothuru‡

August 2026

Abstract

We provide causal evidence on the effects of equity-based CEO incentive pay on corporate innovation activity. To capture exogenous variation in equity-based pay, we exploit a large sample of public-company relative performance share plans under which realized equity payouts vary discontinuously around performance cutoffs. Using a regression discontinuity research design, we find that CEO stock award payouts lead to significant increases in a firm's patenting activity. Moreover, these stock award payouts increase the novelty, quality, and value of a firm's patents. We also document that payout-induced increases in innovation activity are more pronounced in the presence of managerial short-termism or restrictions on selling, hedging, or pledging of shares. Overall, our findings show that equity-based pay for top management is an important driver of corporate innovation.

JEL Classification: G30, G32, G34, J33, O31, O32

Keywords: CEO compensation, equity-based pay, regression discontinuity (RD), performance-based share awards, causal effects, innovation, patents, R&D, managerial short-termism, share pledging

*J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30303 USA. Email: machen@gsu.edu

†J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30303 USA. Email: eercan@gsu.edu

‡J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30303 USA. Email: ypothuru@gsu.edu

1 Introduction

Equity-based forms of incentive pay, such as stock and option awards, play a central role in the compensation of CEOs of public U.S. companies. By aligning the interests of top management with those of shareholders, equity-based pay can overcome agency problems and increase firm value (Jensen and Meckling 1976; Diamond and Verrecchia 1982; Fama and Jensen 1983). Theoretical arguments suggest that equity pay can be especially useful for incentivizing top management when investment projects involve a high risk of failure (Manso 2011) or are long-term in nature (Holmström and Tirole 1993).

While the theoretical effects of stock-based CEO compensation are relatively well-understood, much less is known about how this type of compensation affects corporate outcomes in practice. A key reason for the lack of direct empirical evidence on the effects of stock-based pay is the endogenous nature of compensation contracts. Indeed, compensation contracts are the result of the decisions of directors, executives, and consultants in response to many observable and unobservable characteristics. Thus, any study of the effects of stock-based pay must contend with the fact that compensation contracts are correlated with unobservable firm, industry, and CEO characteristics that may themselves influence firm outcomes.

In this paper, we provide causal empirical evidence on the effects of CEO equity pay on long-term corporate investment. We focus on corporate investment in the form of innovation activity, which is recognized as a key driver of economic growth (Schumpeter 1942; Solow 1956; Romer 1990), firm value (Jaffe 1986; Hall et al. 2005), and competitive advantage (Aghion and Howitt 1992). To the extent that corporate innovation involves risky, long-term investments, examining the causal link between equity pay and innovation can shed light on whether stock-based incentives do in fact align CEO interests with long-term value creation as theory suggests should be the case.

Our identification strategy relies on a particular type of long-term incentive pay that has become the predominant form of executive compensation at large U.S. public companies, namely,

performance share plans.[1] In the typical performance share plan, an executive receives an award of stock upfront, but the actual number of shares eventually paid out can scale up or down (e.g., between 0% and 200% of the initial award) as a function of how well the firm reaches one or more pre-specified performance goals. Most plans assess performance in terms of the firm's stock returns, accounting performance, or both, and performance is almost always measured over a fixed period of time (e.g., three years after the grant date). Also, a critical feature for our purposes is that performance share plans often specify two or three performance cutoff levels ex ante, and the quantity of stock eventually paid out to an executive can exhibit a discrete jump if the firm barely meets a performance goal versus barely missing it.

We exploit the cutoff rules in a sample of performance share plans from 1998 to 2024 to implement a regression discontinuity (RD) design that estimates the causal effect of stock awards on various innovation outcomes. Our sample consists of share plans with performance goals tied to well-defined stock return metrics, such as market-adjusted or peer-group-adjusted Total Shareholder Return (TSR). By focusing on share awards linked to stock return performance rather than accounting performance, our empirical design greatly mitigates the concern that boards or managers can precisely manipulate performance around key cutoff levels.[2] In our sample, meeting a stock performance goal translates into a sizeable and economically meaningful jump in the realized amount of the stock award. Indeed, barely passing a goal versus barely missing it translates on

[1]The economic importance of share based incentives has increased dramatically compared to option based incentives. Among the first 100 S&P 500 proxy filers in 2025, performance-based equity accounted for 60% of the average long-term incentive (LTI) mix, compared with 24% for time-based RSUs and 16% for stock options. Thus, performance-based equity represented 2.5 times the share of RSUs and 3.75 times the share of options, and exceeded the two other LTI vehicles combined by 50%. The median value of CEO long-term incentives also increased from $11.72 million in 2023 to $12.49 million in 2024, an increase of approximately 7%. The latter amount is equivalent to about 71% of median total CEO compensation of $17.68 million. See Pearl Meyer (2025). More recent evidence shows a further increase in the dollar importance of equity pay: Equilar reports that median S&P 500 CEO stock awards rose 11.5% in 2025 to $10.9 million, equivalent to approximately 62% of median total CEO compensation of $17.7 million. See Equilar and Associated Press (2026).

[2]Accounting-based performance measures are potentially subject to both accrual and real manipulation around salient thresholds (Healy 1985; Holthausen et al. 1995; Burgstahler and Dichev 1997; Degeorge et al. 1999; Graham et al. 2005; Roychowdhury 2006). Most directly in the executive-compensation setting, Bennett et al. (2017) document bunching just above contractual performance goals, particularly earnings goals, together with higher abnormal accruals and reductions in R&D or SG&A among firms that narrowly exceed them.

average into a 10 percent increase in the CEO's stock wealth,based on their pre-event beneficial ownership.

Our RD analysis reveals that CEO stock awards have positive effects on various measures of innovation activity. Using local linear regressions with a triangular kernel function and a 7.5% bandwidth, we find that winning a stock award leads to a significantly greater volume of patent applications and patent grants. For instance, successfully meeting a performance goal and obtaining the share payout causes the firm's patent grants and patent applications to increase by 70% and 62%, respectively, during the subsequent three years. Although there is no evidence that realized stock awards significantly affect R&D spending, we find that they do result in increased breadth and originality of the firm's granted patents. Furthermore, stock award payouts cause significant increases in the total market value and forward citation counts of a firm's new patents over the subsequent three years.

In order for our RD design to be valid, variation in treatment status (i.e., whether or not a stock award payout is high or low) should be effectively random in a neighborhood of each cutoff threshold. This will hold true if individuals do not precisely manipulate stock performance around the cutoff levels (Lee and Lemieux 2010). To check the plausibility of this condition, we conduct tests on the density of the assignment variable (stock performance). Both McCrary (2008) tests and Cattaneo et al. (2018) tests fail to detect density discontinuities at cutoff levels, thus supporting our premise of no manipulation. Furthermore, we verify that the other covariates have smooth distributions around thresholds, which helps to rule out the possibility that the observed jumps in innovation outcomes are driven by jumps in the covariates.

We also examine whether our main results continue to hold under alternative regression specifications, subsamples, and estimation methods. Robustness tests show that our estimates are qualitatively unchanged when we use different kernel functions, bandwidths, and bias corrections. Furthermore, we estimate RD regressions after excluding any award events for which it is unclear whether the firm interpolates share payouts when setting the CEO incentive scheme. We find that our qualitative results continue to hold when we use this more restrictive, no-interpolation subsample.

Next, to gain a better understanding of the nature of incentive effects from stock awards, we examine how a CEO's short-term or long-term orientation at the time of a realized stock award moderates the award's impact. If stock awards foster corporate innovation mainly by overcoming managerial myopia, then the impact of a stock award should be larger when the CEO previously exhibited a short-term orientation. To examine this idea, we condition on several empirical measures of CEO short-termism, including measures based on the tenure of the CEO, the relative delta sensitivity or value from vested and unvested equity holdings, and a short-termism score constructed by applying a Large Language Model to the Management's Discussion and Analysis (MD&A) in annual reports. Subsample RD regressions reveal that stock awards generally have positive effects on innovation, but only when the CEO has a high degree of short-term orientation. Thus, we conclude that stock awards are particularly useful for addressing incentive problems that stem from managerial short-termism.

In further tests, we examine the role of ex post managerial flexibility in shaping the incentive effects of realized share awards. We hypothesize that share awards are most effective at driving long-term innovation when the CEO is at least somewhat restricted in their ex post ability to sell, hedge, or pledge their awarded shares. To explore this hypothesis, we assemble data on key firm-level policies that are in place at the time that stock awards are paid out. These firm-level policies include the following: (1) stock ownership guidelines; (2) anti-hedging restrictions; and (3) prohibitions against the pledging of shares as collateral for personal loans. We find that, when such restrictions are present, stock awards continue to significantly boost innovation activity. However, when restrictions on ex post CEO behavior are absent, stock awards are generally ineffective at driving innovation. Overall, these findings suggest that share payouts alone may be insufficient for motivating a CEO to innovate. Instead, the combination of share payouts with mechanisms that encourage or commit the CEO to hold them is what is likely to be most effective at driving innovation.

We contribute to the literature that examines the implications of equity-based CEO pay for firm behavior. Prior studies in this area focus mainly on the link between equity-based CEO

compensation and corporate risk-taking. For example, a large literature (e.g., Coles et al. 2006; Tchistyi et al. 2011) measures overall CEO wealth vega, or sensitivity to stock volatility, and how vega affects risk-taking through investment and financial policies.[3] Other work focuses specifically on executive stock options, using the option expensing mandate brought about by FAS 123-R as a quasi-natural experiment. Studies in this vein find that executive stock options have ambiguous effects on risk-taking (Hayes et al. 2012) and patenting (Mao and Zhang 2018; Biggerstaff et al. 2019) but appear to increase new product development (Faurel et al. 2024). Another line of research examines vesting restrictions tied to CEO equity, showing that managers exhibit greater short-term orientation in their investment policies when options and shares approach pre-arranged vesting dates (Edmans et al. 2017a) or when executive options have accelerated vesting (Ladika and Sautner 2020). Our paper differs from these studies in that we focus on the effects of exogenous changes in CEO stockholdings, rather than changes in stock option portfolios or equity vesting provisions. To the best of our knowledge, we are the first to study the causal effects of discontinuous CEO stock awards on firms' innovation and investment decisions.[4]

Our paper is also related to work that explores important design features of CEO equity compensation, including ex ante mechanisms that link the vesting or granting of equity awards to performance. For instance, a stream of research examines the characteristics and determinants of performance-vesting features in equity-based awards (Carter et al. 2009; Bettis et al. 2010) as well as the impact of performance-vesting awards on the delta and vega sensitivity of executive compensation (Bettis et al. 2018). Other papers exploring performance-based awards in CEO compensation contracts include De Angelis and Grinstein (2015), Pawliczek (2021), and Core and Packard (2022). Brisley (2006) theoretically studies the risk-taking effects and incentive benefits of option awards that vest progressively as a function of stock price performance. Shue and Townsend (2017) exploit institutional features in multi-year option compensation plans to show that increases

[3] Papers that examine the association between CEO stock and option holdings and changes in firm risk include Agrawal and Mandelker (1987) and Core et al. (1999), among many others.

[4] Gao (2026) studies a sample of discontinuities in CEO performance pay and shows that narrowly missing a relative performance goal leads to an increase in CEO insider trading profits. While our study focuses exclusively on goals tied to stock price performance, most of the compensation goals in Gao (2026) are tied to accounting performance, which is a priori much more susceptible to precise manipulation around performance cutoffs.

in stock options lead to increased equity volatility. Bennett et al. (2017) study the use of accounting performance goals in executive compensation and document that firms often exceed goals by a small margin, leading to higher abnormal accounting accruals and lower R&D and SG&A expenditures. In theoretical work, Manso (2011) analyzes how stock options with long vesting provisions can be part of the optimal contract for motivating corporate innovation. Unlike much of this literature that studies the ex ante incentives provided by specialized vesting and granting features, we exploit these features to capture capture exogenous ex post variation in CEO equity holdings. In doing so, our work complements the existing literature by offering researchers and practitioners direct, credible evidence on the broader implications of CEO stock awards, which in recent years have eclipsed stock option awards as the single largest source of CEO equity pay among public U.S. companies.

The remainder of the paper is organized as follows. Section 2 describes our data sources and provides descriptive statistics on firm and CEO characteristics, equity awards, performance metrics, and innovation outcomes. Section 3 outlines our RD methodology and our regression specifications. We present the main empirical results in Section 4. Section 5 details the results of additional analyses on managerial short-termism and firm-level restrictions on the ex post selling, hedging, or pledging of shares. Section 6 concludes.

2 Data

Our analysis relies on publicly available data on compensation plans, contractual performance rules, realized stock performance, and subsequent innovation outcomes. We retrieve these data from several primary sources. Parts of our data on executive compensation plans come from ISS Incentive Lab, which codes the terms of individual equity and cash awards to named executive officers of large U.S. issuers from proxy statements and related filings over fiscal years 1998 through 2024, including structured fields and analyst method notes that describe each award's contractual performance goal. Because those fields are incomplete or missing altogether for many awards, we extend the data with the DEF 14A proxy statements themselves, retrieved from the SEC's EDGAR

system. The same filings, together with Form 10-K reports, also supply the governance provisions and managerial disclosures we classify. Realized stock performance for issuers and their benchmark peers comes from CRSP daily returns, and firm fundamentals come from Compustat. Innovation outcomes combine USPTO patent-application records with the extended patent-level data of Kogan et al. (2017), which supply patent grants, forward citations, and market-implied patent values. We supplement these sources with Form 144 notices for proposed insider sales, Federal Reserve interest rates for CEO portfolio calculations, and the Revelio Labs corporate hierarchy for entity resolution in the patent match. In the remainder of this section, we describe these sources, the methods we use to merge them, and the large language model (LLM) and agentic artificial intelligence (AI) procedures we use to extract variables from unstructured textual disclosures.

2.1 Executive compensation

2.1.1 Stock awards

To construct a sample of performance-vesting share plans, we start with the full Incentive Lab award universe.[5] The unit of observation throughout the paper is an award: one canonical grant identifier, with its performance metric, performance period, payout schedule, and vesting terms.

From the full universe of 778,858 awards at 4,802 firms, we first retain equity awards, then awards whose vesting or payout depends on performance or accelerated-vesting conditions, and then the restricted stock unit subset of those awards, which yields 172,916 RSU performance awards. Of these, 79,856 specified payout based on performance relative to a peer group or index rather than on an absolute goal, and 71,060 use a stock-return-based performance measure such as total shareholder return, stock return, stock price, or market value. Requiring a one- or two-performance-group structure that the payout calculation can represent leaves 25,211 award events. Further restricting attention to awards held by the CEO produces a design universe of 5,683 award events at 1,059

[5] Incentive Lab collects grant-level compensation data for the named executive officers of approximately the largest 750–1,000 U.S. public firms by market capitalization in each year, back-filling firms that enter the top group so that coverage includes the S&P 500 and much of the broader large-cap universe; see Bettis et al. (2018) for a detailed description of the database.

firms. Panel A of Table 1 reports each filtering step; the rows are nested, and the firm column counts the distinct issuers represented at each stage.

2.1.2 Contractual performance goals and realized stock performance

Evaluating each award requires the ex ante contractual performance goal and realized performance measured against that goal. The Incentive Lab data include basic structural aspects of stock awards, including the performance metric underlying the share award; the threshold, target, and maximum payout goals; which companies serve as peer companies; and free-text descriptions of how performance comparisons are calculated, recorded in analyst comments and comparison-method fields. However, some data needed to calculate the performance comparison are only partly available or missing altogether. Examples of incomplete or missing fields in Incentive Lab include the return basis, averaging windows, annualization scheme, the exact benchmark used, and the rank or percentile convention. Even when we restrict attention to awards with stock-based performance rules, there is considerable variation in whether the metric is cumulative TSR, annualized TSR, or an average stock price; whether the benchmark is the S&P 500 or a bespoke peer group of public companies; and whether the measurement window is a few months or several years.

We convert this combined evidence into contract-specific performance cutoffs in three steps. First, an LLM interpretation stage first converts the assembled Incentive Lab and proxy method text into structured calculation fields. Second, an agentic AI workflow then standardizes each rule into an executable calculation. Third, it computes the issuer's and each disclosed peer's total shareholder return over the contractual window from CRSP, applies the contract's rank or percentile convention, and expresses the threshold, target, and maximum cutoffs on the same scale as realized performance, so that the distance between the two is well defined for every award. Internet Appendix Section IA.1 documents the extraction, the standardization, and the treatment of incomplete peer data in detail.

Panel B of Table 1 shows caharacteristics of the resulting sample. Combining each rule with realized returns yields 1,380 unique awards with at least one calculable cutoff. The empirical sample then filters observations by bandwidth and retains exactly one cutoff per award, under a fixed priority

that takes the target cutoff first, then the threshold, then the maximum. Applying this rule within a symmetric window of ± 7.5 percentage points, the maintained bandwidth throughout the paper, yields 431 awards from 183 firms, with performance realized between 2005 and 2022; narrower and wider windows yield 139 awards at ± 2.5 points, 278 at ± 5 points, and 534 at ± 10 points. Within the preferred sample, 246 observations are centered at a target cutoff, 105 at a threshold, and 80 at a maximum, and the running variable is measured exactly for 256 observations and within an identified interval for the remaining 175.

2.2 Firm, CEO, and performance characteristics

Table 2 reports descriptive statistics for the preferred sample in three panels: firm characteristics, CEO characteristics, and performance and assignment characteristics. Panel A covers firm characteristics. We merge the award sample with CRSP and Compustat by firm identifier to obtain size, profitability, leverage, cash holdings, valuation, prior stock returns, prior idiosyncratic volatility, and pre-period R&D intensity, each measured before the performance realization. We also use LLM classification of each firm's proxy statements to extract governance provisions that shape what a CEO can do with earned shares: 65.3% of matched awards are at firms whose latest pre-event DEF 14A discloses a restriction on pledging company stock, 79.8% disclose an anti-hedging policy, 39.7% disclose a CEO stock-ownership or holding guideline, and 25.1% give the compensation committee discretion over stock-award payouts.[6] Log assets, the prior 12-month return, and prior idiosyncratic volatility enter the RD regressions as controls. Other variables, including ROA, leverage, cash/assets, market-to-book, and pre-period R&D intensity, are used to validate RD design but are not used in the RD regressions.

Panel B covers CEO characteristics. We measure CEO tenure, beneficially owned shares, directly priceable common-stock wealth, and scheduled vesting from awards other than the focal one, each dated before the performance realization, from Incentive Lab and the underlying filings.

[6]Each filing is submitted independently to the language model under a purpose-built prompt restricted to the provision in question, and the model returns the classification together with the supporting filing language. Internet Appendix Section IA.3.2 documents the pledging-restriction and anti-hedging classifications, and Section IA.3.1 documents the committee-discretion classification; Appendix Table A.1 defines the holding-guideline coding.

Pre-event common-stock wealth averages $39.2 million against a median of $15.4 million, the right skew typical of executive wealth. Realized vesting flows during the realization year and the year after it are scaled by these predetermined positions. The cumulative vesting-share ratio, which divides cumulative vesting shares by pre-event shareholdings, averages 0.391, and the corresponding value-based intensity averages 0.442. We again use LLMs to construct text-based CEO measures. A managerial-myopia score assigned to each firm's MD&A disclosure on an ordered one-to-five scale, measured in the fiscal year before performance realization, and the pledging status and pledged-share amounts of the CEO extracted from beneficial-ownership disclosures.[7] Finally, we collect the CEO's proposed insider sales from Form 144 notices filed over the post-realization window.

Panel C covers performance and assignment characteristics. For each award we compute realized performance against the extracted contractual rule using CRSP returns for the issuer and for the disclosed peer group or index over the contractual measurement window, apply the contract's rank or percentile convention, and measure the signed distance between realized relative performance and the selected payout cutoff. The resulting distance is the running variable; it averages -0.661 percentage points within the maintained window, and 41.8% of awards clear their cutoff. The award jump variable, defined as the treatment indicator times the contract-implied change in share entitlement at the cutoff scaled by pre-event CEO stock wealth, has an unconditional mean of 0.032.

Measuring realized relative performance requires the returns of every disclosed peer, and identifying peer groups is a nontrivial measurement problem. Some disclosed peers cannot be matched to return data because they are foreign, private, delisted, or identified ambiguously. Where every constituent return is observed, realized relative performance is identified as a point, and we label the observation exact. Where a subset of the peer group cannot be matched, the firm's rank is bounded but not pinned down. We store the identified interval and record the observation at the interval's numerical upper endpoint under a fixed convention, never describing it as an exact

[7] Internet Appendix Section IA.3.3 documents the myopia score and Section IA.3.2 the CEO pledging extraction. The vested and unvested components of each CEO's option and stock portfolio, used in Section 5.1, are constructed from Incentive Lab holdings tables with Black–Scholes–Merton sensitivities as documented in Internet Appendix Section IA.4.

realization.[8] Similarly, contracts that disclose interpolation between cutoffs generate a smaller entitlement change at the cutoff than step-schedule contracts, and interpolation status enters the estimation as a fixed effect. Section 4.2 shows that the results are robust to dropping every interval-identified observation and to excluding contracts that disclose interpolation. All continuously valued analysis variables are winsorized at the 1st and 99th percentiles of the pre-2023 performance-realization universe.

2.3 Innovation activity

The innovation data combine USPTO pre-grant patent-application records with the extended patent-level data of Kogan et al. (2017), hereafter KPSS, which link granted patents to public firms and supply, for each patent, its forward citations and its market-implied economic value. We aggregate outcomes over the years following each award's performance realization, using horizons of $t+1$, $t+1$ through $t+2$, and $t+1$ through $t+3$, where t is the performance-realization year.

Patent applications present a linkage problem that granted patents do not. Whereas the KPSS data already identify the firm behind each granted patent, application records identify assignees by name only, and no common identifier connects USPTO assignees to CRSP/Compustat firms. This entity-resolution problem is one manifestation of the broader measurement difficulties that arise when patent records are aggregated to the firm level (Lerner and Seru 2022). Applications are frequently filed under a subsidiary or acquired entity whose legal name bears little resemblance to that of the listed parent, so exact name matching misses valid links, while fuzzy string matching at any fixed similarity threshold both accepts false matches, because a high similarity score is not proof of common economic identity, and rejects true ones, because a subsidiary's name may share almost nothing with its parent's. We therefore link application assignees to firms with an agentic AI matching procedure that separates candidate generation from identity adjudication. A fuzzy stage proposes ranked candidates from the CRSP/Compustat universe and the Revelio Labs corporate

[8] Internet Appendix Section IA.1.3 details the peer matching, the interval construction, and the conditions under which an award enters the calculable-cutoff panel.

hierarchy, exact matches are accepted directly, and language-model agents adjudicate every non-perfect candidate using corporate-family evidence, with unresolved cases left unmatched rather than forced. This matching applies to patent applications only. For applications that eventually grant, the resulting links agree with the independently constructed KPSS firm assignments approximately 98.6% of the time. Internet Appendix Section IA.2 describes the procedure in detail.

We measure innovation along several dimensions, following the patent literature (Hall et al. 2001; Bernstein 2015). The most basic measures are counts: patent grants and patent applications, summed over the horizon and transformed as $\log(1+x)$. Both series are dated by application filing year, so the grant count for a year is the number of applications filed in that year that had issued by the data cutoff. The grant series is right-truncated for the latest cohorts, whose applications have had the least time to issue, while the application series is complete for every cohort. We report both and treat their agreement as the more informative fact.

Since raw counts cannot distinguish valuable patents from marginal ones, we next measure the economic value of patenting using the real economic value of KPSS, which prices each patent from the issuing firm's stock-market reaction to the grant announcement, and its scientific impact using forward citations, whose link to firm value Hall et al. (2005) establish. Beyond quantity and impact, we examine the character of the underlying research. Following Hall et al. (2001) and a large subsequent finance literature that uses citation dispersion to characterize the technological nature and breadth of corporate innovation (Lerner et al. 2011; Bernstein 2015; Brav et al. 2018), we measure adjusted patent originality. Originality captures the breadth of technological knowledge on which an invention builds: patents that draw their backward citations from a wider range of technology fields receive higher originality scores. Our measure is constructed as an application-support-weighted patent attribute and incorporates the finite-sample adjustment for citation-based Herfindahl measures (Hall et al. 2001). We complement originality with technological breadth, defined as the number of distinct CPC subclasses spanned by the firm's patenting in a given year. Finally, we measure the observable innovation input, R&D intensity, formed annually as R&D expenditure over assets, averaged across the available post years, and transformed as $\log(1+x)$.

Table 3 reports the distribution of the innovation outcomes. They are heavily right-skewed with a large mass of structural zeros. The median award in this local sample is followed by no patent grants, applications, real economic value, or citations over the three-year horizon. The discontinuities we report in real economic value and forward citations are identified mainly from a right tail and we weigh them accordingly in Section 4. The attribute measures are defined only for firm-awards with sufficient patent-level support, which is why adjusted originality is available for only 106 observations against 431 for the count and value outcomes.

3 Empirical Design

3.1 Discontinuities in stock award payouts

Identifying the effect of equity incentives on innovation is challenging because equity is not assigned at random. Boards decide how much equity to grant while they are forming views about the firm's investment opportunities, so the size of a CEO's equity stake and the firm's innovative prospects are determined together, and a correlation between the two could reflect the board's anticipation of those prospects, the firm's life cycle, or persistent differences across firms rather than any effect of the equity itself (Edmans et al. 2017b). Economy-wide shocks to compensation practice move much else at the same time, and the evidence they have produced remains contested (Mao and Zhang 2018; Biggerstaff et al. 2019). What identification requires is variation in a CEO's equity that is economically large, that is fixed before the board can observe the outcomes it might respond to, and that arrives at firms which are otherwise comparable. Performance cutoffs based on stock returns provide exactly the type of variation needed for it.

Consider two firms whose CEOs hold awards with the same peer group and the same three-year measurement period, and whose realized relative returns end the period a quarter of a percentage point apart on either side of the target. Nothing about their realized relative performance differs in any economically meaningful sense, and five or seven percentage points of relative return over three years is noise from the perspective of any theory of managerial effort or ability. The contract,

written three years earlier, converts that difference into a discrete difference in shares earned. Since the contract was fixed before performance was observable, which side of the cutoff a firm lands on is as good as random in the immediate neighborhood of the cutoff, and comparing outcomes across the two sides isolates the consequence of the equity increment.

Figure 1 illustrates the contractual structure that generates the discontinuity. A typical award specifies not one goal but several: a threshold below which the performance-based portion does not vest, a target, and a maximum that caps the payout, with the number of shares earned increasing as realized performance crosses successive contractual performance levels. In a hypothetical reconstructed schedule, performance below threshold earns nothing, reaching threshold earns 30% of the target number of shares, reaching target earns 100%, and reaching maximum earns 200%.

Apple Inc. provides a particularly transparent example of a discrete relative total shareholder return (rTSR) contract. Tim Cook's 2011 restricted stock unit (RSU) award, as modified in 2013, conditioned a portion of vesting on Apple's TSR relative to companies that remained in the S&P 500 throughout the applicable performance period. Performance in the bottom third of the comparison group resulted in 0% of the performance-based RSUs vesting, performance in the middle third resulted in 50% vesting, and performance in the top third resulted in 100% vesting. Thus, crossing one of these contractual boundaries changed the number of shares earned discretely even when the difference in underlying relative performance was arbitrarily small. For the three-year performance period from August 25, 2015 through August 24, 2018, 280,000 of Cook's performance-based RSUs were subject to this Relative-TSR condition. Apple's realized Relative TSR placed it at the 88th percentile of the eligible S&P 500 firms, and all 280,000 performance-based RSUs consequently vested on August 24, 2018.[9]

[9]Apple Inc., *Notice of 2019 Annual Meeting of Shareholders and Proxy Statement*, Schedule 14A, made available to shareholders January 8, 2019, p. 33-34. The proxy states that Mr. Cook's 2011 RSU Award, as modified in 2013, contained specified vesting of 0%, 50%, and 100% for performance in the bottom, middle, and top thirds, respectively, of the eligible S&P 500 comparison group. For the August 25, 2015–August 24, 2018 performance period, 280,000 RSUs were subject to the condition; Apple's Relative TSR ranked at the 88th percentile and all 280,000 RSUs vested. Available from the U.S. Securities and Exchange Commission, https://www.sec.gov/Archives/edgar/data/320193/000119312519004664/d667873ddef14a.htm.

Another important feature of our RD design is that realized performance is measured by a stock-return metric, which is less subject to manipulation than the accounting metrics that threshold designs in compensation settings must otherwise confront. A manager who is measured on earnings has reporting discretion over earnings and real discretion over the decisions that produce them. Bennett et al. (2017) document bunching around accounting-based compensation goals, Graham et al. (2005) report survey evidence that executives sacrifice real value to meet earnings benchmarks, and Dechow and Sloan (1991) show that managers cut R&D when their horizons shorten. A relative stock-return goal is a much harder object to hit precisely. The final quantity is the firm's rank against peers or an index whose returns are realized simultaneously, so landing on a specific rank requires forecasting other firms' outcomes as well as one's own, and the residual noise in relative returns is large compared with the precision that sorting across a cutoff would require. Malenko and Shen (2016) exploit the same property in a different setting, using a firm's TSR relative to its industry median as the forcing variable in a regression-discontinuity design around ISS say-on-pay recommendations.[10] Managers retain discretion over disclosure timing, repurchases, and other price-relevant actions, and Bloomfield et al. (2024) show that relative performance evaluation can induce strategic disclosure aimed at peers. Our specifications therefore control for the firm's prior 12-month return, which absorbs pre-realization price effects of this kind, and Section 3.3 tests the continuity conditions directly.

3.2 Empirical specification

Our empirical strategy compares innovation outcomes at firms whose realized performance fell just short of a contractual cutoff with outcomes at firms whose performance just cleared one. We implement the comparison with a sharp regression-discontinuity design (Imbens and Lemieux 2008; Lee and Lemieux 2010). Local linear regressions estimated separately on the two sides of the cutoff within a symmetric bandwidth, triangular kernel weights that give the awards closest

[10] Gao (2026) studies insider trading around narrowly missed compensation goals in a sample that pools relative stock-return and relative accounting goals. Because our sample is restricted to stock-return cutoffs, the reporting discretion that makes an accounting goal a difficult running variable does not enter, and the two designs isolate different margins. Flammer and Bansal (2017) exploits close votes on shareholder proposals in a related discontinuity spirit.

to the cutoff the greatest influence, fixed effects that absorb predetermined cohort, industry, and contract-type variation, and standard errors clustered by firm. The coefficient of interest is the jump in the conditional mean of the outcome at the cutoff.

We note three key aspects of how we implement the RD design. First, we estimate the discontinuity with local linear regression inside a bandwidth rather than with a high-order global polynomial. Global polynomials of high degree place substantial weight on observations far from the cutoff and can generate spurious discontinuities from the shape of the fitted curve rather than from the data near zero. The methodological literature is close to unanimous in recommending low-order local polynomials with a bandwidth and kernel instead (Imbens and Lemieux 2008; Lee and Lemieux 2010; Gelman and Imbens 2019; Calonico et al. 2014). Second, we allow separate slopes on the two sides of the cutoff, because constraining the slope to be equal forces any change in the local relationship between relative performance and the outcome into the estimated level jump, which is precisely the contamination a discontinuity design is meant to avoid (Lee and Lemieux 2010). Third, we weight observations with a triangular kernel, the weighting under which local linear estimation has optimal boundary properties (Calonico et al. 2014). In Section 4.2, we show robustness of our main results to variations along all three of these design dimensions.

Formally, for grant g at firm i, let X_{ig} be the Relative-TSR distance from the selected contractual cutoff, measured in percentage points and centered at zero, and let $D_{ig} = \mathbf{1}[X_{ig} \geq 0]$ indicate that the cutoff was cleared. The design is a sharp regression discontinuity in which the contract maps the sign of X_{ig} into the payout entitlement deterministically, so treatment status is a known function of the running variable rather than a probabilistic one. The maintained specification is

$$Y_{ig} = \alpha + \tau D_{ig} + \beta_L X_{ig} + \beta_\Delta (D_{ig} \times X_{ig}) + \Gamma' Z_{ig} + \lambda_t + \lambda_j + \lambda_c + \lambda_s + \varepsilon_{ig}, \tag{1}$$

where Y_{ig} is the outcome aggregated over the stated post-realization horizon, τ is the discontinuity in the conditional mean of Y at $X = 0$, β_L is the slope below the cutoff, and $\beta_L + \beta_\Delta$ is the slope above it. Z_{ig} contains log assets, the prior 12-month stock return, and prior idiosyncratic volatility,

together with the documented outcome-specific predetermined controls. The four fixed effects are λ_t for the performance-realization year, λ_j for the FF12 industry, λ_c for the cutoff type, and λ_s for the interpolation status of the contract. Elements of Z_{ig} and the fixed effects are measured prior to the conclusion of the performance period. ROA, leverage, cash/assets, and market-to-book never enter Equation (1). We estimate the equation by triangular-kernel-weighted least squares within $h = 0.075$ and cluster standard errors by firm, because the 431 awards come from 183 issuers.

The identifying assumption is that the conditional expectations of potential outcomes are continuous in X at zero. Because D is a deterministic function of X, the assumption cannot be tested directly, but it has two testable implications: the density of the running variable should not jump at a cutoff, and predetermined characteristics should not jump at a cutoff. In the next section, we examine each of these implications in detail.

3.3 Validity

3.3.1 No manipulation of the running variable

The continuity assumption would fail if CEOs or boards could place awards precisely on the favorable side of a cutoff, and the standard diagnostic is whether the density of the running variable jumps at zero (McCrary 2008). Figure 2 shows the distribution of the running variable within ± 10 percentage points, separating exact observations from interval upper endpoints. The distribution is not smooth in the way a continuous economic variable would be, because measured relative performance inherits mass points from rank-based contracts. When a contract ranks the firm against a peer group of a few dozen firms, feasible ranks are discrete and distances from a cutoff cluster accordingly. This is a property of the contracts rather than an artifact of measurement, and it is the reason the formal density tests use a mass-point-adjusted estimator. Nevertheless, as seen in the figure, there is no pile-up of observations just above zero relative to just below it.

Figure 3 reports two formal tests on the maintained sample. Panel A is a McCrary-style diagnostic that fits separate local-quadratic curves to binned log densities inside the analysis

window; the estimated log-density jump is -0.630 with a p-value of 0.259. Panel B reports the local-polynomial density estimator of Cattaneo et al. (2020) with the manipulation test of Cattaneo et al. (2018), run with its mass-point adjustment; at MSE-optimal side-specific bandwidths the test statistic is -0.090 with a p-value of 0.928. Appendix Table A.4 reports the complete prespecified sensitivity grid across seven specifications that vary the estimator, bin width, fitting window, and polynomial order, every test statistic is negative and none rejects continuity at the ten-percent level, with p-values between 0.105 and 0.928. The uniformly negative point estimates indicate a density that is slightly lower just above the cutoff than just below it, the opposite of what sorting into the favorable side would produce, and also the opposite of what the interval upper-endpoint convention could generate mechanically, since that convention can move observations toward the cutoff only from below. The tests are run on the selected sample, so they bound sorting conditional on the one-cutoff-per-award selector. Together with the economic argument that a relative stock-return is extremely difficult to manipulate to within a fraction of a percentage point, the density evidence supports the assumption that awards do not sort across the cutoff.

3.3.2 Continuity in predetermined covariates

If awards near the cutoff are locally randomized, firms on the two sides should be observably similar before treatment. Table 4 tests this implication directly. Each row is a separate triangular-kernel local-linear RD within $h = 0.075$ containing only the treatment indicator, the running variable, and their interaction, so each coefficient asks whether the named predetermined characteristic itself changes discretely at the cutoff. The nine covariates span firm size, profitability, leverage, cash holdings, valuation, the prior 12-month return, prior idiosyncratic volatility, pre-period R&D intensity, and pre-period patent applications, the last of which is the most demanding covariate for an innovation design because it tests whether firms that were already patenting more sort above the cutoff.

None of the nine covariates shows a significant discontinuity. The smallest p-value is 0.148, for cash/assets, and the pre-period patenting discontinuity is an essentially zero point estimate,

at -0.060 with a standard error of 0.437. A joint Wald test that all nine discontinuities are zero, estimated in one stacked regression after standardizing each covariate, returns $\chi^2(9) = 7.953$ with a p-value of 0.539, so the covariates neither individually nor jointly move at the cutoff. The individual tests are not uniformly precise, and three point estimates are not negligible relative to the covariates' own dispersion, but their direction is mixed and the joint test comfortably fails to reject. The balance evidence is thus consistent with the local comparability that the design requires. Awards just above and just below the cutoff belong to firms that look alike on every predetermined dimension we can measure, including their prior innovation. Section 4.2 complements this evidence with placebo cutoffs at which the innovation outcomes themselves show no discontinuity.

3.3.3 Discontinuities in stock awards

Our basic premise is that clearing a performance cutoff changes the CEO's equity, and this premise is directly observable. Figure 4 presents the evidence graphically, plotting raw binned means with separate local-linear fits, so the compensation discontinuity is visible before any model is imposed. Panel A plots Award Jump, the contract-implied entitlement change scaled by pre-event CEO stock wealth; it is zero below the cutoff by construction, so the panel displays the size of the assigned entitlement change, a step of roughly ten percent of the CEO's pre-award stock wealth at the cutoff. Panel B plots the Cumulative Vesting Share Ratio, whose control side is not mechanically zero because CEOs who narrowly miss still earn shares under interpolation provisions and other awards; the visible step at zero shows that the equity CEOs actually receive, and not merely their entitlement, jumps at the cutoff.

Table 5 quantifies both margins across the different specifications. Panel A reports the full-specification Award Jump coefficient of 0.085, with estimates between 0.085 and 0.099 across the range of specifications. It records that clearing a cutoff assigns a block of stock worth approximately 8.5% of the value of the CEO's pre-award common-stock holdings. Ceteris paribus, a single award event moves the CEO's stock ownership by nearly a tenth. Panels B and C report the realized vesting response. Cumulative Vesting Intensity, which divides vesting value over t and $t+1$ by

pre-event stock wealth, rises by 0.244 against a control-side mean of 0.437, and the Cumulative Vesting Share Ratio rises by 0.221 against a control-side mean of 0.391. Both are increases of roughly 56–57% of the average level among comparable CEOs below the cutoff, and both remain distinguishable from zero throughout the range of sepcifications. The formal robust bias-corrected estimator in Appendix Table A.1 confirms the pattern, with 0.090 for Award Jump, significant at the one-percent level, and 0.246 for the share ratio, significant at the ten-percent level, while the value-based intensity measure retains its magnitude at 0.255 without conventional significance; we lean on the share ratio because it does not depend on the price at which pre-event wealth is valued.

The three panels rest on 221, 338, and 356 awards against the 414 used for the innovation outcomes, because the pre-event wealth proxy requires a recent beneficial-ownership snapshot with a matched price and the vesting ratios require a strictly positive denominator. The compensation and innovation samples therefore overlap but are not identical, and we do not scale the innovation estimates by the equity response. The earned shares that vest may in principle be sold, a possibility Section 5.3 examines directly and finds no evidence for. The evidence establishes the validity condition that matters most for interpretation, in which narrowly clearing a Relative-TSR cutoff produces an economically large and statistically significant increase in the CEO's stock awards.

4 Main Empirical Results

4.1 Effect of stock awards on innovation

4.1.1 Main innovation outcomes

We first present the graphical evidence and then the regression estimates for the main innovation outcomes. Figure 5 plots the four principal innovation outcomes, the logs of one plus patent grants, patent applications, real economic value, and forward citations over the three post-realization years, against the running variable, with covariate-adjusted binned means and separate local-linear fits on the two sides of the cutoff. Visual inspection reveals a discontinuity in all four panels: each outcome

shifts up discretely at zero, and the fitted level changes equal the full-specification coefficients reported below. Given the identification assumption and the validity evidence in Section 3.3, we attribute these discontinuities to the causal effect of the additional stock awards earned at the cutoff. For comparison, no similar discontinuity appears at placebo cutoffs where no contractual payout changes (Section 4.2).

Table 6 shows how patent counts and innovation inputs respond to realized stock awards. In terms of economic magnitude of effect, the full specification implies that clearing a cutoff raises $\log(1+\text{patent grants})$ by 0.528 and $\log(1+\text{patent applications})$ by 0.481. Converted to percentage terms, firms whose CEOs narrowly clear a cutoff produce approximately 70% more patent grants and 62% more patent applications over the following three years than firms that narrowly miss. The table also shows that the response builds over time: for patent grants, the coefficient on barely passing a cutoff is 0.352 at $t+1$, 0.451 over two years, and 0.528 over three, with applications following a similar time-horizon profile at 0.320, 0.400, and 0.481. The fact that the estimates grow with the horizon is consistent with the idea that changes in innovation policy take time to manifest as patent grants and applications.

We note that these estimated magnitudes require careful interpretation because the outcome is $\log(1+y)$ and the median stock award in the sample is followed by no patenting. The exponentiated coefficient describes a proportional shift across the local population of awards; it is driven by the firms that have positive patenting, and the estimand is a local comparison in a 431-award window. Note that the observed response τ is the raw effect of clearing a cutoff rather than a response scaled by the size of the equity change, so we do not read it as an elasticity of patenting with respect to CEO equity. Nonetheless, given sample means of 0.537 and 0.647 for the transformed outcomes, it is evident that the estimated impacts of stock awards are large and economically meaningful.

The positive and significant estimates of the effects on innovation are consistent with the predictions of long-horizon incentive theories. Manso (2011) shows that motivating exploration requires tolerance for early failure and reward for eventual success, and that equity-like claims with long horizons dominate short-term performance pay for that purpose; Holmström (1989) makes

the complementary argument that innovation is unusually costly to motivate with conventional incentive instruments. Empirically, Lerner and Wulf (2007) find that innovative output responds to how incentives are structured within firms, Chang et al. (2015) tie broad-based non-executive employee options to patenting, and Mao and Zhang (2018) link CEO option-portfolio sensitivities to patenting, though Biggerstaff et al. (2019) report no innovation response to the same FAS 123R shock once patent truncation is addressed. It is important to note that our tests do not require a complete understanding of why a firm initially awarded the performance shares that it did; we are more interested in the impact of ex post changes in equity. However, it must be acknowledged that clearing a cutoff could convey information about the firm's competitive position, or the board could use ex post discretion to adjust a CEO's compensation. Evidence from the Section 5.2 sample split regressions does not support this alternative explanations.

Unlike innovation output, R&D intensity does not appear to respond to stock award realizations. The discontinuity in $\log(1 + \text{mean R\&D intensity})$ is 0.001 in the maintained specification, small at every horizon, and never distinguishable from zero, while the formal robust bias-corrected estimator in Appendix Table A.1 returns 0.004, marginally significant and close to the 0.005 sample mean of the transformed outcome. What accounts for the different effects on patenting and R&D intensity? R&D expenditure is an accounting aggregate set through budgeting processes, committed in advance, and reported annually; patents reflect which projects were pursued, continued, and prosecuted within whatever budget exists. March (1991) frames the relevant margin as exploration versus exploitation, Katila and Ahuja (2002) find that the scope of search across new knowledge domains and the depth of reuse are complements, each related non-monotonically to new-product output, and Lerner and Wulf (2007) document that the productivity of a research organization depends on how its incentives are arranged. Overall, the regression results in Table 6 suggest that an equity shock to the CEO may change project selection, continuation decisions, risk tolerance, or organizational attention, but it does so without changing the budget line, which is consistent with the breadth response that we document in the next section.

4.1.2 Technological scope

In Table 7, we examine how CEO equity shocks affect the nature of the firm's subsequent patenting. First, we examine technological breadth, which measures how widely the firm's patenting spans distinct CPC subclasses. The regression estimates show that breadth rises by 4.195 over the three-year horizon, and the horizon block is nearly flat, at 3.773 at $t+1$ and 4.179 at $t+2$, each significant. All three cells are on an annual scale, so the flat profile indicates a persistent annual widening of the firm's technological footprint rather than simply deepening innovation in existing technology areas. The table also shows that adjusted patent originality rises by 0.462 at the three-year horizon on 91 observations, having been 0.213 and insignificant before fixed effects, and by 0.501 at $t+2$ on 65.

These results on innovation scope connect the patent volume results to the theory of managerial horizons. In March (1991), a manager under short-horizon pressure exploits existing knowledge, while a manager with a claim on distant outcomes explores new combinations. Likewise, Manso (2011) shows that tolerance for early failure is what makes exploration incentive-compatible; and Katila and Ahuja (2002) document that searching across new knowledge domains is the dimension of R&D behavior most tightly linked to new-product output. An increase in breadth and originality of patenting following an increment of long-dated equity is consistent with the general implications of these studies.

4.1.3 Patent value and quality

Given that stock awards drive innovation volume, a natural question to ask is whether the increased innovation also leads to higher innovation quality. To address this question, we consider patent citations and patents' market values, which have become the standard metrics for gauging innovation quality in the literature (Hall et al. 2005; Kogan et al. 2017; Bernstein 2015). If the additional patenting at the cutoff consisted of marginal filings, counts would rise while value and impact remain unchanged. Table 8 shows that the opposite is true. Indeed, clearing a cutoff raises the log of one plus market-implied patent value by 0.853 and the log of one plus forward citations by 0.453, both on 414 observations: the implied value of the firm's patenting more than doubles, a 135%

increase, and forward citations rise by approximately 57%. Against sample means of 0.643 and 0.274, these responses are larger relative to the outcome scale than either of the patent-count results. The specification profile matches the count results, with raw estimates of 0.485 and 0.327 that are positive but imprecise, rising and tightening once year and industry effects are absorbed.

These outcomes address the possibility that equity awards motivate a CEO to generate more innovation output but with patents of lower value. The real economic value measure of Kogan et al. (2017) prices each patent from the issuing firm's stock-market reaction to the grant announcement, so it aggregates investors' assessment of the patent's economic significance, and forward citations capture the patent's influence on subsequent technology, the margin Hall et al. (2005) show is priced into firm value. Positive discontinuities in both, alongside the breadth and originality results, indicate that the equity increment shifts firms toward innovation that both the market and later inventors judge more significant. The pattern is again what exploration theories predict: projects selected for long-run payoff carry higher value when they succeed (Manso 2011).

Of course, the measures by themselves do not completely prove value creation (the value measure inherits what the market believed at grant), but the general agreement between implied market value and the citation-based impact is the strongest evidence available at this level of aggregation. Notably, the estimates across the horizons are 0.649, 0.731, and 0.853 for value and 0.426, 0.459, and 0.453 for citations, so value accumulates with the horizon in the way a genuine flow response would while citations are flat after $t+2$. Applications carry no truncation at all and track the grant estimate closely at every horizon and in every robustness column, which constitutes evidence that truncation is not generating the count results.

4.2 Robustness

4.2.1 Alternative bandwidths

Table 9 re-estimates the four principal outcomes at symmetric bandwidths from 5% to 10%, reapplying the one-cutoff-per-award selector inside each window so that the sample construction

is internally consistent at every column. All four outcomes are positive at every bandwidth. Most of the estimates are significant. Indeed, 20 out of 28 estimates are significant at the five-percent level, while 26 out of 28 are significant at the ten-percent level; the two insignificant cells are patent applications at the narrowest and the widest window.

The patent-grant discontinuity declines from 0.532 at 5% to 0.361 at 10%, peaking at 0.570 at 7%, with the maintained 7.5% estimate of 0.528 just below that peak. Real economic value runs from 0.684 at 10% to 0.889 at 7%, and citations from 0.305 to 0.495 across the same columns. The gradient is mild and has the sign one expects: point estimates rise slightly from 5% to 7% and then decline modestly at the wider windows, which is what happens when observations further from the cutoff, where the local linear approximation is worse and the contractual comparison is less tight, are given weight. The results are therefore not an artifact of the maintained bandwidth choice.

4.2.2 Functional form

Panel A of Table 10 varies the kernel and the local polynomial order on the preferred sample. Across triangular, uniform, and Epanechnikov kernels at local linear order, the patent-grant estimate ranges from 0.384 to 0.528, real economic value from 0.773 to 0.875, and citations from 0.319 to 0.453. Moving to local quadratic order raises eleven of the twelve kernel-by-outcome estimates, with patent grants between 0.761 and 0.904 and citations between 0.663 and 0.764, at some cost in precision as the additional flexibility must be paid for. The most flexible specifications, local quadratic at any kernel, produce the largest estimates, and the uniform-kernel local-linear specification, which gives observations at the edge of the window the same weight as observations at the cutoff, produces the smallest. Every one of the twenty-four kernel-by-order estimates is positive and significant at least at the ten-percent level, so the sign and economic magnitude of the results are robust to the choice of kernel and polynomial order.

4.2.3 Running-variable certainty and interpolation

Panel B changes the sample rather than the estimator, and it addresses the two measurement features described in Section 2.2. Excluding the awards that disclose interpolation leaves 311 observations and raises the estimates, to 0.631 for patent grants, 0.559 for applications, 0.990 for real economic value, and 0.471 for citations. This is the direction the contractual logic predicts: a schedule that interpolates between cutoffs generates a smaller level change at the cutoff than a step schedule, so pooling the two should attenuate the estimate, and removing the disclosed-interpolation awards should raise it. The headline results are therefore not generated by the interpolating contracts.

Restricting to the 240 complete cases whose running variable is exactly identified, which drops the interval-endpoint convention entirely, gives 0.451 for patent grants, 0.876 for real economic value, and 0.486 for citations, each remaining significant, while patent applications fall to 0.430 and lose significance. The applications result is a loss of precision rather than a sign reversal: the point estimate falls by about a tenth on 58% of the estimation sample. This is the measurement check we regard as most important, because it removes any dependence on where within its identified interval an observation is placed, and it indicates that the point convention is not generating the results.

4.2.4 Placebo cutoffs and falsification tests

If the innovation discontinuities reflect the contractual payout change rather than some smooth feature of the relationship between relative performance and patenting, they should appear only where a payout actually changes. Appendix Table A.5 reports two falsification designs. The first fixes a fake cutoff at $+5$ percentage points, where no contract in the sample assigns a payout change, and re-estimates the full specification within a symmetric five-point window on one side of the true cutoff; none of the four principal innovation outcomes shows a discontinuity, with p-values between 0.311 and 0.912. The second draws a placebo cutoff at random from a prespecified grid constructed so that every candidate window lies strictly on one side of the true cutoff; at the realized draw of $+8.5$ percentage points, all four estimates are again small and indistinguishable from zero. The seed, grid, bandwidth, and outcome set were fixed before estimation. Together with the density and

covariate evidence in Section 3.3, these tests locate the discontinuity at the contractual cutoff and nowhere else we looked.

4.2.5 Predetermined vesting and wealth controls

Appendix Table A.3 adds predetermined scheduled vesting from awards other than the focal one, and pre-event CEO stock wealth, to the maintained specification. This addresses a specific alternative: that what we attribute to the focal award reflects other equity vesting on a schedule that happens to correlate with the performance period, in the spirit of the vesting-based identification in Edmans et al. (2017a) and Ladika and Sautner (2020). Adding scheduled nonfocal vesting gives 0.533 for patent grants and 1.020 for real economic value on 363 observations; adding pre-event stock wealth gives 0.486 and 0.876 on 348; adding both gives 0.526 and 1.074 on 303. The grant, value, and citation results are unaffected in sign and magnitude. Patent applications lose significance when scheduled nonfocal vesting enters, as the complete-case requirement removes between 12 and 27 percent of the sample, but remain significant when wealth alone is added. The specification is fixed in advance and its stability indicates that the focal award, rather than coincident vesting from other grants, drives the results.

5 Additional Analysis

5.1 Managerial short-termism

A central theme in the compensation literature is managerial short-termism: managers under pressure to deliver near-term results underinvest in projects whose payoffs are further in the future (Stein 1989; Graham et al. 2005). Equity that vests in the near term intensifies that pressure (Gopalan et al. 2014; Edmans et al. 2017a; Jochem et al. 2018; Ladika and Sautner 2020), while long-dated equity works against it. If the innovation response we document operates mainly by lengthening the CEO's effective horizon, it should be strongest where short-horizon pressures are otherwise most binding and where an additional block of long-dated stock represents a larger change in

the CEO's long-horizon exposure. This subsection examines several empirical proxies for CEO short-termism, including the CEO's position in the career cycle, structural aspects of the CEO's existing equity portfolio, and a text-based measure of the short-term orientation expressed in the firm's own MD&A disclosures; the stock-ownership requirements that operate on the same margin are examined together with the governance policies in Section 5.3. Table 11 reports split-sample estimates for each of the four principal innovation outcomes.

Panel A splits on CEO tenure at the sample median of four years. The discontinuity in $\log(1+$ patent grants$)$ is 0.075 for short-tenure CEOs and 0.839 for long-tenure CEOs, on 200 and 213 observations, and for real economic value it is 0.126 against 1.259; all four long-tenure estimates are positive and significant while all four short-tenure estimates are near zero. Tenure proxies for firm-specific knowledge, accumulated authority over the research organization, and the ability to redirect projects (Barker and Mueller 2002), so the concentration of the response among long-tenure CEOs is consistent with the idea that incentives matter most for executives who are in a position to act on a multi-year research agenda. Also, career-horizon models suggest that reputational incentives weaken as a CEO approaches retirement and has less of a vested interest in long-term outcomes (Dechow and Sloan 1991; Gibbons and Murphy 1992). The estimates in Panel A showing that stock awards have a greater innovation impact for long-tenured CEOs are consistent with these hypothesized effects.

Panels B and C split the sample based on the composition of the CEO's existing equity portfolio, using the ratio of exercisable to unexercisable option delta and the broader ratio of vested to unvested portfolio delta. We predict that CEOs whose portfolios are already weighted toward unvested, long-dated claims have substantial long-horizon exposure before the award vests, so an additional stock award has little effect. In contrast, a CEO whose exposure is concentrated in already-vested, liquid positions experiences the largest proportional lengthening of horizon from the same award (Gopalan et al. 2014; Edmans et al. 2017a). Our findings are consistent with this prediction. For example, in Panel B the estimated effect is 1.923 for grants versus -0.677 in the low-ratio group on 53 observations per group; its count cells are significant at the one percent level and its citations cell at

the five percent level. Both ratios are measured before the performance realization, so the splits are predetermined with respect to treatment. Panel C corroborates the findings in Panel B, with the low-ratio group showing small negative and imprecise estimates across all four outcomes while the high-ratio group shows 1.158 for grants, 1.004 for applications, 1.470 for real economic value, and 0.656 for citations, all significant on 123 observations.

Panel D provides the most direct test of the short-termism mechanism. We split the sample on the LLM-based managerial-myopia score, which rates the short-term orientation expressed in each firm's MD&A disclosure on an ordered five-point scale in the fiscal year before performance realization. The high-myopia group shows larger estimates on every outcome, at 1.057 against 0.242 for patent grants and 1.594 against 0.734 for real economic value, on 132 and 263 observations. When the firm's own disclosures reveal a more short-term orientation, a discrete increment of long-dated equity has the most room to change behavior, and the award is most impactful at overcoming the baseline myopic tendency. This shows that stock award realizations can effectively counteract the managerial short-termism studied theoretically in Stein (1989) and documented empirically with time vesting of equity options as in Ladika and Sautner (2020) and Edmans et al. (2017a).

Taken together, the various proxies provide a consistent picture. The innovation response to cutoff passage is concentrated among CEOs with the authority to redirect research, among CEOs for whom the marginal block is the largest proportional change in long-horizon exposure, and at firms whose disclosures reveal the strongest short-term orientation. The consistency of these orderings across independent proxies is what a managerial short-termism mechanism predicts, and it is difficult to reconcile with alternatives in which cutoff passage matters for reasons unrelated to incentives, since such reasons would operate symmetrically across the groups of every split.

5.2 Board discretion

Relative-TSR contracts frequently reserve discretion to the compensation committee, allowing it to adjust the payout that the formula would otherwise produce. Where such discretion exists, the mapping from realized performance to shares earned is no longer mechanical. a CEO who narrowly

misses may be made whole, and a CEO who narrowly clears may be adjusted downward. This is a measurement implication of the contract rather than a claim about board behavior, and it yields a sharp prediction. The discontinuity should be present where the contract leaves the committee no stock-award discretion and attenuated where it does not.

Panel A of Table 12 reports prevalence in the preferred sample. Of 431 awards, 323 carry no stock-award discretion and 108 carry some. Within the latter group, three awards specify discretion that can only increase the payout, 31 discretion that can only decrease it, 20 two-sided discretion, and 54 discretion whose direction is not stated. The direction-specific cells are small and serve only to describe the composition of the sample. Absent or unmentioned discretion is coded as no discretion, which is the conservative coding for our purposes: it places contracts with unstated discretion in the group where we expect to find an effect, biasing against the prediction.

Panel B reports the split estimates, and they line up with the prediction. In the no-discretion group the discontinuity is 0.603 for patent grants, 0.593 for applications, 0.975 for real economic value, and 0.442 for citations, on 309 observations, all significant and close to or above the pooled estimates. In the discretion group all four estimates are near zero or negative and none is distinguishable from zero, on 105 observations. The innovation response is present exactly where the contract converts performance into shares mechanically, and absent where the committee can undo the formula.

This ordering also bears on the leading alternative interpretation of the main results. Information conveyed by clearing a performance cutoff, about competitive position or about the CEO's standing with the board, is conveyed in both groups, since performance is realized and observable regardless of what the committee may do about the payout. What differs across the groups is only whether clearing the cutoff mechanically determines the shares earned. The concentration of the response in the no-discretion group is therefore the pattern an equity-exposure channel predicts and an information channel does not. Because discretion status is a contract feature that firms choose, we read the split as mechanism evidence within the design rather than as an experiment in contract

design, and the presence of discretion in a contract is evidence about the mapping from performance to pay, not evidence that discretion was exercised.

5.3 Selling, pledging, holding and anti-hedging

An important feature of equity-based incentives is that their effect depends on whether the CEO remains exposed to the shares after they are received. Prior work shows that managers may offset new equity compensation by subsequently selling shares (Ofek and Yermack 2000), reduce their effective ownership through hedging transactions (Bettis et al. 2001; Gao 2010), or obtain liquidity against their holdings by pledging shares as collateral (Dou et al. 2019). Stock-ownership guidelines can work in the opposite direction by encouraging executives to retain company stock, particularly when the requirements are binding (Core and Larcker 2002; Korczak and Liu 2014; Benson et al. 2016). Our findings are consistent with these arguments. Narrowly clearing a performance cutoff does not lead to a detectable increase in proposed selling or share pledging, and the positive effects of stock awards on innovation are concentrated or more precisely estimated when firm policies restrict pledging or hedging, or when CEOs are subject to stock-ownership requirements. Taken together, these results suggest that stock awards are most effective at encouraging innovation when the additional shares remain part of the CEO's exposure to long-run firm value rather than being monetized or otherwise offset.

Table 13 examines the first question using Form 144 filings and share-pledging disclosures over the first two post-realization years. Neither the value of proposed insider sales nor the probability of any proposed sale shows a discontinuity distinguishable from zero, and pledging, an alternative route to monetizing equity without selling it, shows no detectable discontinuity either. The probability of any CEO pledge is the closest outcome to conventional significance, at 0.067 against a standard error of 0.045, and remains indistinguishable from zero. Form 144 records proposed rather than executed transactions, so these are nulls on the observable signal of intent to sell; within that limit, we find no evidence that CEOs respond to narrowly earning more equity by moving to monetize it.

Table 14 examines the second question by splitting the innovation regressions on predetermined governance provisions: firm-level policies matched to the exact DEF 14A filing year preceding performance realization, and the CEO's own stock-ownership requirement. A pledging restriction prevents the CEO from posting company stock as loan collateral, an anti-hedging policy prevents the CEO from neutralizing the stock's economic exposure with derivative positions, and a holding requirement obliges the CEO to retain earned shares as ownership.

Panels A and B report the policy splits. Among awards at firms that disclose a pledging restriction, the discontinuity is 0.460 for patent grants, 0.459 for applications, and 0.425 for real economic value, each significant on 240 observations, while the 125 awards without a disclosed restriction show estimates near zero. The anti-hedging split shows the same asymmetry in significance: the 296 awards in the policy group are significant on all four outcomes, at 0.489 for patent grants and 0.633 for real economic value, while the 69 awards without a disclosed policy are estimated too imprecisely to distinguish from zero.

Panel C turns to stock-ownership requirements, which operate on the same margin from inside the CEO's own contract. A holding requirement converts newly earned shares from monetizable wealth into retained exposure, which is precisely the difference between a wealth shock and an incentive shock. The panel compares awards with no disclosed requirement to awards whose CEO faces a requirement and meets it in full; the two groups are coded from separate disclosure indicators that treat undisclosed compliance differently, so they overlap rather than partition the sample. Both groups show positive estimates, and the estimates are uniformly larger where the requirement is met in full, at 0.671 against 0.304 for patent grants, 0.730 against 0.189 for applications, and 1.021 against 0.529 for real economic value. Where governance ensures that the CEO retains the shares an award delivers, the incentive content of cutoff passage is preserved, and the innovation response is correspondingly larger.

The two tables answer the monetization question from both directions. Cutoff passage does not trigger observable selling or pledging, and the innovation response is concentrated precisely where governance keeps earned shares as exposure, whether through firm-level pledging and

hedging policies or through the CEO's own holding requirement. Both facts are what an incentive interpretation requires, and together with the discretion split they tie the innovation response to the incentive content of the award rather than to the event of passage itself.

6 Conclusion

This paper studies whether CEO equity-based incentives help drive corporate innovation. To overcome endogeneity concerns and identify causal effects, we implement a novel regression discontinuity (RD) design that makes use of a large sample of CEO performance share plans from 1998 to 2024. We capture exogenous variation in stock-based compensation by exploiting discontinuous jumps in equity award payouts around stock return cutoffs. Since we focus only on stock awards with goals based on stock performance (e.g., Total Shareholder Return) rather than accounting performance, it is much less likely that our econometric approach is invalidated by the ability of CEOs to precisely manipulate performance near a cutoff.

The evidence from our analysis supports the idea that realized equity awards are effective at driving innovation activity and shifting CEOs toward long-horizon innovative investment. Moreover, we find that meeting a performance cutoff is economically meaningful: it raises both the equity assigned to a CEO, a block worth roughly 8.5% of pre-award stock holdings, and the equity the CEO actually earns relative to prior holdings. Furthermore, compared to firms that narrowly miss a performance cutoff, meeting one leads to approximately 70% more patenting over three years and causes patents to span a wider set of technology classes and carry higher market-implied value and more forward citations. We provide evidence of the validity of our empirical design by confirming that discontinuities appear only at the contractual cutoff and vanish at placebo cutoffs. Furthermore, the running variable shows no sign of manipulation, predetermined firm characteristics including prior patenting are continuous across the cutoff, and the results survive alternative bandwidths, kernels, polynomial orders, and the restriction to exactly measured performance.

Heterogeneity tests reveal that stock awards seem to spur innovation activity by lengthening the horizon of top management. In particular, under several structural and text-based measures of CEO short-termism, the innovation effects of stock awards are significantly stronger when the CEO's pre-existing myopia is high. Moreover, we find that the innovation effects of realized share awards are stronger when ex post board discretion is absent and when the firm has governance policies in place that foster continued equity exposure by restricting the CEO's ability to pledge, hedge, or sell shares. Measured R&D intensity does not detectably move, consistent with the incentive operating through project selection and continuation within existing budgets rather than through the budget line itself.

Finally, we acknowledge two limitations of our study. First, as with any regression-discontinuity design, the estimates are local: they describe the innovation effects of narrowly clearing a Relative-TSR cutoff among firms that write such contracts, rather than the innovation effects of stock compensation for firms that experience performance far away from a cutoff. Second, some observed effects of new equity-based awards likely take considerable time to manifest, and thus the patent grant and citation results rest partly on cohorts whose outcomes are still accruing. Direct evidence on realized share transactions, cutoff-specific treatment effects, and pooled interaction tests for the mechanism splits would sharpen our interpretations further and shed additional light on the impact of CEO equity awards on the firm's subsequent innovation.

References

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Agrawal, A. and Mandelker, G. N. (1987). Managerial incentives and corporate investment and financing decisions. *Journal of Finance*, 42(4):823–837.

Barker, V. L. and Mueller, G. C. (2002). CEO characteristics and firm R&D spending. *Management Science*, 48(6):782–801.

Bennett, B., Bettis, J. C., Gopalan, R., and Milbourn, T. (2017). Compensation goals and firm performance. *Journal of Financial Economics*, 124(2):307–330.

Benson, B. W., Lian, Q., and Wang, Q. (2016). Stock ownership guidelines for CEOs: Do they (not) meet expectations? *Journal of Banking & Finance*, 69:52–71.

Bernstein, S. (2015). Does going public affect innovation? *Journal of Finance*, 70(4):1365–1403.

Bettis, J. C., Bizjak, J., Coles, J. L., and Kalpathy, S. (2010). Stock and option grants with performance-based vesting provisions. *Review of Financial Studies*, 23(10):3849–3888.

Bettis, J. C., Bizjak, J., Coles, J. L., and Kalpathy, S. (2018). Performance-vesting provisions in executive compensation. *Journal of Accounting and Economics*, 66(1):194–221.

Bettis, J. C., Bizjak, J. M., and Lemmon, M. L. (2001). Managerial ownership, incentive contracting, and the use of zero-cost collars and equity swaps by corporate insiders. *Journal of Financial and Quantitative Analysis*, 36(3):345–370.

Biggerstaff, L., Blank, B., and Goldie, B. (2019). Do incentives work? Option-based compensation and corporate innovation. *Journal of Corporate Finance*, 58:415–430.

Bloomfield, M. J., Heinle, M. S., and Timmermans, O. (2024). Relative performance evaluation and strategic peer-harming disclosures. *Journal of Accounting Research*, 62(3):877–933.

Brav, A., Jiang, W., Ma, S., and Tian, X. (2018). How does hedge fund activism reshape corporate innovation? *Journal of Financial Economics*, 130(2):237–264.

Brisley, N. (2006). Executive stock options: Early exercise provisions and risk-taking incentives. *Journal of Finance*, 61(5):2487–2509.

Burgstahler, D. and Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1):99–126.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Carter, M. E., Ittner, C. D., and Zechman, S. L. C. (2009). Explicit relative performance evaluation in performance-vested equity grants. *Review of Accounting Studies*, 14(2–3):269–306.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, 18(1):234–261.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Chang, X., Fu, K., Low, A., and Zhang, W. (2015). Non-executive employee stock options and corporate innovation. *Journal of Financial Economics*, 115(1):168–188.

Coles, J. L., Daniel, N. D., and Naveen, L. (2006). Managerial incentives and risk-taking. *Journal of Financial Economics*, 79(2):431–468.

Core, J. E., Holthausen, R. W., and Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*, 51(3):371–406.

Core, J. E. and Larcker, D. F. (2002). Performance consequences of mandatory increases in executive stock ownership. *Journal of Financial Economics*, 64(3):317–340.

Core, J. E. and Packard, H. A. (2022). Non-price and price performance vesting provisions and CEO incentives. *The Accounting Review*, 97(7):109–134.

De Angelis, D. and Grinstein, Y. (2015). Performance terms in CEO compensation contracts. *Review of Finance*, 19(2):619–651.

Dechow, P. M. and Sloan, R. G. (1991). Executive incentives and the horizon problem: An empirical investigation. *Journal of Accounting and Economics*, 14(1):51–89.

Degeorge, F., Patel, J., and Zeckhauser, R. (1999). Earnings management to exceed thresholds. *Journal of Business*, 72(1):1–33.

Diamond, D. W. and Verrecchia, R. E. (1982). Optimal managerial contracts and equilibrium security prices. *Journal of Finance*, 37(2):275–287.

Dou, Y., Masulis, R. W., and Zein, J. (2019). Shareholder wealth consequences of insider pledging of company stock as collateral for personal loans. *Review of Financial Studies*, 32(12):4810–4854.

Edmans, A., Fang, V. W., and Lewellen, K. A. (2017a). Equity vesting and investment. *Review of Financial Studies*, 30(7):2229–2271.

Edmans, A., Gabaix, X., and Jenter, D. (2017b). Executive compensation: A survey of theory and evidence. In Hermalin, B. E. and Weisbach, M. S., editors, *The Handbook of the Economics of Corporate Governance*, volume 1, pages 383–539. North-Holland, Amsterdam.

Fama, E. F. and Jensen, M. C. (1983). Separation of ownership and control. *Journal of Law and Economics*, 26(2):301–325.

Faurel, L., Li, Q., Shanthikumar, D., and Teoh, S. H. (2024). Bringing innovation to fruition: Insights from new trademarks. *Journal of Financial and Quantitative Analysis*, 59(2):474–520.

Flammer, C. and Bansal, P. (2017). Does a long-term orientation create value? Evidence from a regression discontinuity. *Strategic Management Journal*, 38(9):1827–1847.

Gao, H. (2010). Optimal compensation contracts when managers can hedge. *Journal of Financial Economics*, 97(2):218–238.

Gao, M. (2026). Get the money somehow: The effect of missing performance goals on insider trading. *Journal of Financial and Quantitative Analysis*. First View.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Gibbons, R. and Murphy, K. J. (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy*, 100(3):468–505.

Gopalan, R., Milbourn, T., Song, F., and Thakor, A. V. (2014). Duration of executive compensation. *Journal of Finance*, 69(6):2777–2817.

Graham, J. R., Harvey, C. R., and Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40(1–3):3–73.

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

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 36(1):16–38.

Hayes, R. M., Lemmon, M., and Qiu, M. (2012). Stock options and managerial incentives for risk taking: Evidence from FAS 123R. *Journal of Financial Economics*, 105(1):174–190.

Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7(1–3):85–107.

Holmström, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization*, 12(3):305–327.

Holmström, B. and Tirole, J. (1993). Market liquidity and performance monitoring. *Journal of Political Economy*, 101(4):678–709.

Holthausen, R. W., Larcker, D. F., and Sloan, R. G. (1995). Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics*, 19(1):29–74.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5):984–1001.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Jochem, T., Ladika, T., and Sautner, Z. (2018). The retention effects of unvested equity: Evidence from accelerated option vesting. *Review of Financial Studies*, 31(11):4142–4186.

Katila, R. and Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal*, 45(6):1183–1194.

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

Korczak, P. and Liu, X. (2014). Managerial shareholding policies and retention of vested equity incentives. *Journal of Empirical Finance*, 27:116–129.

Ladika, T. and Sautner, Z. (2020). Managerial short-termism and investment: Evidence from accelerated option vesting. *Review of Finance*, 24(2):305–344.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Lerner, J. and Seru, A. (2022). The use and misuse of patent data: Issues for finance and beyond. *The Review of Financial Studies*, 35(6):2667–2704.

Lerner, J., Sørensen, M., and Strömberg, P. (2011). Private equity and long-run investment: The case of innovation. *Journal of Finance*, 66(2):445–477.

Lerner, J. and Wulf, J. (2007). Innovation and incentives: Evidence from corporate R&D. *Review of Economics and Statistics*, 89(4):634–644. NBER Working Paper No. 11944.

Malenko, N. and Shen, Y. (2016). The role of proxy advisory firms: Evidence from a regression-discontinuity design. *Review of Financial Studies*, 29(12):3394–3427.

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

Mao, C. X. and Zhang, C. (2018). Managerial risk-taking incentive and firm innovation: Evidence from FAS 123R. *Journal of Financial and Quantitative Analysis*, 53(2):867–898.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Ofek, E. and Yermack, D. (2000). Taking stock: Equity-based compensation and the evolution of managerial ownership. *Journal of Finance*, 55(3):1367–1384.

Pawliczek, A. (2021). Performance-vesting share award outcomes and CEO incentives. *The Accounting Review*, 96(5):337–364.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3):335–370.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. Harper & Brothers, New York.

Shue, K. and Townsend, R. R. (2017). How do quasi-random option grants affect CEO risk-taking? *Journal of Finance*, 72(6):2551–2588.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94.

Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics*, 104(4):655–669.

Tchistyi, A., Yermack, D., and Yun, H. (2011). Negative hedging: Performance-sensitive debt and CEOs' equity incentives. *Journal of Financial and Quantitative Analysis*, 46(3):657–686.

Appendix

This appendix collects supplementary estimates, robustness tests, and variable definitions. Table A.2 reports formal robust bias-corrected estimates at the maintained bandwidth for every mechanism and innovation outcome. Table A.3 adds alternative predetermined vesting and wealth controls to the maintained innovation regressions. Table A.4 reports the complete density-test sensitivity grid. Table A.5 reports the placebo-cutoff falsification tests, and Table A.1 defines every variable used in the paper.

Table A.1: Definition of Variables

Variable	Definition
RD design and assignment variables	
Relative-TSR distance from cutoff	$X =$ focal TSR $-$ required focal TSR; use the exact point when the identified interval is degenerate and otherwise the stored finite upper endpoint.
Barely passes cutoff	$D = \mathbf{1}[X \geq 0]$.
Triangular kernel weight	$w_i = (1 - \lvert X_i \rvert / h)\mathbf{1}[\lvert X_i \rvert \leq h]$, with $h = 0.075$ unless a robustness column states otherwise.
Exact running-variable observation	$\mathbf{1}[\text{lower bound} = \text{upper bound} = \text{exact point}]$.
Estimated interval upper endpoint	$\mathbf{1}[\text{interval point}]$; the analysis point is `running_variable_upper_bound` and is never labeled exact.
Performance-realization-year fixed effect	year(performance_end_date).
FF12-industry fixed effect	FF12(SIC) using the published Fama–French 12-industry SIC ranges and the latest Compustat date strictly before performance start.
Cutoff-type fixed effect	Categorical indicators for threshold, target, and maximum payout cutoffs.
Interpolation-status fixed effect	Y→interpolation disclosed; N→no interpolation disclosed; NC→not clear; blank→source missing.
CEO equity and vesting variables	
Award Jump	$D \times$ ex ante cutoff jump value$/(\text{sharesOwned}_{t-1} \times P_{t-1})$, requiring a positive denominator.
Cumulative Vesting Intensity	$(\text{valueVesting}_t + \text{valueVesting}_{t+1})/(\text{sharesOwned}_{t-1} \times P_{t-1})$.
Cumulative Vesting Share Ratio	$(\text{numVesting}_t + \text{numVesting}_{t+1})/\text{sharesOwned}_{t-1}$.
Pre-event CEO common-stock wealth	$\text{sharesOwned}_{t-1} \times P_{t-1}$ for directly priceable CEO shares.
Pre-event beneficially owned shares	Total beneficially owned shares at the selected pre-event CEO snapshot.

Continued on next page

Table A.1, continued

Variable	Definition
Predetermined scheduled nonfocal vesting	Scheduled pre-event vesting from awards other than the focal award; retained only when deterministically observed.
Innovation outcomes	
Patent grants	$\log[1+\sum_{s=1}^{3} n_granted_{t+s}]$; each annual count is the number of applications filed that year and observed as granted by the frozen data cutoff.
Patent applications	$\log[1+\sum_{s=1}^{3} n_applications_{t+s}]$.
Mean R&D intensity	$\log[1+m^{-1}\sum_{s\in\mathcal{O}}(XRD_{t+s}/AT_{t+s})]$, where $\mathcal{O}\subseteq\{1,2,3\}$ is the set of observed years and $m=\lvert\mathcal{O}\rvert$.
Adjusted patent originality	$O_{1:3}=\sum_{s=1}^{3} w_s O_s$, where $w_s=n_s/\sum_{r=1}^{3} n_r$ and n_s is annual application support. Annual $O_s=\text{mean}[N/(N-1)(1-HHI_{backward})]$ for applications with $N>1$ usable backward citations.
Adjusted patent generality	$G_{1:3}=\sum_{s=1}^{3} w_s G_s$, where $w_s=n_s/\sum_{r=1}^{3} n_r$ and n_s is annual application support. Annual $G_s=\text{mean}[N/(N-1)(1-HHI_{forward})]$ for applications with $N>1$ usable forward citations.
Technological breadth	$m^{-1}\sum_{s\in\mathcal{O}}$ count distinct CPC subclasses$_{t+s}$ across observed years $t+1$ to $t+3$; annual breadth counts the union within that firm's annual application portfolio.
Real Economic Value	$\log[1+\sum_{s=1}^{3}$ xi_real_total$_{t+s}]$, the inflation-adjusted market-based value of matched patent applications.
Forward Citations	$\log[1+\sum_{s=1}^{3}$ cites_total$_{t+s}]$ for patent applications matched to the patent-value/citation source.
Control variables	
Firm size	$\log(AT)$ for the latest Compustat fiscal observation strictly before performance start, requiring $AT>0$.
ROA	NI/AT using the latest Compustat observation strictly before performance start.
Leverage	$[\text{coalesce}(DLC,0)+\text{coalesce}(DLTT,0)]/AT$.
Cash/assets	CHE/AT.
Market-to-book	$(CSHO\times PRCC_F)/CEQ$, requiring $CEQ>0$.

Continued on next page

Table A.1, continued

Variable	Definition
Prior 12-month return	$\prod_d(1+RET_d)-1$ over the 366 calendar days before performance start, requiring at least 100 valid CRSP trading days.
Prior idiosyncratic volatility	Sample SD of residuals from $RET_d = \alpha + \beta MKTRET_d + u_d$ over the pre-start 366-day window, requiring at least 100 paired days.
Pre-period R&D intensity	Arithmetic mean of annual XRD/AT over observed years $t-3$ through $t-1$.
Pre-period adjusted originality	Application-support-weighted mean of annual adjusted originality over $t-3$ through $t-1$.
Pre-period adjusted generality	Application-support-weighted mean of annual adjusted generality over $t-3$ through $t-1$.
CEO heterogeneity variables	
CEO tenure	Years from the CEO appointment date to the pre-performance-start reference date; when appointment is unavailable, pre-event fiscal year minus first observed CEO firm-year (left-censored proxy).
CEO holding requirement	Indicator that a CEO ownership/holding guideline exists; source missing is coded zero/no requirement for the requested heterogeneity test.
Holding requirement met in full	Indicator that the realization-year CEO requirement is disclosed as fully met; observed partial/no compliance equals zero and undisclosed compliance remains missing.
Exercisable/unexercisable option-delta ratio	exercisable option delta$_{t-1}$/unexercisable option delta$_{t-1}$ when the denominator is positive.
Vested-proxy delta at t-1	$\Delta^{VP}_{t-1} = \Delta^{\text{owned stock}}_{t-1} + \Delta^{\text{exercisable options}}_{t-1}$, measured as the dollar change in CEO equity value for a 1% stock-price increase. Owned-stock delta is $0.01P_{t-1} \times$ sharesOwned$_{t-1}$; exercisable-option delta sums $0.01P_{t-1} \times$ numExercisable $\times$ $e^{-qT}\Phi(d_1)$ across option tranches. Beneficially owned stock is an explicit liquid/vested proxy, not provider-reported vested restricted stock; unavailable components remain missing.

Continued on next page

Table A.1, continued

Variable	Definition
Unvested delta at t-1	$\Delta^{U}_{t-1} = \Delta^{\text{time-unvested stock}}_{t-1} + \Delta^{\text{unexercisable options}}_{t-1}$, measured as the dollar change in CEO equity value for a 1% stock-price increase. Time-unvested-stock delta is $0.01 P_{t-1} \sum_j \text{numUnvested}_{j,t-1}$ and unexercisable-option delta uses the corresponding Black–Scholes delta summed across tranches. Performance-unearned stock and options are excluded; unavailable components remain missing.
Broad vested-proxy/unvested delta ratio	vested proxy delta_{t-1}/unvested delta_{t-1} when the denominator is positive.
LLM managerial-myopia score	Documented 1–5 ordered score matched by normalized CIK and exact fiscal year $t-1$; the displayed split uses the main-sample median.
Governance-policy moderators	
Pledging restriction policy at t-1	Indicator equal to one when any supplied firm filing in the exact calendar year $t-1$ discloses a meaningful ban or restriction on pledging company securities, and zero when the matched filing discloses no qualifying restriction. Firm-years without a supplied filing remain missing.
Anti-hedging policy at t-1	Indicator equal to one when any supplied firm filing in the exact calendar year $t-1$ discloses an anti-hedging restriction covering company securities, and zero when the matched filing discloses no qualifying policy. Firm-years without a supplied filing remain missing.
Both policies at t-1	Product of the pledging-restriction and anti-hedging indicators for a matched $t-1$ filing year. It equals one only when both are disclosed and zero when fewer than both are disclosed; unmatched firm-years remain missing. A zero is not proof that a policy did not exist outside the supplied filing.
Board discretion and CEO selling outcomes	
Board stock-award discretion	$\mathbf{1}$[stock awards or stock awards and bonus]; absent or unmentioned discretion is coded zero. Direction indicators classify increase, reduce, or two-sided authority.
Proposed Form 144 sale value	$\log[1 + \text{stock-only value proposed for sale}_{t+1:t+2}]$, conditional on complete five-year source-window coverage.

Continued on next page

Table A.1, continued

Variable	Definition
Any proposed Form 144 sale	Indicator for at least one eligible CEO stock-only proposed-sale notice in $t+1$ or $t+2$, conditional on complete source-window coverage.
Any CEO share pledge	$\max(\text{pledge flag}_{t+1}, \text{pledge flag}_{t+2})$; an annual documented yes dominates no and information-not-found is coded zero under the user-specified rule.
Mean annual CEO shares pledged	$\log[1 + (\text{shares pledged}_{t+1} + \text{shares pledged}_{t+2})/2]$; unavailable annual share counts are coded zero under the documented construction.

Table A.2: Maintained-Bandwidth Formal rdrobust Estimates

Outcome	Bias-corrected coef.	Robust SE	p-value	Bandwidth	Observations
Award Jump	0.090***	(0.031)	0.003	0.075	210
Cumulative Vesting Intensity	0.255	(0.161)	0.114	0.075	322
Cumulative Vesting Share Ratio	0.246*	(0.137)	0.072	0.075	340
Log(1 + patent grants)	0.783**	(0.315)	0.013	0.075	396
Log(1 + patent applications)	0.658*	(0.363)	0.070	0.075	396
Log(1 + mean R&D intensity)	0.004*	(0.002)	0.069	0.075	371
Adjusted patent originality	0.131**	(0.053)	0.014	0.075	85
Adjusted patent generality	-0.382***	(0.042)	0.000	0.075	51
Technological breadth	6.941**	(3.181)	0.029	0.075	396
Log(1 + Real Economic Value)	0.966*	(0.497)	0.052	0.075	396
Log(1 + Forward Citations)	0.618**	(0.291)	0.034	0.075	396
Specification indicators					
Local polynomial / bias polynomial		$p = 1$ / $q = 2$			
Triangular kernel; mass-point adjustment		Yes			
All four maintained fixed effects		Yes			
Full outcome-specific controls		Yes			
Firm-clustered inference		Yes			

Notes: The unit is one selected award with performance-realization year below 2023. The running variable is centered Relative-TSR distance and treatment is one at or above zero. This appendix table reports only the maintained covariate-adjusted formal rdrobust specification. Both the estimation and bias bandwidths are fixed at 0.075; the local polynomial is linear ($p = 1$), the bias polynomial is quadratic ($q = 2$), the kernel is triangular, and mass points are adjusted. Coefficients are robust bias-corrected estimates, parentheses contain robust standard errors, and stars use the 1%, 5%, and 10% two-sided levels. The covariate set contains log assets, prior 12-month return, prior idiosyncratic volatility, and the documented outcome-specific predetermined controls. The fixed effects are performance-realization year, FF12 industry, cutoff type, and interpolation status; mechanism outcomes retain their documented predetermined vesting/wealth controls. Inference is CR1-clustered by firm. Observations are the total effective left-plus-right sample inside the fixed bandwidth. Continuous analysis variables are winsorized at 1%/99%. No reduced-model fallback is used; a dash denotes a saved non-estimation. Table A.1 gives exact variable formulas.

Table A.3: Alternative Predetermined Vesting-Control Specifications

Control set	Log(1 + patent grants)	Log(1 + patent applications)	Log(1 + Real Economic Value)	Log(1 + Forward Citations)
Baseline	0.528**	0.481*	0.853***	0.453**
	(0.208)	(0.246)	(0.298)	(0.185)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Add scheduled nonfocal vesting	0.533**	0.458	1.020***	0.515**
	(0.238)	(0.284)	(0.328)	(0.215)
	[$N = 363$]	[$N = 363$]	[$N = 363$]	[$N = 363$]
Add pre-event CEO common-stock wealth	0.486**	0.432*	0.876***	0.423**
	(0.219)	(0.258)	(0.316)	(0.205)
	[$N = 348$]	[$N = 348$]	[$N = 348$]	[$N = 348$]
Add both vesting and common-stock wealth	0.526**	0.443	1.074***	0.475**
	(0.259)	(0.305)	(0.354)	(0.242)
	[$N = 303$]	[$N = 303$]	[$N = 303$]	[$N = 303$]

Notes: Each cell is the Barely passes cutoff coefficient from an award-level Sharp RD, followed by the firm-clustered standard error and number of observations. The limited ladder is fixed ex ante for this report and does not search across controls for significance. Baseline includes the maintained four fixed effects, separate local-linear slopes, log assets, prior 12-month return, and prior idiosyncratic volatility. Subsequent rows add predetermined scheduled nonfocal vesting, directly priceable pre-event CEO common-stock wealth, or both. All additions are measured before performance realization. Every model uses $h = 0.075$, triangular weights, the one-cutoff-per-award selector, the same complete-case logic, and 1%/99% winsorization. Baseline estimates are reused from the main machine output; other rows are new full-specification estimates. Stars use the 1%, 5%, and 10% two-sided levels. Table A.1 defines the added controls.

Table A.4: Density-Specification Sensitivity

Method	Specification	Polynomial	Bandwidth / window	Statistic	p-value	Observations
CJM mass-point adjusted	MSE-optimal bandwidth	2	0.045 / 0.033	-0.090	0.928	431
McCrary-style binned log density	0.25-pp bins; 7.5-pp window; local linear	1	0.075 / 0.075	-1.327	0.184	431
McCrary-style binned log density	0.5-pp bins; 7.5-pp window; local linear	1	0.075 / 0.075	-1.420	0.155	431
McCrary-style binned log density	1-pp bins; 7.5-pp window; local linear	1	0.075 / 0.075	-1.137	0.256	431
McCrary-style binned log density	0.5-pp bins; 7.5-pp window; local quadratic	2	0.075 / 0.075	-1.129	0.259	431
McCrary-style binned log density	0.5-pp bins; 5-pp window; local linear	1	0.050 / 0.050	-1.338	0.181	259
McCrary-style binned log density	0.5-pp bins; 10-pp window; local linear	1	0.100 / 0.100	-1.623	0.105	534

Notes: The sample contains selected awards with performance-realization year below 2023, and every test uses the same centered Relative-TSR running variable as the main analysis. CJM rows use mass-point adjustment, triangular kernels, MSE-optimal or stated side-specific bandwidths, and jackknife inference. McCrary-style rows use binned log densities, a 0.5-count continuity correction, the stated polynomial and fitting window, and HC3 inference. Statistics test equality of the left and right density limits at zero. Observations count selected awards with a nonmissing running variable. Failure to reject is not proof that manipulation is impossible.

Table A.5: Fake-Cutoff Falsification Tests

Outcome	Fake-cutoff jump	Clustered SE	p-value	Observations
Panel A. Fixed fake cutoff at +5 percentage points				
Log(1 + patent grants)	0.022	(0.202)	0.912	201
Log(1 + patent applications)	0.076	(0.248)	0.760	201
Log(1 + Real Economic Value)	-0.407	(0.402)	0.311	201
Log(1 + Forward Citations)	-0.138	(0.221)	0.533	201
Panel B. Randomized fake-cutoff design				
Log(1 + patent grants)	-0.144	(0.312)	0.646	194
Log(1 + patent applications)	-0.212	(0.323)	0.512	194
Log(1 + Real Economic Value)	0.062	(0.468)	0.894	194
Log(1 + Forward Citations)	0.128	(0.413)	0.757	194
Specification indicators for every row				
Panel A symmetric bandwidth	5.0 percentage points			
Panel B symmetric bandwidth	5.0 percentage points			
Panel B randomized cutoff	$+8.5$ percentage points			
Panel B candidate grid	$[-10,-5.5]\cup[+5.5,+10]$, in 0.5-point increments			
Separate local slopes at fake cutoff	Yes			
All four fixed effects	Yes			
Full outcome-specific controls	Yes			
Firm-clustered inference	Yes			

Notes: Panel A fixes the fake cutoff at $+5$ percentage points and uses a symmetric 5-percentage-point bandwidth, giving support $[0,+10]$. Panel B is a reproducible randomized placebo. Before estimation, seed 20260826 draws one cutoff from the symmetric 0.5-percentage-point grid $[-10,-5.5]\cup[+5.5,+10]$, which excludes Panel A's cutoff and ensures that every candidate's complete 5-point window lies strictly on one side of the true zero cutoff. The realized cutoff is $+8.5$ percentage points, giving support $[3.5,13.5]$. Panel A retains the bandwidth-first ± 10-point selector used by the fixed design; Panel B alone uses the expanded ± 15-point selector required to supply every candidate window. The maintained sample and every non-placebo analysis remain at $h = 0.075$. Within each same-side support, $X_i^F = X_i - c$ and $F_i = \mathbf{1}[X_i \geq c]$, and the reported fake jump τ^F comes from $Y_i = \alpha + \tau^F F_i + \beta X_i^F + \gamma F_i X_i^F + \mathrm{FE}_i + \mathrm{Controls}_i + \varepsilon_i$. Every row uses triangular weights centered on the fake cutoff, performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, log assets, prior 12-month return, prior idiosyncratic volatility, and any documented outcome-specific predetermined controls. Standard errors are clustered by firm, observations are awards, and continuous analysis variables retain the 1%/99% winsorization. The seed, candidate grid, bandwidth, and outcome set are fixed independently of the estimated coefficients and significance levels. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

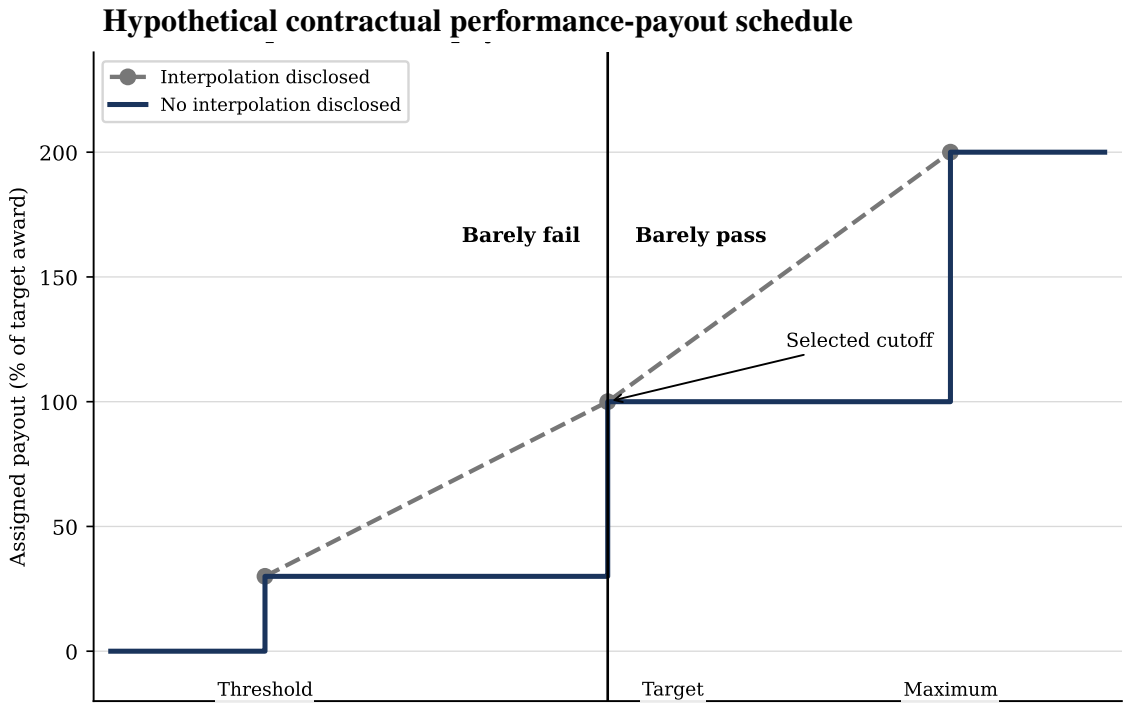


Figure 1: Intuition for the Relative-TSR Performance-Vesting Discontinuity. The figure depicts a hypothetical example of a contractual payout schedule with discrete jumps in share payouts at the threshold, target, and maximum goals. Realized Relative-TSR performance is on the horizontal axis and the assigned payout, as a percentage of the target award, is on the vertical axis. The step function depicts a schedule that provides for no interpolation and the dashed line a schedule for which interpolation between goals is disclosed. Barely-fail and barely-pass observations sit on either side of the selected cutoff, illustrating the local comparison the design uses.

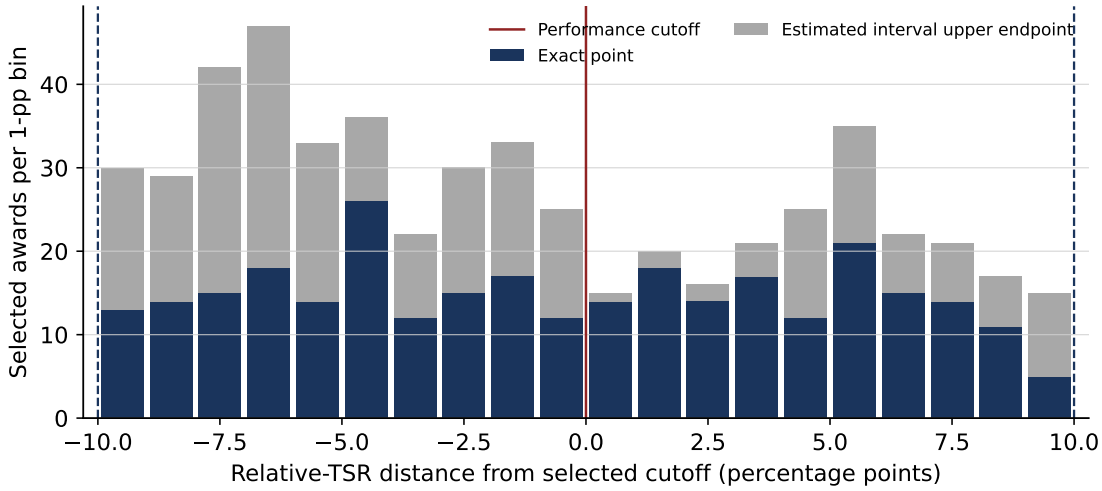


Figure 2: Distribution and Composition of the Running Variable. Bars show one-percentage-point bins of Relative-TSR distances from the selected contractual cutoff within ± 10 percentage points, separating observations whose identified interval is degenerate (exact running variable) from observations recorded at the upper endpoint of a non-degenerate identified interval. The solid vertical marker is the cutoff at zero and the dashed markers are the ± 10-percentage-point display-window boundaries. The plotted window contains 534 awards, 297 of them exact points and 237 recorded at an interval upper endpoint; within the maintained ± 7.5-percentage-point analysis bandwidth the sample contains 256 exact and 175 interval-endpoint observations.

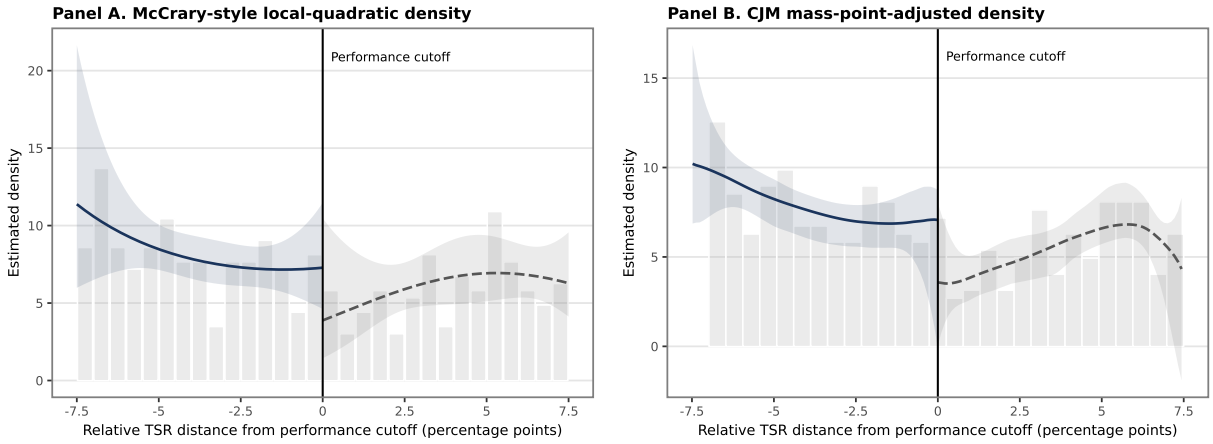


Figure 3: Density Continuity (Manipulation) Tests. Both panels test continuity of the running-variable density at the cutoff, which would be violated by precise sorting across a payout cutoff, on the maintained 431-award sample; in both, light bars display the empirical running-variable distribution. Panel A is a McCrary-style diagnostic using 0.5-percentage-point bins inside the maintained 7.5-percentage-point window and separate local-quadratic fits to log bin density; the estimated log-density jump is -0.630 (SE $= 0.558$, $p = 0.259$). Panel B plots the mass-point-adjusted local-polynomial density estimator of Cattaneo et al. (2020) with pointwise 95% confidence regions, together with the manipulation test of Cattaneo et al. (2018), which is the version the paper relies on because rank-based contracts place substantial mass at repeated values of the running variable; at MSE-optimal side-specific bandwidths of 0.045 and 0.033 the test statistic is -0.090 ($p = 0.928$). Appendix Table A.4 reports the complete sensitivity grid, in which no specification rejects continuity at the ten-percent level. Failure to reject is evidence against a detectable density discontinuity under a given specification.

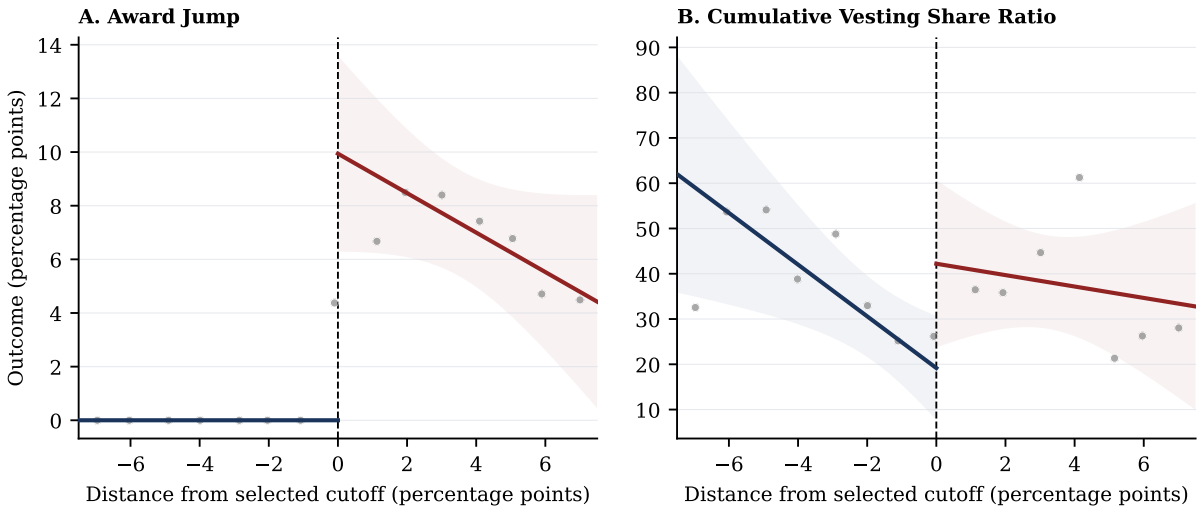


Figure 4: Discontinuities in CEO Equity Outcomes. Panel A plots Award Jump and Panel B the Cumulative Vesting Share Ratio against the Relative-TSR distance from the selected contractual cutoff. Points are one-percentage-point raw binned means. Lines are fitted separately below and above zero by triangular-kernel weighted least squares inside $h = 0.075$, with separate local slopes but no fixed effects and no controls. Shaded regions are firm-clustered 95% confidence bands. Each panel uses the complete-case sample required by that outcome's full-controls column in Table 5, and the fitted discontinuity equals the corresponding different-slopes coefficient in that table (0.099 for Award Jump and 0.230 for the Cumulative Vesting Share Ratio). The vertical axes are expressed in percent of the relevant pre-event base, so the plotted jumps of roughly 10 and 23 are those coefficients multiplied by 100. Panel A displays the size of the assigned entitlement change rather than an estimated discontinuity since Award Jump is zero below the cutoff by construction.

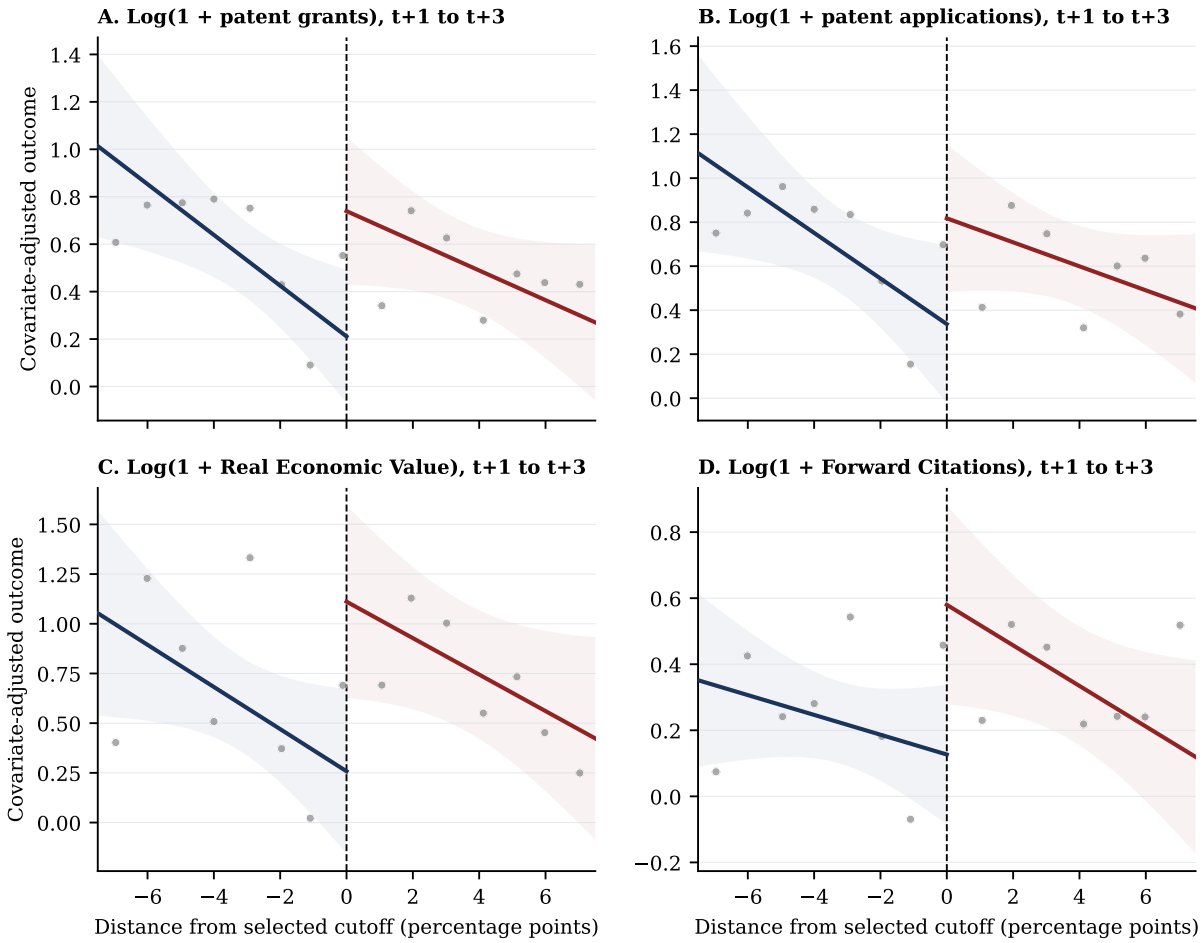


Figure 5: Discontinuities in Innovation Outcomes. Panels plot the log of one plus patent grants, patent applications, Real Economic Value, and Forward Citations over $t+1$ through $t+3$ against the Relative-TSR distance from the selected contractual cutoff. Points are one-percentage-point covariate-adjusted binned means. Lines average predictions from the maintained triangular-kernel local-linear model over the observed covariate and fixed-effect distribution, fitted separately below and above zero inside $h = 0.075$; shaded regions are firm-clustered 95% delta-method confidence bands. Every panel includes performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, log assets, prior 12-month return, prior idiosyncratic volatility, and the documented outcome-specific predetermined controls, so the plotted level discontinuity equals the corresponding full-controls coefficient in Tables 6 and 8.

Table 1: Sample Construction and RD Coverage

Sample restriction	Awards	% of all awards	Firms
Panel A. Incentive Lab award sample funnel			
All awards	778,858	100.0	4,802
Equity awards	546,371	70.2	4,775
Performance or accelerated equity awards	186,362	23.9	3,926
RSU performance or accelerated equity awards	172,916	22.2	3,712
Relative RSU performance awards	79,856	10.3	2,259
Awards using TSR, stock return, stock price, or market value as performance measures	71,060	9.1	2,110
Awards with one or two performance-group structures	25,211	3.2	1,110
CEO awards	5,683	0.7	1,059

RD sample construction	Awards
Panel B. Running-variable construction and final RD sample	
All unique awards with a calculable cutoff	1,380
Selected sample by symmetric bandwidth	
$h = 2.5$ percentage points	139
$h = 5.0$ percentage points	278
$h = 7.5$ percentage points	431
$h = 10.0$ percentage points	534
Composition of the preferred $h = 7.5$ percentage-point sample	
Cutoff-type decomposition (sums to 431 awards)	
Target cutoff	246
Threshold cutoff	105
Maximum cutoff	80
Running-variable-source decomposition (sums to 431 awards)	
Exact running-variable observation	256
Estimated interval upper endpoint	175

Notes: Panel A covers Incentive Lab fiscal years 1998-2024 and is an unweighted, strictly cumulative award-level funnel. An award is one canonical grant identifier, and the firm count is the number of issuers represented at each stage. Each successive row imposes the named restriction on the preceding row, counts therefore are nested and are not additive. The award funnel ends at 5,683 CEO awards. Panel B is the empirical RD universe and the preferred sample spans performance-realization years 2005-2022. The running variable is focal Relative-TSR less the award-specific required Relative-TSR at a threshold, target, or maximum payout landmark. Bandwidth filtering is applied before retaining one cutoff per award under the target-threshold-maximum priority; the maintained window is +/-7.5 percentage points. The cutoff-type and running-variable-source blocks are alternative decompositions of the same 431-award preferred sample. Each block sums to 431 and the blocks must not be added together. Panel A and Panel B therefore report distinct analytical universes. Appendix Table A.1 defines the selector fields and their source tables.

Table 2: Descriptive Statistics: Firm, CEO, and Performance Characteristics

Variable	N	Mean	SD	Min.	Median	Max.
Panel A. Firm characteristics						
Log assets	415	8.776	1.383	5.654	8.640	12.432
ROA	305	0.030	0.086	-0.358	0.039	0.190
Leverage	415	0.331	0.197	0.000	0.315	0.938
Cash/assets	305	0.063	0.077	0.000354	0.037	0.500
Market-to-book	290	3.860	7.735	0.413	2.211	68.856
Prior 12-month return	414	0.130	0.313	-0.675	0.115	1.767
Prior idiosyncratic volatility	414	0.014	0.008	0.006	0.011	0.055
Pre-period R&D intensity	415	0.006	0.018	0.000	0.000	0.132
Pledging restriction	377	0.653	0.477	0.000	1.000	1.000
Anti-hedging policy	377	0.798	0.402	0.000	1.000	1.000
CEO stock-ownership or holding guideline	431	0.397	0.490	0.000	0.000	1.000
Board discretion over stock awards	431	0.251	0.434	0.000	0.000	1.000
Panel B. CEO characteristics						
CEO tenure (years)	430	5.455	5.443	0.000	4.003	23.285
LLM managerial-myopia score (1-5)	411	2.394	0.593	2.000	2.000	4.000
Cumulative Vesting Intensity	394	0.442	0.789	0.000	0.135	5.131
Cumulative Vesting Share Ratio	421	0.391	0.674	0.000	0.147	4.070
Pre-event CEO common-stock wealth ($ millions)	349	39.223	66.948	0.024	15.446	389.913
Pre-event beneficially owned shares (millions)	398	1.217	3.073	0.000250	0.481	25.668
Predetermined scheduled non-focal vesting (thousand shares)	380	53.701	80.280	0.000	24.373	384.798
Panel C. Performance and assignment characteristics						
Relative-TSR distance from cutoff (p.p.)	431	-0.661	4.557	-7.480	-1.385	7.434
Barely passes cutoff	431	0.418	0.494	0.000	0.000	1.000
Award Jump	274	0.032	0.074	0.000	0.000	0.600

Notes: Statistics use the unrestricted selected h=0.075 sample before outcome-specific restrictions. The unit is one selected award; N is the non-missing award count for the row. Pre-period R&D intensity is the three-year mean of annual R&D/assets. Pledging and anti-hedging indicators are matched to filing year t-1. The source matches 377 of 431 awards and unmatched firm-years remain missing. Source-missing holding requirements are coded no guideline, and absent or unmentioned stock-award discretion is coded no discretion. In Panel B, the LLM myopia score, common-stock wealth, and beneficial ownership are measured at t-1; CEO tenure uses the pre-performance reference date; and non-focal vesting is predetermined. Higher LLM scores indicate greater apparent short-term orientation. Cumulative Vesting Intensity divides valueVesting summed over t and t+1 by common-stock wealth at t-1; Cumulative Vesting Share Ratio divides numVesting over the same years by shares owned at t-1. Missing annual vesting numerators are zero, while a positive numerator with a non-positive or unavailable denominator remains missing. Panel C reports assignment and performance characteristics. Award Jump is treatment times the contract-implied ex ante cutoff jump divided by sharesOwned at t-1 times matched CRSP price at t-1. Continuously valued outcomes and covariates are winsorized at the 1st and 99th percentiles of the pre-2023 universe. Log assets, prior return, and prior idiosyncratic volatility are treatment controls; ROA, leverage, cash/assets, market-to-book, and pre-period R&D intensity are balance covariates only. Exact formulas appear in Appendix Table A.1.

Table 3: Descriptive Statistics: Innovation

Variable	N	Mean	SD	Min.	Median	Max.
Innovation outcomes, t+1 to t+3						
Log(1 + patent grants)	431	0.537	1.164	0.000	0.000	5.878
Log(1 + patent applications)	431	0.647	1.303	0.000	0.000	6.370
Log(1 + mean R&D intensity)	393	0.005	0.016	0.000	0.000	0.138
Adjusted patent originality	106	0.446	0.257	0.000	0.474	1.000
Adjusted patent generality	68	0.353	0.246	0.000	0.382	1.000
Technological breadth: distinct CPC subclasses	431	2.303	7.822	0.000	0.000	68.047
Log(1 + Real Economic Value)	431	0.643	1.845	0.000	0.000	9.225
Log(1 + Forward Citations)	431	0.274	1.040	0.000	0.000	6.087

Notes: The unit is one selected award in the maintained h=0.075 sample and N is the number of nonmissing award observations for each row. Grants and applications are annual patent counts summed over t+1 through t+3 and transformed as log(1+x). Mean R&D intensity is the arithmetic mean of annual R&D spending/assets over observed t+1 through t+3 years and is then transformed as log(1+x). Adjusted originality and generality are application-support-weighted means of their annual finite-sample-adjusted patent indexes; technological breadth is the equal-year mean of annual distinct-CPC-subclass counts. Real Economic Value and Forward Citations are annual source totals summed over t+1 through t+3 and transformed as log(1+x). Missing source values are excluded variable by variable and documented no-filing years retain structural zeros. The reported minimum and maximum are after the maintained 1%/99% winsorization of continuous analysis variables. Appendix Table A.1 gives the exact formulas.

Table 4: Predetermined Covariate Continuity at the Relative-TSR Cutoff

Predetermined covariate	Coefficient	Clustered SE	p-value	Observations
Firm size (log assets)	-0.354	(0.383)	0.355	415
ROA	0.006	(0.024)	0.814	305
Leverage	0.002	(0.055)	0.971	415
Cash/assets	0.026	(0.018)	0.148	305
Market-to-book	1.409	(1.408)	0.317	290
Prior 12-month return	0.100	(0.088)	0.255	414
Prior idiosyncratic volatility	-0.000314	(0.002)	0.899	414
Pre-period R&D intensity	0.002	(0.005)	0.688	415
Log(1 + pre-period patent applications)	-0.060	(0.437)	0.891	417
Joint standardized Wald test	$\chi^2(9) = 7.953$	–	0.539	417
Specification indicators				
Separate local slopes		Yes		
Fixed effects		No		
Additional covariates		No		

Notes: Each covariate row is a separate award-level triangular-kernel local-linear RD within $h = 0.075$. The dependent variable is the predetermined covariate named in that row. The regression contains Barely passes cutoff, the centered Relative-TSR running variable, and their interaction, which permits separate slopes below and above zero. The selector first applies the bandwidth and then retains one cutoff per award under the target-threshold-maximum priority. Continuous covariates are winsorized at 1%/99%; the running variable and treatment indicator are not. Standard errors are clustered by firm, and observations are award observations with non-missing values for the row covariate. The final row jointly tests that all 9 treatment discontinuities equal zero in one stacked regression after standardizing each covariate on its own displayed complete-case sample. It uses the same separate slopes, triangular weights, no fixed effects, no additional covariates, and firm clustering as the individual rows. Its observation entry is the 417 distinct awards represented across 3,390 stacked covariate-award records. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Exact covariate formulas appear in Appendix Table A.1.

Table 5: Economic Magnitude and CEO Vesting Intensity Around Cutoff Passage

Outcome	Main	Different slopes	Simple FE	Rich FE	Full controls
Panel A. Award Jump					
Barely passes cutoff, τ	0.096***	0.099***	0.091***	0.091***	0.085***
	(0.017)	(0.019)	(0.016)	(0.016)	(0.014)
Control-side mean	0.000	0.000	0.000	0.000	0.000
Observations	221	221	221	221	221
Panel B. Cumulative Vesting Intensity					
Barely passes cutoff, τ	0.271**	0.251**	0.234*	0.268*	0.244*
	(0.123)	(0.122)	(0.141)	(0.146)	(0.142)
Control-side mean	0.437	0.437	0.437	0.437	0.437
Observations	338	338	338	338	338
Panel C. Cumulative Vesting Share Ratio					
Barely passes cutoff, τ	0.250**	0.230**	0.209*	0.208*	0.221**
	(0.102)	(0.102)	(0.111)	(0.113)	(0.112)
Control-side mean	0.391	0.391	0.391	0.391	0.391
Observations	356	356	356	356	356
Specification indicators					
Separate local slopes	No	Yes	Yes	Yes	Yes
Performance-realization-year FE	No	No	Yes	Yes	Yes
FF12-industry FE	No	No	Yes	Yes	Yes
Cutoff-type FE	No	No	No	Yes	Yes
Interpolation-status FE	No	No	No	Yes	Yes
Outcome-specific controls	No	No	No	No	Yes

Notes: Each panel reports the discontinuity at zero for Barely passes cutoff from an award-level Sharp RD. All five columns within a panel use the common complete-case sample required by that panel's Full controls specification. Main contains treatment and a common local-linear running-variable slope. Different slopes adds the treatment-by-running-variable interaction. Simple FE adds performance-realization-year and FF12-industry fixed effects; Rich FE adds cutoff-type and interpolation-status fixed effects; Full controls adds log assets, prior 12-month return, prior idiosyncratic volatility, and the documented mechanism-specific predetermined controls. The Award Jump full-control model additionally includes directly priceable pre-event CEO common-stock wealth and predetermined scheduled nonfocal vesting; both vesting-ratio models additionally include predetermined scheduled nonfocal vesting. Award Jump is the treatment indicator times the contract-implied ex ante cutoff jump divided by positive directly priceable pre-event CEO common-stock wealth, sharesOwned at t-1 times the matched CRSP price at t-1. Cumulative Vesting Intensity is vesting value summed over t and t+1 divided by the same directly priceable CEO common-stock wealth. Cumulative Vesting Share Ratio is vesting shares summed over t and t+1 divided by pre-event shares owned. Missing annual vesting numerators are coded zero. A zero cumulative numerator yields a zero ratio, while positive vesting requires a strictly positive denominator. All outcomes use weighted least squares, h=0.075, triangular kernel weights, the one-cutoff-per-award selector, 1%/99% winsorization, and firm-clustered inference. Cells combine the coefficient and clustered standard error. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Appendix Table A.1 gives the exact formulas.

Table 6: Innovation Output and Input

Outcome	Main	Different slopes	Simple FE	Rich FE	Full controls
Panel A. Log(1 + patent grants)					
Barely passes cutoff, τ	0.267	0.251	0.530**	0.518**	0.528**
	(0.265)	(0.287)	(0.222)	(0.217)	(0.208)
Observations	414	414	414	414	414
Panel B. Log(1 + patent applications)					
Barely passes cutoff, τ	0.204	0.184	0.481*	0.467*	0.481*
	(0.309)	(0.328)	(0.268)	(0.267)	(0.246)
Observations	414	414	414	414	414
Panel C. Log(1 + mean R&D intensity)					
Barely passes cutoff, τ	0.002	0.002	0.001	0.001	0.001
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Observations	388	388	388	388	388
Specification indicators					
Separate local slopes	No	Yes	Yes	Yes	Yes
Performance-realization-year FE	No	No	Yes	Yes	Yes
FF12-industry FE	No	No	Yes	Yes	Yes
Cutoff-type FE	No	No	No	Yes	Yes
Interpolation-status FE	No	No	No	Yes	Yes
Outcome-specific controls	No	No	No	No	Yes

Full-controls horizon comparison Outcome	$t+1$	$t+1:t+2$	$t+1:t+3$
Log(1 + patent grants)	0.352**	0.451**	0.528**
	(0.169)	(0.198)	(0.208)
	[$N=414$]	[$N=414$]	[$N=414$]
Log(1 + patent applications)	0.320	0.400*	0.481*
	(0.203)	(0.234)	(0.246)
	[$N=414$]	[$N=414$]	[$N=414$]
Log(1 + mean R&D intensity)	0.002	0.001	0.001
	(0.002)	(0.002)	(0.002)
	[$N=388$]	[$N=388$]	[$N=388$]

Notes: Each coefficient is the estimated discontinuity at zero for Barely passes cutoff in an award-level Sharp RD. All columns within an outcome panel use the common complete-case sample required by that panel's Full controls column, so changes across the ladder reflect specification rather than sample composition. Main includes D and one local-linear running-variable slope. Different slopes adds D x X. Simple FE retains the separate slopes and adds performance-realization-year and FF12-industry fixed effects. Rich FE further adds cutoff-type and interpolation-status fixed effects. Full controls further adds log assets, prior 12-month return, prior 12-month idiosyncratic volatility, and any documented outcome-specific predetermined control (the matching pre-period adjusted patent attribute for originality or generality). All regressions use the one-cutoff-per-award sample selected after bandwidth filtering, h=0.075, triangular kernel weights, and standard errors clustered by firm. Continuous analysis variables are winsorized at 1%/99%. Cells report the coefficient with the firm-clustered standard error in parentheses and observations are award observations. The horizon block reports only the Full controls specification and states its horizon in the column heading. Stars denote $^{***}p<0.01$, $^{**}p<0.05$, and $^{*}p<0.10$. Exact constructions appear in Appendix Table A.1.

Table 7: Originality and Technological Breadth

Outcome	Main	Different slopes	Simple FE	Rich FE	Full controls
Panel A. Adjusted patent originality					
Barely passes cutoff, τ	0.213	0.165	0.608***	0.587***	0.462**
	(0.131)	(0.117)	(0.178)	(0.174)	(0.222)
Observations	91	91	91	91	91
Panel B. Technological breadth: distinct CPC subclasses					
Barely passes cutoff, τ	3.149	3.371	4.074**	3.988**	4.195**
	(2.426)	(2.733)	(1.981)	(1.890)	(1.897)
Observations	414	414	414	414	414
Specification indicators					
Separate local slopes	No	Yes	Yes	Yes	Yes
Performance-realization-year FE	No	No	Yes	Yes	Yes
FF12-industry FE	No	No	Yes	Yes	Yes
Cutoff-type FE	No	No	No	Yes	Yes
Interpolation-status FE	No	No	No	Yes	Yes
Outcome-specific controls	No	No	No	No	Yes

Full-controls horizon comparison Outcome	$t+1$	$t+2$	$t+1:t+3$
Adjusted originality	0.294	0.501***	0.462**
	(0.271)	(0.189)	(0.222)
	[$N=81$]	[$N=65$]	[$N=91$]
Technological breadth	3.773**	4.179**	4.195**
	(1.874)	(1.898)	(1.897)
	[$N=414$]	[$N=414$]	[$N=414$]

Notes: Each coefficient is the estimated discontinuity at zero for Barely passes cutoff in an award-level Sharp RD. All columns within an outcome panel use the common complete-case sample required by that panel's Full controls column, so changes across the ladder reflect specification rather than sample composition. Main includes D and one local-linear running-variable slope. Different slopes adds D x X. Simple FE retains the separate slopes and adds performance-realization-year and FF12-industry fixed effects. Rich FE further adds cutoff-type and interpolation-status fixed effects. Full controls further adds log assets, prior 12-month return, prior 12-month idiosyncratic volatility, and any documented outcome-specific predetermined control (the matching pre-period adjusted patent attribute for originality). All regressions use the one-cutoff-per-award sample selected after bandwidth filtering, h=0.075, triangular kernel weights, and standard errors clustered by firm. Continuous analysis variables are winsorized at 1%/99%. Cells report the coefficient with the firm-clustered standard error in parentheses and observations are award observations. The horizon block reports only the Full controls specification and states its horizon in the column heading. Stars denote $^{***}p<0.01$, $^{**}p<0.05$, and $^{*}p<0.10$. Exact constructions appear in Appendix Table A.1.

Table 8: Patent Real Economic Value and Forward Citations

Outcome	Main	Different slopes	Simple FE	Rich FE	Full controls
Panel A. Log(1 + Real Economic Value)					
Barely passes cutoff, τ	0.485	0.493	0.801**	0.800**	0.853***
	(0.409)	(0.441)	(0.348)	(0.335)	(0.298)
Observations	414	414	414	414	414
Panel B. Log(1 + Forward Citations)					
Barely passes cutoff, τ	0.327	0.361	0.452**	0.443**	0.453**
	(0.263)	(0.292)	(0.191)	(0.184)	(0.185)
Observations	414	414	414	414	414
Specification indicators					
Separate local slopes	No	Yes	Yes	Yes	Yes
Performance-realization-year FE	No	No	Yes	Yes	Yes
FF12-industry FE	No	No	Yes	Yes	Yes
Cutoff-type FE	No	No	No	Yes	Yes
Interpolation-status FE	No	No	No	Yes	Yes
Outcome-specific controls	No	No	No	No	Yes

Full-controls horizon comparison Outcome	$t+1$	$t+1:t+2$	$t+1:t+3$
Log(1 + Real Economic Value)	0.649**	0.731**	0.853***
	(0.276)	(0.295)	(0.298)
	[$N=414$]	[$N=414$]	[$N=414$]
Log(1 + Forward Citations)	0.426***	0.459***	0.453**
	(0.160)	(0.177)	(0.185)
	[$N=414$]	[$N=414$]	[$N=414$]

Notes: Each coefficient is the estimated discontinuity at zero for Barely passes cutoff in an award-level Sharp RD. All columns within an outcome panel use the common complete-case sample required by that panel's Full controls column, so changes across the ladder reflect specification rather than sample composition. Main includes D and one local-linear running-variable slope. Different slopes adds D x X. Simple FE retains the separate slopes and adds performance-realization-year and FF12-industry fixed effects. Rich FE further adds cutoff-type and interpolation-status fixed effects. Full controls further adds log assets, prior 12-month return, prior 12-month idiosyncratic volatility, and any documented outcome-specific predetermined control (the matching pre-period adjusted patent attribute for originality or generality). All regressions use the one-cutoff-per-award sample selected after bandwidth filtering, h=0.075, triangular kernel weights, and standard errors clustered by firm. Continuous analysis variables are winsorized at 1%/99%. Cells report the coefficient with the firm-clustered standard error in parentheses and observations are award observations. The horizon block reports only the Full controls specification and states its horizon in the column heading. Stars denote $^{***}p<0.01$, $^{**}p<0.05$, and $^{*}p<0.10$. Exact constructions appear in Appendix Table A.1.

Table 9: Bandwidth Robustness

Outcome	5%	6%	7%	7.5%	8%	9%	10%
Log(1 + patent grants)	0.532**	0.541**	0.570***	0.528**	0.506***	0.398**	0.361**
	(0.258)	(0.224)	(0.215)	(0.208)	(0.196)	(0.183)	(0.180)
	[$N = 264$]	[$N = 326$]	[$N = 386$]	[$N = 414$]	[$N = 443$]	[$N = 481$]	[$N = 516$]
Log(1 + patent applications)	0.445	0.483*	0.514**	0.481*	0.468**	0.369*	0.340
	(0.307)	(0.266)	(0.254)	(0.246)	(0.229)	(0.211)	(0.208)
	[$N = 264$]	[$N = 326$]	[$N = 386$]	[$N = 414$]	[$N = 443$]	[$N = 481$]	[$N = 516$]
Log(1 + Real Economic Value)	0.690*	0.752**	0.889***	0.853***	0.864***	0.743***	0.684**
	(0.367)	(0.314)	(0.307)	(0.298)	(0.287)	(0.282)	(0.281)
	[$N = 264$]	[$N = 326$]	[$N = 386$]	[$N = 414$]	[$N = 443$]	[$N = 481$]	[$N = 516$]
Log(1 + Forward Citations)	0.480**	0.466**	0.495**	0.453**	0.419**	0.322*	0.305*
	(0.233)	(0.196)	(0.192)	(0.185)	(0.180)	(0.169)	(0.172)
	[$N = 264$]	[$N = 326$]	[$N = 386$]	[$N = 414$]	[$N = 443$]	[$N = 481$]	[$N = 516$]
Specification indicators							
Separate local slopes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Performance-realization-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FF12-industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cutoff-type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interpolation-status FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Full outcome-specific controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell is the discontinuity on Barely passes cutoff from a separate award-level Sharp RD. Columns report symmetric bandwidths from 5% to 10%; the maintained 7.5% column is shown in the center. Each cell reports the coefficient, firm-clustered standard error in parentheses, and the number of observations. For every bandwidth, filtering precedes target-first one-cutoff-per-award selection. All models use triangular-kernel local-linear separate slopes, performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, and full outcome-specific controls. The common controls are log assets, prior 12-month return, and prior idiosyncratic volatility; outcome-specific controls follow Appendix Table A.1. Continuous analysis variables are winsorized at 1%/99%. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

Table 10: Functional Form and Sample-Definition Robustness

Specification	Log(1 + patent grants)	Log(1 + patent applications)	Log(1 + Real Economic Value)	Log(1 + Forward Citations)
Panel A. Kernel and polynomial robustness				
Triangular, local linear	0.528**	0.481*	0.853***	0.453**
	(0.208)	(0.246)	(0.298)	(0.185)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Uniform, local linear	0.384*	0.381*	0.773**	0.319*
	(0.198)	(0.229)	(0.305)	(0.185)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Epanechnikov, local linear	0.527**	0.497**	0.875***	0.435**
	(0.206)	(0.241)	(0.300)	(0.184)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Triangular, local quadratic	0.761**	0.607*	0.848*	0.663**
	(0.300)	(0.360)	(0.438)	(0.266)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Uniform, local quadratic	0.904***	0.775**	1.015**	0.764***
	(0.316)	(0.365)	(0.438)	(0.280)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Epanechnikov, local quadratic	0.829***	0.694*	0.935**	0.717***
	(0.308)	(0.362)	(0.443)	(0.275)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Panel B. Running-variable certainty and sample restrictions				
Unrestricted pooled main sample	0.528**	0.481*	0.853***	0.453**
	(0.208)	(0.246)	(0.298)	(0.185)
	[$N = 414$]	[$N = 414$]	[$N = 414$]	[$N = 414$]
Excluding disclosed interpolation	0.631***	0.559*	0.990***	0.471**
	(0.241)	(0.288)	(0.332)	(0.226)
	[$N = 311$]	[$N = 311$]	[$N = 311$]	[$N = 311$]
Exact-point observations only	0.451*	0.430	0.876**	0.486**
	(0.236)	(0.266)	(0.371)	(0.220)
	[$N = 240$]	[$N = 240$]	[$N = 240$]	[$N = 240$]
Specification indicators for every row				
Separate local slopes		Yes		
All four fixed effects		Yes		
Full outcome-specific controls		Yes		

Notes: Cells report the discontinuity on Barely passes cutoff, the firm-clustered standard error in parentheses, and the number of award observations. Panel A changes the kernel and local-polynomial order while retaining the preferred sample: local-linear rows include X and D x X; local-quadratic rows add X-squared and D x X-squared. Panel B changes only the running-variable source sample. Every estimate uses the full specification with performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects plus the full outcome-specific controls. The exact-point restriction uses only observations with degenerate stored running-variable intervals. Kernel weights are recomputed within h=0.075. Continuous analysis variables are winsorized at 1%/99%, and Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

Table 11: Managerial Short-Termism Heterogeneity

Group	Log(1 + patent grants)	Log(1 + patent applications)	Log(1 + Real Economic Value)	Log(1 + Forward Citations)
Panel A. CEO tenure				
Short-tenure CEOs	0.075 (0.287) [$N = 200$]	-0.069 (0.360) [$N = 200$]	0.126 (0.450) [$N = 200$]	-0.023 (0.273) [$N = 200$]
Long-tenure CEOs	0.839*** (0.308) [$N = 213$]	0.851** (0.342) [$N = 213$]	1.259** (0.519) [$N = 213$]	0.697** (0.304) [$N = 213$]
Panel B. Exercisable / unexercisable option delta				
Low option-delta ratio	-0.677 (1.208) [$N = 53$]	-1.713 (1.304) [$N = 53$]	0.947 (1.780) [$N = 53$]	0.341 (1.161) [$N = 53$]
High option-delta ratio	1.923*** (0.718) [$N = 53$]	2.015*** (0.780) [$N = 53$]	1.023 (1.561) [$N = 53$]	1.235** (0.625) [$N = 53$]
Panel C. Broad vested-proxy / unvested delta				
Low vested/unvested delta	-0.399 (0.362) [$N = 124$]	-0.582 (0.427) [$N = 124$]	-0.134 (0.630) [$N = 124$]	-0.257 (0.371) [$N = 124$]
High vested/unvested delta	1.158*** (0.405) [$N = 123$]	1.004*** (0.389) [$N = 123$]	1.470** (0.710) [$N = 123$]	0.656** (0.297) [$N = 123$]
Panel D. LLM managerial myopia				
Low LLM myopia	0.242 (0.260) [$N = 263$]	0.155 (0.319) [$N = 263$]	0.734* (0.377) [$N = 263$]	0.406* (0.239) [$N = 263$]
High LLM myopia	1.057** (0.460) [$N = 132$]	1.153** (0.487) [$N = 132$]	1.594** (0.731) [$N = 132$]	0.849* (0.487) [$N = 132$]

Notes: Each cell is the Barely passes cutoff coefficient from a separate median split-sample Sharp RD. The firm-clustered standard error and number of award observations appear beneath the estimate. Every subgroup regression preserves h=0.075, triangular weights, separate local-linear slopes, performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, plus log assets, prior 12-month return, and prior idiosyncratic volatility. Continuous moderators are measured at t-1, winsorized at 1%/99%, and split at the maintained selected-sample median. Formal cross-group claims require pooled interaction/equality tests in the source report. Panel thresholds are tenure 4.003 years, option ratio 0.000, broad vested/unvested ratio 1.012, and LLM myopia score 2.000. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Moderator formulas appear in Appendix Table A.1.

Table 12: Board Discretion: Prevalence and RD Results

Panel A. Descriptive frequency

Category	Awards	Awards (%)
No stock-award discretion	323	74.9
Any stock-award discretion	108	25.1
Positive discretion	3	0.7
Negative discretion	31	7.2
Two-sided discretion	20	4.6
Direction unstated / other	54	12.5

Panel B. RD estimates by stock-award discretion status

Group	Log(1 + patent grants)	Log(1 + patent applications)	Log(1 + Real Economic Value)	Log(1 + Forward Citations)
No stock-award discretion	0.603**	0.593**	0.975***	0.442**
	(0.242)	(0.285)	(0.366)	(0.221)
	[$N = 309$]	[$N = 309$]	[$N = 309$]	[$N = 309$]
Stock-award discretion	0.032	-0.217	-0.409	-0.167
	(0.278)	(0.354)	(0.319)	(0.213)
	[$N = 105$]	[$N = 105$]	[$N = 105$]	[$N = 105$]

Notes: Panel A reports raw award counts in the selected h=0.075 sample. Stock-award discretion equals one if the grant-year source identifies discretion over stock awards alone or stock awards and bonus; absent or unmentioned discretion is coded no discretion under the requested rule. Positive, negative, and two-sided categories come from the source direction fields and are mutually exclusive after deterministic resolution. Panel B reports full-specification split-sample RD estimates with the coefficient, firm-clustered standard error, and number of observations. Every regression uses separate local-linear slopes, the four maintained fixed effects, and the full outcome-specific controls within h=0.075 with triangular weights. Continuous analysis variables are winsorized at 1%/99%. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Appendix Table A.1 gives the variable definitions.

Table 13: Does Narrow Cutoff Passage Change Proposed CEO Selling or Share Pledging?

Outcome	Coefficient	Clustered SE	Observations
Panel A. Proposed Form 144 selling			
Log(1 + proposed Form 144 sale value)	0.025	(1.381)	413
Any proposed Form 144 sale	0.003	(0.090)	414
Panel B. CEO share pledging			
Any CEO share pledge	0.067	(0.045)	414
Log(1 + mean annual CEO shares pledged)	0.546	(0.626)	414

Notes: Each coefficient is the Barely passes cutoff discontinuity from an award-level Sharp RD. Form 144 outcomes measure proposed selling, not realized transactions or sale proceeds, and require complete five-year source-window coverage. All four outcomes aggregate the first two post-realization years. Every row uses the maintained h=0.075 triangular-kernel local-linear RD with separate slopes, performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, and controls for log assets, prior 12-month return, and prior idiosyncratic volatility. Pledge outcomes additionally control for the matching pre-event pledge measure. Standard errors are clustered by firm; observations are award observations. Continuous variables are winsorized at 1%/99%. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Exact outcome and control constructions appear in Appendix Table A.1.

Table 14: Policy and Holding-Requirement Heterogeneity

Group	Log(1 + patent grants)	Log(1 + patent applications)	Log(1 + Real Economic Value)	Log(1 + Forward Citations)
Panel A. Pledging restriction				
No disclosed pledging restriction	0.115 (0.410) [$N = 125$]	-0.094 (0.493) [$N = 125$]	0.193 (0.565) [$N = 125$]	0.237 (0.358) [$N = 125$]
Disclosed pledging restriction	0.460*** (0.137) [$N = 240$]	0.459*** (0.163) [$N = 240$]	0.425** (0.208) [$N = 240$]	0.140 (0.086) [$N = 240$]
Panel B. Anti-hedging policy				
No disclosed anti-hedging policy	0.908 (0.830) [$N = 69$]	1.081 (0.797) [$N = 69$]	1.157 (0.956) [$N = 69$]	1.138 (0.818) [$N = 69$]
Disclosed anti-hedging policy	0.489** (0.195) [$N = 296$]	0.424* (0.235) [$N = 296$]	0.633** (0.275) [$N = 296$]	0.310* (0.187) [$N = 296$]
Panel C. CEO holding-requirement status				
No holding requirement (missing=0)	0.304** (0.148) [$N = 252$]	0.189 (0.208) [$N = 252$]	0.529** (0.242) [$N = 252$]	0.266** (0.107) [$N = 252$]
Holding requirement met in full	0.671** (0.316) [$N = 222$]	0.730** (0.344) [$N = 222$]	1.021** (0.466) [$N = 222$]	0.477 (0.295) [$N = 222$]

Notes: Each cell is the Barely passes cutoff coefficient from a separate binary split-sample Sharp RD; the firm-clustered standard error and number of award observations appear beneath the estimate. The policy moderator is predetermined: each selected award is matched by normalized firm CIK to the exact DEF 14A filing year t-1. Among duplicate firm-year filings, a policy equals one if any supplied filing discloses the restriction. The source matches 377 of 431 selected awards; unmatched firm-years remain missing and are excluded rather than coded as no policy. Panel A compares disclosed pledging restrictions with no disclosed pledging restriction; Panel B analogously compares anti-hedging policies. A zero means no qualifying restriction was disclosed in the supplied filing and is not proof that no policy existed elsewhere. Pledging restrictions are policy constraints, not actual CEO pledging, and anti-hedging policies cover company securities rather than operating hedges. Panel C reports the two holding-status groups from the full-specification split analysis. Source-missing holding-requirement disclosures are coded zero/no requirement, whereas met in full requires observed full compliance at performance realization. Holding subsamples are separate, overlapping status codings rather than complementary subsamples. 134 of 431 selected awards satisfy both before outcome-specific complete-case filters. Every subgroup regression preserves h=0.075, triangular weights, separate local-linear slopes, performance-realization-year, FF12-industry, cutoff-type, and interpolation-status fixed effects, and the maintained controls: log assets, prior 12-month return, and prior idiosyncratic volatility. Continuous analysis variables are winsorized at 1%/99%. Stars denote $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$. Appendix Table A.1 defines the policy moderators.

Internet Appendix for

"Taking Stock: Do Equity-Based Managerial Incentives Drive Corporate Innovation?"

August 2026

This Internet Appendix describes the construction of the measures used in the paper. Section IA.1 documents how we extract the calculation details of relative total shareholder return (Relative-TSR) performance awards and convert them into contract-specific performance cutoffs, combining large language model (LLM) interpretation of disclosure text with an agentic artificial intelligence (AI) workflow in which all return calculations are deterministic. Section IA.2 documents the agentic AI procedure that links patent-application assignees to public firms. Section IA.3 documents the compensation-committee discretion, CEO share-pledging, pledging-restriction, anti-hedging, and managerial-myopia measures constructed through an LLM batch-classification process. Section IA.4 documents the construction of CEO portfolio delta and vega from ISS Incentive Lab holdings data.

Introduction

The paper's central variables cannot be looked up in any standard database. Whether a CEO narrowly cleared a Relative-TSR payout cutoff, which public firm stands behind a patent application filed under a subsidiary's name, and whether a proxy statement grants the committee discretion, restricts pledging, or reveals a short-term managerial orientation are all facts that live in text: in analyst notes, proxy footnotes, benchmark descriptions, ownership tables, and management commentary, each written in its own firm-specific language. Two contracts that mean exactly the same thing rarely say it the same way. This appendix explains, in one place, how we turn that text into research variables that are precise enough to support a regression-discontinuity design, and why the resulting measures can be trusted.

Our approach rests on a simple division of labor: language models read, and code calculates. For the two hardest tasks, reconstructing each award's contractual calculation rule (Section IA.1) and resolving patent-assignee identities (Section IA.2), we use *agentic AI workflows*, in which large language models act as reviewing agents inside a deterministic pipeline. Each agent receives a narrowly defined bundle of evidence, reaches a structured decision under a restricted rubric, and records the evidence behind that decision; where a task requires information beyond the bundle, as in tracing a subsidiary to its public parent, the agent searches external sources, and difficult cases are judged by several independent agents whose conclusions are reconciled. For the filing-based governance and managerial measures (Section IA.3), we use a simpler *LLM batch process*: every filing is sent to the model as its own request, under a prompt that asks one narrow question with a fixed set of admissible answers, so that tens of thousands of filings are classified uniformly and no filing can influence the classification of another. The CEO portfolio delta and vega measures (Section IA.4) involve no language model at all; they are standard Black–Scholes–Merton calculations from structured holdings and market data.

Three safeguards make the measures dependable. First, models never touch a number that matters: every return, rank, cutoff, and sensitivity is computed by deterministic code after the model's interpretive decision is fixed. Second, interpretation is blind to outcomes: no model ever sees a realized return while reading a contract, so no reading can be chosen, consciously or not, because it places a firm on a favorable side of a cutoff. Third, uncertainty is preserved rather than papered over: when the evidence does not support a value, the record stays unresolved instead of being filled with a zero or an average, and every judgment retains its evidence and provenance so that it can be re-examined and stress-tested. Table IA.1 summarizes the four measurement blocks and the division of labor in each.

Table IA.1. Overview of the Measurement Procedures

Component	Measurement challenge	AI stage	Deterministic stage
Relative-TSR cutoffs (Section IA.1)	Contract terms are distributed across provider fields, proxy text, comments, and benchmark descriptions.	Agent review of deduplicated evidence signatures recovers the comparison rule, return basis, averaging windows, benchmark, and supported assumptions.	Map securities, measure award-specific TSR, invert the contractual hurdle, construct exact or bounded running variables, and validate the arithmetic.
Patent–firm linkage (Section IA.2)	Application assignees may be subsidiaries, acquired entities, or operating companies whose names differ from the listed parent.	Agents adjudicate non-perfect candidate matches using corporate-family and public-source evidence; multiple agents judge difficult records.	Generate ranked fuzzy candidates, accept literal perfect matches, attach the accepted CIK, and retain match provenance.
Filing-based measures (Section IA.3)	Proxy and 10-K language mixes behavior, policy, compensation instruments, and statements about persons other than the CEO.	Batch LLM classification applies restricted rubrics to committee discretion, CEO pledging, pledging restrictions, anti-hedging policies, and managerial horizon.	Parse structured output, standardize identifiers and dates, resolve duplicate source routes, and retain evidence and missingness states.
CEO delta and vega (Section IA.4)	Holdings must be aligned to event time, fiscal-year ends, prices, volatility, dividends, and interest rates.	None; no language-model input enters the numerical sensitivities.	Perform fiscal-date matching, Black–Scholes–Merton valuation, tranche aggregation, component validation, and portfolio decomposition.

Notes: This table summarizes the four measurement blocks documented in this appendix. The AI stage determines an economic rule or identity; it does not calculate realized returns, choose an interpretation according to its distance from a cutoff, or impute unresolved numerical inputs. All reported numerical measures are produced by deterministic code after the interpretive decision has been fixed.

IA.1 Extracting the Relative-TSR Performance-Measure Calculation Details

For each Relative-TSR award, we recover the focal-firm total shareholder return that would have been just sufficient to attain a disclosed threshold, target, or maximum under the award's comparison rule. The comparator may be an observed peer at a specified percentile or ordinal rank, a peer mean or median, a named index, or a benchmark entering through a percentage-point or ratio formula. For each contractual cutoff, we measure the focal firm's realized TSR, construct the required focal-firm TSR, and define their signed difference as the running variable. This contract-level treatment is consistent with the institutional evidence on performance-vesting provisions in Bettis et al. (2018) and De Angelis and Grinstein (2015).

The starting universe contains 6,528 CEO relative-performance RSU awards issued by 1,160 firms between 1998 and 2024. ISS Incentive Lab supplies grant, executive, objective, performance-period, payout-goal, and disclosed-peer information. SEC proxy evidence is merged by exact normalized CIK and calendar year and supplements, rather than overwrites, the Incentive Lab fields. CRSP daily total returns provide the focal and comparator return paths, and historical CRSP and CRSP–Compustat identifiers provide date-valid links to PERMNO. The resulting measurement panel contains 4,328 grant–objective–period–cutoff observations across 1,750 grants and 397 firms. Of these observations, 2,435 have point-identified cutoffs and 1,893 have finite two-sided cutoff intervals caused by incomplete peer coverage.

The construction is organized as a deterministic–agentic production architecture, summarized in Figure IA.1. Section IA.1.1 describes the LLM stage that recovers each award's contractual calculation rule from disclosure text. Section IA.1.2 describes the deterministic stages of the agentic workflow that measure returns, invert benchmarks, and standardize every contract into common units. Section IA.1.3 describes the treatment of missing peers.

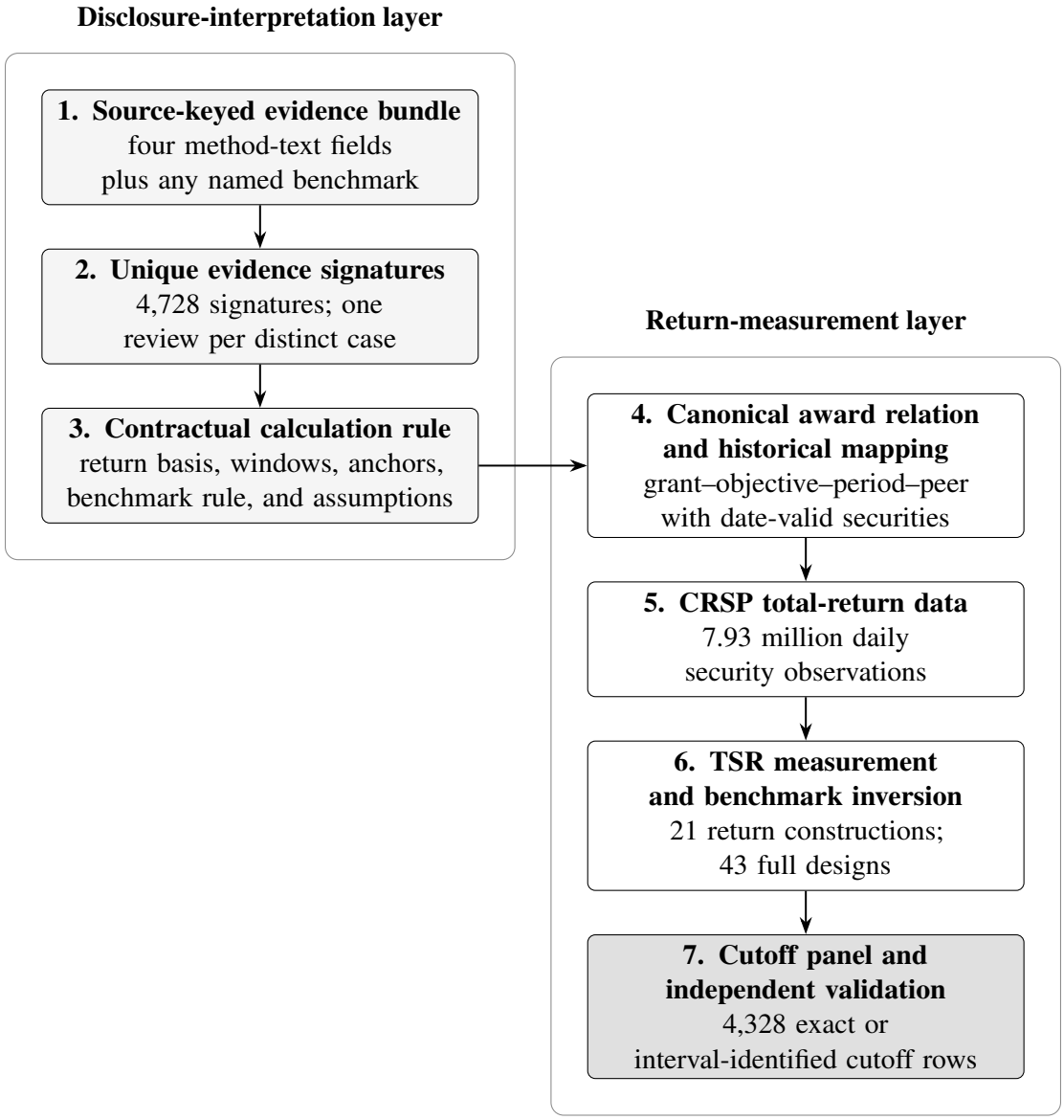


Figure IA.1. Construction Sequence for Relative-TSR Cutoffs. This figure summarizes the deterministic–agentic production architecture. The evidence bundle contains contract instructions, not realized returns. Agent review is confined to deduplicated evidence signatures and produces structured method specifications; it does not process daily returns, choose among methods based on proximity to zero, or alter peer membership. Security mapping, daily-return arithmetic, award-specific window construction, peer aggregation, benchmark inversion, missing-peer bounds, and validation are deterministic and are fixed before the focal firm's distance from any cutoff is examined.

IA.1.1 Recovering the Contractual Calculation Rule with Large Language Models

Evidence compilation. The relevant contract terms are rarely stored in one standardized field. One source may identify the metric as percentile-based, another may state that TSR includes

reinvested dividends, a proxy footnote may define a 20- or 60-trading-day averaging window, and a benchmark field may name the peer group or index. Reading only one field can therefore produce an economically different calculation from the one described by the contract.

We first assemble, with deterministic code, a source-keyed evidence bundle for each award from four method-text fields,

`relative_objective_analyst_comments`

`relative_compare_method_other`

`tsr_measurement_method_and_averaging_window`

`objective_summary_compare_methods`

and, when it identifies a contractual index or benchmark series, a separately normalized named-benchmark field. The evidence bundle contains contractual instructions and no realized-return information. Blank values and provider placeholders are removed, but every informative statement remains attached to its original source field, so that "Percentile" can be recognized as an order-statistic rule while "20 trading days preceding the beginning" is recognized as a return-measurement window. Table IA.2 lists the contractual attributes that the interpretation stage recovers from this bundle.

Grouping identical contractual evidence for agent review. We collapse identical source-labeled evidence bundles into unique method signatures by hashing a canonical serialization of the retained evidence. Two awards share a signature only when the contractual evidence used for interpretation is identical; grant identifiers, hurdle values, peer returns, and the observed distance from any cutoff play no role in this grouping. The 6,528-award universe contains 4,728 distinct signatures. Forty-six signatures inherit previously adjudicated rules, and 589 are resolved by deterministic parsing without model input. The 4,139 signatures not resolved deterministically, representing 5,792 award rows, are reviewed by the language-model agent once per unique signature rather than once per award. The review resolves or partially resolves 3,658 signatures; 481 signatures, representing 642 award rows, remain unresolved.

Each resolved case yields a structured calculation rule and a concise explanation of the supporting evidence. The retained attributes include the comparison family, return basis, annualization, beginning and ending windows, window units and anchors, comparator aggregation, focal-firm inclusion convention, named benchmark, confidence, assumption flag, and unresolved reason, and the final record retains the signature hash, rule identifier, and inference source so that every interpretation can be traced to its evidence. Reviewing identical evidence only once materially reduces redundant interpretation: the 1,798 grant–objective–period observations in the measurement sample contain 671 distinct source-text signatures, but only 21 effective return-measurement constructions

Table IA.2. Contractual Attributes Recovered from the Evidence Bundle

Attribute	Examples of supported values	Economic role
Performance metric and return basis	Total shareholder return; price appreciation excluding dividends; annualized return	Defines the security-level return to be measured.
Comparison class	Percentile; ordinal rank; peer mean; peer median; percentage-point spread; percentage ratio; named index	Determines how peer or index performance enters the hurdle.
Performance period	Exact start and end dates; multiple objective-specific periods	Defines the return window and realization year.
Averaging windows	Trading-day or calendar-day windows; beginning and ending anchors	Determines whether endpoint wealth or average wealth is compared.
Peer aggregation	Marginal order statistic; equal-weight mean; median; index series	Defines the comparator after security-level returns are measured.
Focal-firm inclusion	Focal excluded from or included in the ranked population	Can change the relevant order statistic in rank-based awards.
Contractual goals	Threshold, target, and maximum; percentile, rank, spread, or ratio units	Supplies the hurdle that is inverted into focal-TSR units.
Inference metadata	Evidence source, rule identifier, confidence, assumption flag, unresolved reason	Makes the method decision reviewable and supports sensitivity tests.

Notes: This table lists the structured attributes produced by the disclosure-interpretation stage for each Relative-TSR award. The evidence bundle supplied to the reviewing agent is assembled deterministically from four Incentive Lab and SEC method-text fields and, where applicable, a normalized named-benchmark field. The bundle contains contractual instructions only; realized returns never enter the interpretation stage.

and 43 full calculation designs. The review rubric is deliberately conservative. A generic statement that TSR is measured relative to peers is not treated as evidence of a peer mean, a payout percentage is not converted into a TSR hurdle without formula evidence, and a named index is not replaced by an equal-weight peer portfolio unless the disclosure establishes that construction; signatures whose evidence identifies the comparison class but not the required arithmetic are routed to adjudication or left unresolved.

Safeguards against outcome-driven interpretation. The method specification is fixed before realized focal and comparator returns are used to calculate the cutoff. The reviewing agent does not observe daily returns, does not alter the disclosed peer set, and cannot choose among plausible interpretations according to which one places the focal firm closest to a cutoff, because no return information exists in its input. When the text supports more than one economically plausible reading, we retain each supported alternative. A material change in pass/fail classification across supported readings produces an assumption-sensitive observation rather than an unqualified exact cutoff.

IA.1.2 Deterministic Calculation and Standardization in the Agentic Workflow

Canonical unit and date-valid security mapping. The data are first built at the most granular economic relation: grant $\times$ relative objective $\times$ exact performance period $\times$ disclosed peer. A grant can contain more than one objective or performance period, so returns are measured before any one-row-per-grant selection. The focal firm and every disclosed peer are linked to a security that is valid during the award's performance period. The preferred mapping sequence uses direct PERMNO, CIK, GVKEY through the CRSP/Compustat link, historical CUSIP, historical ticker, CRSP name history, exact normalized name, and corroborated high-confidence name recovery. Each mapping retains the source identifier, matching method, confidence, validity interval, and ambiguity flag.

Duplicate aliases that map to the same PERMNO are collapsed before peer counts and ranks are formed. The focal security is excluded from the primary disclosed-peer population. When the contract does not state whether the focal firm participates in the rank population, we also calculate a focal-included sensitivity result and flag material classification changes.

Security-level total shareholder return. For a simple buy-and-hold award, security j's cumulative total shareholder return over the award-specific trading dates is

$$\mathrm{TSR}_j = \prod_{d \in \mathscr{D}_j} (1 + RET_{jd}) - 1, \tag{IA.1}$$

where RET_{jd} is the CRSP daily total return and $\mathscr{D}_j$ is the set of trading dates in the award's performance period. Cash distributions are therefore included, and the CRSP return path incorporates the return consequence of delisting. Calendar endpoints are moved only to the first and last available trading observations consistent with the contractual period, and the resulting date gaps are retained.

Two variations cover the remaining contracts, and both are simple transformations of this buy-and-hold return rather than new formulas. When the contract expresses performance as an annualized return over D calendar days, the cumulative return is compounded to an annual rate as $(1 + \mathrm{TSR}_j)^{365.25/D} - 1$. When the contract compares average values around the endpoints rather than the endpoints themselves, the calculation builds a cumulative total-return wealth index for every security, averages that index over the disclosed beginning and ending windows, and takes the ratio of ending to beginning average wealth, minus one, as the return; trading-day windows use the stated number of trading observations, and calendar-day windows span the stated calendar interval. Whichever variation applies, the same transformation is applied to the focal firm and every peer before any ranking or aggregation, so that all securities are compared under the contract's own return convention.

Coverage is method-specific. A delisting can complete a buy-and-hold return because the terminal economic loss appears in the CRSP return path; it cannot, by itself, supply an ending averaging window years after the delisting. A security without enough observations for the required window is therefore incomplete unless the contract provides a survivor or removal rule.

Effective return-measurement constructions. A return-measurement construction records whether the contract uses single dates or average-wealth windows, the length and unit of those windows, their anchors relative to the contractual start and end dates, and whether the resulting return is annualized. This object is distinct from the benchmark rule; a full calculation design adds the comparison family, comparator aggregation, focal-rank convention, and ratio formula.

Table IA.3 reports the 21 return-measurement constructions represented in the 1,798 grant–objective–period observations underlying the cutoff panel. Single-date, nonannualized endpoints account for 1,675 observations, or 93.2 percent, and the remaining 123 observations use 20 alternative constructions. The most common alternatives are paired 20-trading-day wealth averages beginning on or after the contractual start (36 observations), annualized single-date returns (21),

paired 20-trading-day wealth averages anchored strictly before the start (17), and paired 30-trading-day wealth averages beginning on or after the start (10).

Table IA.3. Frequency of Effective Return-Measurement Constructions

Effective return measurement	Obj.-periods	Grants	Rule families	Full designs
Single-date endpoints; not annualized	1,675	1,627	6	11
20/20 trading-day wealth averages; on/after start to ending on end; not annualized	36	36	2	3
Single-date endpoints; annualized	21	21	2	3
20/20 trading-day wealth averages; strictly before start to ending on end; not annualized	17	17	2	3
30/30 trading-day wealth averages; on/after start to ending on end; not annualized	10	10	2	3
15/15 trading-day wealth averages; on/after start to ending on end; not annualized	6	6	1	2
20/20 calendar-day wealth averages; on/after start to ending on end; not annualized	6	6	2	2
60/60 calendar-day wealth averages; on/after start to ending on end; not annualized	6	6	2	2
30/30 calendar-day wealth averages; on/after start to ending on end; not annualized	5	5	1	2
30/30 trading-day wealth averages; strictly before start to ending on end; not annualized	5	5	1	1
10/10 calendar-day wealth averages; on/after start to ending on end; not annualized	1	1	1	1
10/10 trading-day wealth averages; on/after start to ending on end; annualized	1	1	1	1
15/15 calendar-day wealth averages; strictly before start to ending on end; not annualized	1	1	1	1
15/15 trading-day wealth averages; ending on start to ending on end; not annualized	1	1	1	1
20/20 calendar-day wealth averages; strictly before start to ending on end; annualized	1	1	1	1

Continued on next page

Table IA.3, continued

Effective return measurement	Obj.-periods	Grants	Rule families	Full designs
50/50 trading-day wealth averages; strictly before start to ending on end; not annualized	1	1	1	1
60/60 calendar-day wealth averages; ending on start to ending on end; not annualized	1	1	1	1
60/60 calendar-day wealth averages; strictly before start to ending on end; not annualized	1	1	1	1
60/60 trading-day wealth averages; on/after start to ending on end; not annualized	1	1	1	1
60/60 trading-day wealth averages; strictly before start to ending on end; not annualized	1	1	1	1
90/90 trading-day wealth averages; on/after start to ending on end; not annualized	1	1	1	1

Notes: The unit is a unique grant × relative objective × exact performance period represented in the cutoff panel; repeated threshold, target, and maximum rows are counted once. Paired lengths report the beginning and ending windows applied to both focal and comparator securities. Trading-day windows contain the stated number of CRSP observations; calendar-day windows span the stated calendar interval. "On/after start" begins with the first eligible trading observation at or after the contractual start, and "strictly before start" excludes the start date; ending windows use the last eligible observations at or before the contractual end. The 21 rows sum to 1,798 objective-periods and generate 43 full calculation designs after benchmark-family and comparator conventions are added.

The distribution has two implications. First, the endpoint calculation is the dominant construction but cannot be imposed universally: 6.8 percent of objective-periods require a different return transformation, and that transformation can change both the focal return and the ordering of peers. Second, one return construction can support several benchmark families, so the 21 return constructions expand to 43 full calculation designs once percentile, rank, mean, median, spread, ratio, named-index, aggregation, and focal-inclusion rules are distinguished.

Standardizing different contracts into common units. The contractual threshold, target, and maximum fields are goals, not necessarily benchmark returns. We standardize percentile goals to fractions, ordinal goals to descending ranks, and percentage-point and ratio goals to decimal relative-performance units, and we then invert each contractual rule into the focal-firm TSR required to attain that goal. This puts every retained running variable in the same economic unit: focal TSR minus required focal TSR. Table IA.4 summarizes the inversion applied to each method class.

Table IA.4. From Contract Language to a Focal-TSR Cutoff

Method class	Evidence needed	Cutoff in focal-TSR units
Percentile	Percentile, quartile, quintile, or tercile language and a peer population	Return of the observed peer at descending rank $N - \lceil qN \rceil + 1$.
Ordinal rank	Explicit rank or rank schedule	Return of the peer at the disclosed descending rank.
Peer mean	Explicit average or mean formula	Equal-weight mean of security-level peer TSRs, plus any disclosed spread.
Peer median	Explicit median comparator	Median peer TSR, plus any disclosed spread.
Percentage-point spread	Focal TSR minus comparator TSR equals hurdle h	Cutoff$(h) = C + h$, where C is the supported comparator.
Percentage ratio	Contract describes relative performance as a ratio	Primary wealth-ratio inversion Cutoff$(h) = (1 + h)(1 + C) - 1$; raw-return-ratio and spread alternatives are retained.
Named index	A specific index and any required combination rule	The supported index series under the disclosed measurement convention; otherwise no cutoff is manufactured.
Generic relative TSR	Relative-to-peer language without arithmetic	Unresolved unless another evidence field identifies rank, aggregation, or index arithmetic.

Notes: N is the number of complete distinct peers, q is the normalized percentile, C is the evidence-supported comparator, and h is the normalized hurdle. An observed marginal peer is used for rank and percentile awards; the calculation does not create an interpolated synthetic quantile. These rules are intentionally conservative: they reduce the number of exact cutoffs but prevent a contractually unsupported comparator from entering the running variable.

For a percentile award, complete peers are sorted from highest to lowest TSR. If the normalized percentile is q and there are N complete distinct peers, the primary marginal rank is

$$r(q,N) = N - \lceil qN \rceil + 1, \tag{IA.2}$$

and the benchmark is the return of the observed peer at rank $r(q,N)$. For an ordinal award, the benchmark is the observed peer at the disclosed descending rank.

The remaining method classes need no formulas beyond those shown in Table IA.4. For mean, median, and percentage-point awards, the required focal return is simply the evidence-supported comparator plus the disclosed spread, $C + h$. For a percentage-ratio award, the primary convention reads the hurdle as relative wealth performance, so the required focal return is $(1+h)(1+C)-1$; because a label such as "Percentage Ratio" may not distinguish a wealth ratio from a raw-return ratio or a percentage-point spread, the calculation also stores the raw-return-ratio and spread readings as supported alternatives, and a result is classified as assumption-sensitive when these readings change pass/fail status.

Running variable and treatment side. For grant i, objective-period combination g, and contractual cutoff k (threshold, target, or maximum), the running variable is

$$X_{igk} = \text{TSR}_{ig} - \text{CutoffTSR}_{igk}. \tag{IA.3}$$

A positive value is on the pass side and a negative value is on the fail side. TSR is stored in decimal units, so 0.01 equals one percentage point. Threshold, target, and maximum are retained as separate cutoff rows, which preserves the contractual distinction among the three cutoffs and permits any subsequent sample definition to be imposed without recalculating returns.

Worked contract examples. The following examples illustrate how disclosure language, return measurement, and benchmark inversion jointly determine the running variable.

Example 1: Percentile cutoff defined by the marginal peer. Grant 2,456,613 uses a percentile comparison against the EEI Stock Index peer group over 2021–2023. Incentive Lab reports a target of $q = 0.50$ and 38 complete peers, so Equation (IA.2) gives a marginal rank of $38 - \lceil 0.50 \times 38 \rceil + 1 = 20$. The twentieth-ranked peer is DTE Energy Company with realized TSR of 0.116745, and the focal firm's TSR is 0.116181, so the running variable is $0.116181 - 0.116745 = -0.000564$. The award is 0.0564 percentage points below target and is classified as a fail. The same rule is applied to threshold and maximum, but the marginal rank changes with the contractual goal.

Example 2: Ordinal-rank cutoff. Grant 1,972,385 states an ordinal rank and reports a target rank of six among ten complete peers. The sixth-ranked peer, Conagra Brands, has TSR of 0.314243, and the focal firm's TSR is 0.315878, giving a running variable of $0.315878 - 0.314243 = 0.001635$. The target is passed by 0.1635 percentage points. We do not interpolate between the fifth and sixth peers or substitute the peer-group mean. Because the disclosure does not expressly state whether the focal firm is in the ranked population, a focal-included calculation is retained as a sensitivity result.

Example 3: Explicit peer-mean formula. Grant 100,865 states that payout is calculated from "Average Qwest TSR Percentage minus Average Telecom Peer TSR Percentage." The formula identifies an equal-weight mean of the seven complete telecom-peer returns, and the 100 percent payout target corresponds to a zero relative spread. The peer mean is 0.127616, which is therefore the focal-TSR target cutoff. With focal TSR of 0.582319, the running variable is $0.582319 - 0.127616 = 0.454703$, and the target is passed. The zero-spread target is labeled a strong inference because the payout formula is explicit even though the normalized objective table does not separately report a zero TSR spread.

Example 4: Explicit named-index percentage-point rule. Grant 370,026 illustrates how separate evidence fields establish different parts of one exact rule. The comparison-method field states "Percentage points," which establishes additive spread arithmetic, while the named-benchmark field identifies the Semiconductor Sector Index (SOX). The authorized index dictionary maps that name to the PHLX Semiconductor Sector Total Return series and verifies coverage from February 18, 2014 through February 18, 2016. The realized index TSR is 0.116734.

At target, the disclosed spread is zero, so under the additive spread convention the required focal TSR equals the index return of 0.116734. The focal TSR is 0.319351, producing a running variable of 0.202617 and a pass. At maximum, the disclosed spread is 25 percentage points, so the required focal TSR is 0.366734 and the running variable is -0.047383. The index identity, total-return basis, contractual dates, and additive inversion are explicit and independently validated, so the resulting cutoff is classified as exact and carries no method-assumption flag.

Example 5: Explicitly annualized percentile calculation. Grant 1,889,628 shows that annualization is a return-measurement attribute separate from the percentile rule. The SEC method field states "Three-year annualized TSR," and the objective-summary field states "Percentile." Incentive Lab supplies a performance interval from December 26, 2019 through December 26, 2022, a 50th-percentile target, and 20 disclosed peers, all with complete returns.

Over the 1,096-day interval, the focal firm's compounded total return is 0.454084, and applying the annualization convention gives annualized focal TSR of $(1 + 0.454084)^{365.25/1096} - 1 =$

0.132881. The contractual target is descending rank $20 - \lceil 0.50 \times 20 \rceil + 1 = 11$. The eleventh-ranked peer, The Greenbrier Companies, has unannualized TSR of 0.203982 and annualized TSR of 0.063818. The benchmark is therefore 0.063818, the running variable is 0.069063, and the target is passed. The classification is unchanged by the focal-inclusion sensitivity and by missing-peer perturbations, so the cutoff is retained as a high-confidence exact observation.

Example 6: 60-trading-day averaging windows. Grant 273,699 states that TSR uses average closing-price and reinvested-dividend wealth over the 60 trading days ending at the performance-period end relative to the corresponding 60 trading days preceding the beginning. The performance period is January 7, 2011 through January 7, 2014. The beginning window contains October 13, 2010 through January 6, 2011, and the ending window contains October 11, 2013 through January 7, 2014.

For the focal firm, beginning and ending average wealth are 2.539016 and 2.091153, so the averaging-window convention gives focal TSR of $2.091153/2.539016 - 1 = -0.176393$. For Murphy Oil, the peer defining the threshold rank, the corresponding averages are 2.965984 and 3.472490, giving TSR of 0.170772. The threshold is descending rank 13 among 18 complete peers, so Murphy Oil's transformed TSR is the benchmark. Subtracting the benchmark from the focal firm's transformed TSR gives a running variable of $X = -0.347164$, and the focal firm fails. The important point is that the same 60-day transformation is applied to every peer before sorting; a simple endpoint return would implement a different contract.

All six examples come from the exact component. Each has complete comparator coverage or an independently validated exact index series, and none has a comparator-sensitive calculation status. Together they span the dominant single-date construction, an annualized endpoint construction, a named-index spread rule, a peer mean, ordinal and percentile order statistics, and a less common averaging-window design. They also show what the procedure does not do: it does not infer a peer mean merely because peers are listed, convert a payout percentage into a TSR hurdle without formula evidence, substitute a generic peer portfolio for a named index, or select the interpretation that minimizes the focal firm's distance from a cutoff.

Method composition and identification status. Percentile awards account for 3,688 of the 4,328 retained cutoff rows, ordinal ranks for 532, percentage-point spreads for 70, percentage ratios for 28, peer medians for nine, and peer means for one. The exact and estimated components are mutually exclusive at the grant–objective–period–cutoff key. Assumption-sensitive status does not mean the arithmetic failed; it means that an otherwise reproducible result depends on a disclosed convention, recovered identifier, focal-inclusion rule, ratio interpretation, or comparator assumption.

IA.1.3 Handling Missing Peers

Peer coverage and the distinction between classification and a continuous cutoff. The canonical peer panel contains 47,735 disclosed peer rows. After historical identifier recovery, 44,411 rows map to PERMNO, or 93.0 percent, and 41,247 have a complete CRSP return window under the award-specific method, or 86.4 percent. The exact component has mean complete-peer coverage of 99.79 percent, the interval-identified component has mean observed-complete coverage of 89.82 percent, and the median contractual peer count in the retained sample is 16.

Missing peers do not always make pass/fail status unknown. Let m be the number of unresolved peers and let the focal firm's rank among observed peers be r. Its best possible full-set rank is r, and its worst possible rank is $r+m$. If both ranks are on the same side of the contractual hurdle, classification is robust to every placement of the missing peers, and the observation is flagged accordingly. Because the marginal peer's return is not uniquely known, however, this classification alone does not provide a continuous running-variable value. If the rank bounds cross the hurdle, the pass/fail classification is treated as sensitive to the unresolved peers.

Finite cutoff intervals for rank and percentile awards. For a rank or percentile contract, let N denote the expected number of distinct contractual peers, n the number of peers with complete returns, $m = N - n$ the unresolved peers, and k the contractual descending marginal rank. Sort observed peer returns as

$$R_{(1)} \geq R_{(2)} \geq \cdots \geq R_{(n)}.$$

When both needed order statistics exist, the unknown marginal-peer cutoff C is sharply bounded by

$$C \in \left[R_{(k)}, R_{(k-m)}\right]. \tag{IA.4}$$

No missing peer return is set to zero, a peer-group average, an industry return, or an interval midpoint. If $k > n$ or $k - m \leq 0$, one endpoint depends entirely on unrestricted missing returns and the row is not admitted to the finite two-sided interval sample.

If the focal return is F, the running variable is then bounded between $F - C_U$ and $F - C_L$, where C_L and C_U are the endpoints of the cutoff interval. We retain both endpoints, the interval width, and an indicator that the running variable is interval identified, and the prespecified point convention records each interval at its numerical upper endpoint. This is a reporting convention fixed before any bandwidth counts were examined, not an estimate that an unidentified peer return equals a boundary value.

A simple illustration conveys the logic of Equation (IA.4). Suppose a contract expects ten peers, eight have complete returns, two are unresolved, and the target is descending rank six. The relevant observed returns imply $C \in [R_{(6)}, R_{(4)}]$. If those order statistics are -0.02 and 0.12 and the focal return is 0.04, then $X \in [-0.08, 0.06]$: the interval crosses zero, so pass/fail is assumption-sensitive. If the focal return instead equals 0.20, then $X \in [0.08, 0.22]$ and the focal firm passes under every placement of the missing peers.

Main, exact-only, and certified samples. The main point sample includes exact rows and estimated rows evaluated at the upper-endpoint convention. Two robustness samples remove or limit dependence on that convention. The exact-only sample drops every nondegenerate interval. The certified sample retains an estimated observation inside a symmetric bandwidth only when the entire feasible interval, both endpoints included, lies inside the window. At ± 3 percentage points, the main point rule yields 235 cutoff rows across 222 grants, while full-interval certification yields 185 rows across 174 grants. The point, exact-only, and certified definitions provide complementary ways to assess sensitivity to interval identification without treating an interval endpoint as an exact realization.

Validation. We validate the Relative-TSR construction at the security-return, exact-cutoff, interval-cutoff, and consolidated-panel levels. The tests cover key uniqueness, date ordering, total-return arithmetic, annualization, averaging-window anchors, rank direction, hurdle normalization, benchmark inversion, running-variable signs, and the identity of interval endpoints. Independent validators reconstruct selected observations directly from the underlying security-level returns, with deliberate oversampling of observations near zero, uncommon calculation designs, recovered identifiers, low peer coverage, and extreme returns.

All 1,893 finite nondegenerate intervals and 13 degenerate intervals created by later peer recovery are reconciled to the underlying ordered peer returns. The consolidated check verifies the 4,328-row union, the absence of duplicate grant–objective–period–cutoff keys, the inclusion of threshold, target, and maximum observations, and the requirement that every retained point value lie inside its stored bounds. Separate reviews cover 100 observations nearest the cutoff, 258 observations selected through a broader risk-based validation sample, and all 112 recovered or corrected exact rows. No reviewed observation fails the numerical reconstruction. Observations that depend on a disclosed mapping, focal-inclusion convention, ratio interpretation, or comparator assumption remain identified as assumption-sensitive even when the arithmetic is correct.

IA.2 Patent Name Matching with an Agentic AI Method

The patent linkage is constructed at the application stage rather than only after a patent is granted. Filing-year assignment aligns the innovation measure with the firm's decision to pursue an invention, but application records often identify a subsidiary, operating company, acquired entity, or legal assignee whose name differs from that of the publicly traded parent. The objective of the matching procedure is therefore to recover economic identity rather than to maximize textual similarity.

The resulting crosswalk links each raw USPTO application assignee to the SEC Central Index Key (CIK) of the relevant public company. Each accepted link retains the route through which it was obtained, the corporate relationship supporting a non-literal match, and sufficient candidate information to review difficult cases. Unsupported records remain unmatched.

IA.2.1 Why a Conventional One-Pass Name Match Is Insufficient

A one-pass fuzzy match ranks names by string similarity and accepts or rejects the strongest candidate at a fixed score. That approach is efficient for candidate generation, but it has two structural limitations in this setting. First, a high similarity score is not proof that two entities are economically identical. Second, a low score does not imply that a match is false when the application is filed by a subsidiary or an acquired company whose legal name differs sharply from the current listed parent.

The reference universe therefore combines public-company information from CRSP and Compustat with the broader company and corporate-hierarchy information available in Revelio Labs. Revelio's parent–child relationships are particularly valuable because patent applications are frequently filed under a subsidiary, operating unit, or acquired entity rather than under the name of the publicly traded parent, and a match limited to the parent names in standard public-firm files can miss an economically valid link even when the assignee name is accurate. The corporate-family structure creates a documented path from a child-entity assignee to its public parent, which improves recall without asking the similarity algorithm to treat two dissimilar strings as though they were the same legal name. Assignee names are also standardized before candidate generation for routine formatting differences such as capitalization, punctuation, spacing, common legal suffixes, and token order; standardization makes the candidate search more reliable but is never itself treated as a decision that two entities are the same.

IA.2.2 Hybrid Fuzzy and Agentic Matching Workflow

The procedure separates high-throughput candidate generation from agentic identity adjudication, and it proceeds in three stages summarized in Figure IA.2.

In the first stage, deterministic code defines the matching universe and generates candidates. The procedure begins with the unique assignee entities appearing in USPTO pre-grant application records and identifies the U.S. subset used for public-company matching. Each U.S. assignee is compared with names in the Revelio and CRSP/Compustat reference universes, and the three strongest candidates are retained together with their similarity information. A literal 100-percent name match is accepted directly as an automated match, subject to identifier and duplicate checks.

In the second stage, every non-perfect candidate set is passed to language-model agents for contextual identity review; a candidate is never accepted merely because its similarity score exceeds a numerical threshold. Each reviewing agent examines the highest-ranked candidates and determines whether one represents the same economic entity as the USPTO assignee, drawing on company context and corporate-family information rather than the similarity score alone. When none of the proposed candidates is valid, the agent searches the internet for a better public-company match, and a match outside the initial candidate set is accepted only when the available evidence supports the economic relationship. Difficult or ambiguous records are judged independently by multiple agents; their conclusions are compared, and disagreements are refined or adjudicated rather than silently resolved by the highest fuzzy score.

In the third stage, the accepted public parent receives its CIK, and the crosswalk records whether the link arose from the direct fuzzy stage or the agentic-review stage. Records with no defensible public-company link remain unresolved.

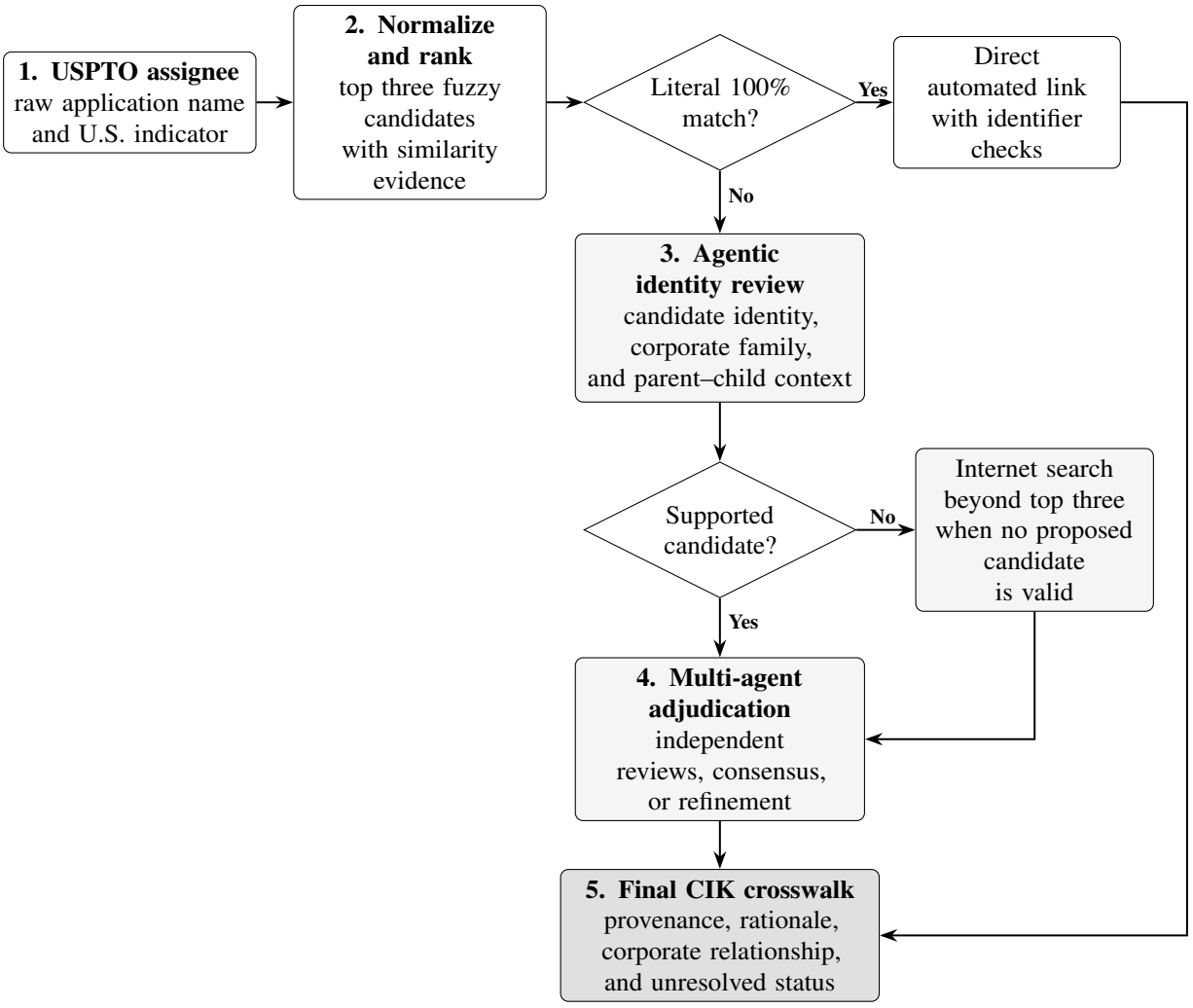


Figure IA.2. Hybrid Fuzzy and Agentic Patent-Assignee Matching Procedure. This figure summarizes the entity-resolution workflow that links USPTO application assignees to public-firm CIKs. Fuzzy similarity generates candidates but does not settle non-perfect identity decisions. A literal perfect match can proceed directly; all other cases enter agentic review, in which language-model agents assess candidate identity using corporate-family and public-source evidence, may search beyond the initial candidate set, and submit difficult records to independent multi-agent adjudication. The final crosswalk records the resolution route and leaves unsupported records unmatched.

Several rules prevent the agentic-review stage from becoming an unconstrained matching exercise. A candidate is accepted only when it represents the same economic entity or a supported parent–child relationship; the existence of a plausible public company in the same industry is not sufficient. An agent can reject all proposed candidates. When a subsidiary is mapped to a parent, the relationship is recorded explicitly rather than hidden behind a low-similarity name match. The final CIK is therefore accompanied by match provenance, candidate information, the agentic decision, and a short rationale where available.

IA.2.3 Matching Funnel and the Incremental Contribution of Agentic Review

The starting application-assignee universe contains 292,435 unique entities, of which 115,892 are identified as U.S. assignees. The fuzzy stage links 29,043 U.S. assignees to public firms. Agentic review adds 6,962 links that the direct fuzzy stage did not resolve, producing 36,005 U.S. assignee entities with a public-company CIK; the remaining 79,887 U.S. assignee entities do not receive a reported public-company link. Table IA.5 reports the funnel.

Table IA.5. Patent-Assignee Matching Funnel

Stage	Assignee entities	Share	Interpretation
Unique assignees filing applications	292,435	100.00%	Starting USPTO application-assignee universe.
Unique U.S. assignees	115,892	39.63% of total	Geographic subset used for public-company linkage.
Matched through the fuzzy procedure	29,043	25.06% of U.S.	Direct recovery from the automated candidate-ranking stage.
Additional matches from agentic review	6,962	6.01% of U.S.	Links recovered through agent candidate review, corporate-family evidence, or supported search beyond the top three candidates.
Final U.S. assignees linked to a CIK	36,005	31.07% of U.S.	Combined direct and agent-reviewed match set.
U.S. assignees without a reported CIK link	79,887	68.93% of U.S.	Residual assignee entities left unmatched.

Notes: Counts refer to assignee entities, not necessarily distinct public issuers, because multiple legal names can map to the same CIK. The agentic-review additions equal 23.97 percent of the fuzzy-only count and 19.34 percent of the final matched set, so approximately one in five final links is attributable to agentic review rather than to direct string similarity.

The funnel should not be interpreted as a target to match every assignee. Many applications are filed by private firms, individuals, universities, foreign entities, or organizations outside the target public-company universe, and leaving such a record unmatched is preferable to forcing a weak public-company attribution.

IA.2.4 Validation Against an Independent Patent–Firm Linkage

The application-assignee links are compared with known public-firm links for applications that eventually become granted patents and appear in the linkage of Kogan et al. (2017), hereafter KPSS. The reported agreement is approximately 98.6 percent. This external comparison is informative because the application-stage match is evaluated against a separately constructed grant-stage firm identity rather than against the same matching rules used to create it.

The 98.6-percent agreement applies only to observations that eventually grant and have a KPSS benchmark link. It is therefore evidence about the validation subsample, not a universal error rate for the full application-assignee universe. The validation summary does not provide the comparison-sample size, a false-positive/false-negative matrix, or separate disagreement rates for direct and agent-reviewed links, and it cannot evaluate application-only observations that never grant, because those records have no KPSS benchmark. We accordingly interpret the exercise as strong external agreement within its support rather than as complete validation of every matched assignee.

IA.2.5 Crosswalk Structure and Analytical Use

The crosswalk retains more than the final CIK. Table IA.6 summarizes the information used to review difficult links and to conduct stage-specific sensitivity tests.

Table IA.6. Information Retained for Review of Application-Assignee Links

Field block	Example content	Purpose
Source identity	USPTO assignee identifier and raw assignee name	Preserves the entity exactly as observed in the application record.
Scope indicator	U.S.-assignee flag	Documents entry into the target matching universe.
Candidate set	Top three names, identifiers, source universe, similarity scores, and ranks	Makes the automated stage reviewable and permits alternative thresholds.
Perfect-match flag	Indicator for a literal 100-percent match	Separates direct automated links from agent-reviewed cases.
Agentic decision	Accepted candidate, no valid candidate, or supported external-search result	Records the contextual decision made after candidate generation.
Corporate relationship	Direct identity, subsidiary-to-parent, acquired entity, or other supported relationship	Documents why a non-literal name can map to the public parent.
Consensus information	Independent agent decisions and adjudicated result	Preserves evidence of multi-agent scrutiny for difficult cases.
Final public-firm key	CIK and, where available, GVKEY or ticker	Supports deterministic merging to SEC and accounting data.
Match provenance	Direct fuzzy or agentic-review stage	Enables stage-specific validation and robustness tests.
Evidence or rationale	Short source-supported explanation	Facilitates manual review of difficult or non-literal matches.

Notes: This table describes the field blocks retained in the assignee-to-CIK crosswalk. The retained candidate rankings, decisions, and provenance make it possible to re-examine any link, to compare direct and agent-reviewed matches separately, and to impose stricter acceptance rules in robustness tests without repeating the matching procedure.

The matching crosswalk and the firm-year patent panel are distinct data objects. The crosswalk contains 36,005 matched U.S. assignee entities. The analytical patent panel contains 397 firms and 10,322 firm-years over 2000–2025 after downstream sample restrictions and aggregation. Because multiple assignee names can map to one public parent and because the analytical panel imposes restrictions beyond entity resolution, the two counts do not form a one-to-one funnel. Patent outcomes are assigned to the application filing year, non-filing firm-years retain structural-zero counts, and unresolved firm identities are never forced into the panel. Every downstream patent measure inherits the quality of the upstream assignee-to-CIK match, so the crosswalk's provenance fields also support robustness tests that restrict attention to direct matches or to agent-reviewed matches.

IA.3 Measures Constructed with an LLM Batch Process

The measures in this section convert DEF 14A and Form 10-K language into structured research variables through a batch-classification process. For the pledging, governance-policy, and managerial-myopia measures, each filing is submitted to the language model as an independent request under a purpose-built prompt that ties the classification to a narrowly defined economic question and a restricted set of admissible values, and the model returns structured output together with the filing evidence supporting its decision; the committee-discretion measure applies the same restricted-rubric design to proxy-statement evidence records at the issuer-period level.[11] Because every observation is processed independently, information from one observation cannot influence the classification of another, and the procedure scales uniformly across tens of thousands of filing observations.

Each classification retains the issuer and period identifiers, the classification itself, and supporting evidence or rationale where applicable. The variables measure what is disclosed in the filing; they do not establish the existence of an undisclosed policy, the execution of a permitted action, or the manager's unobserved private preferences. Table IA.7 summarizes the measures, their source documents, and their native units.

IA.3.1 Board and Compensation-Committee Discretion

Economic construct. The measure captures discretion retained by the compensation committee over compensation instruments and payout adjustments. We refer to this as board discretion because the compensation committee operates under board authority, but the coded object is more specific: it is the committee's disclosed authority to increase, reduce, cancel, modify, or otherwise determine compensation outcomes.

This distinction matters for the Relative-TSR design. When a contract maps realized performance mechanically into shares, crossing a contractual cutoff creates a sharper change in the CEO's equity. When the committee can alter the formulaic result after performance is observed, the running variable maps less tightly into the treatment actually received. The measure records the existence and stated direction of that authority; it does not establish that the committee exercised discretion in a particular case.

[11] The CEO share-pledging and pledging-restriction/anti-hedging classifications use GPT-5 nano accessed through the OpenAI Batch API, with one request per filing, and the managerial-myopia classification also uses GPT-5 nano; low reasoning effort is documented for the pledging and myopia runs. For the pledging extraction, the filing text is normalized for encoding and whitespace, the full filing is retained whenever it fits within the model's context budget, unusually long filings are truncated only at the end of the document, each request carries a unique identifier constructed from the source filename and collection folder, and request- and batch-level manifests are retained for reproducibility; the policy classification uses the same source files and context handling.

Table IA.7. Text-Based Measures, Source Documents, and Native Units

Measure	Source text	Native unit	Main interpretation
Compensation-committee discretion	DEF 14A evidence describing committee authority	CIK–year-month classification record	Compensation instrument covered by discretion and the stated direction of that authority.
CEO share pledging	DEF 14A beneficial-ownership tables and footnotes	Filing observation	Whether the current CEO is identified as actually pledging shares and, where disclosed, the amount and context.
Pledging restriction	DEF 14A policy language	Filing observation	Whether the filing discloses a meaningful ban, limit, approval requirement, or other restriction on pledging company securities.
Anti-hedging policy	DEF 14A policy language	Filing observation	Whether the filing discloses a meaningful restriction on hedging economic exposure to company securities.
Managerial myopia	Form 10-K MD&A	CIK–fiscal/report year	Ordered 1–5 assessment of whether management emphasizes near-term performance at the expense of long-run investment or capabilities.

Notes: This table summarizes the five text-based measures constructed through the LLM classification process. The pledging, governance-policy, and myopia measures are produced by an independent batch request per filing under a purpose-built prompt with a restricted set of admissible values; the committee-discretion measure applies the same restricted-rubric design to issuer-period evidence records. The native unit is the level at which the classification is produced; annual or event-time variables should be formed only afterward under prespecified aggregation rules.

Classification scheme and retained fields. The classification stage applies a restricted rubric to DEF 14A passages previously identified as describing compensation-committee discretion. Each observation receives one mutually exclusive compensation type and, separately, one direction category. Table IA.8 reports the two taxonomies.

Table IA.8. Classification Taxonomy for Compensation-Committee Discretion

Category	Definition
Panel A. Compensation type (mutually exclusive)	
`stock_awards`	Discretion over restricted stock, restricted shares, RSUs, PSUs, performance shares, stock units, grants, and comparable equity awards.
`stock_options`	Explicitly option-only authority, such as option exercisability, strike price, or option-specific vesting.
`bonus`	Cash bonus, annual incentive, cash incentive, incentive pool, or a comparable cash arrangement.
`stock_awards_and_bonus`	One authority that clearly covers both equity and cash-incentive compensation.
`others`	Salary, benefits, severance, general compensation decisions, plan interpretation or administration, or another authority outside the preceding categories.
Panel B. Direction of authority (classified separately from type)	
`increase`	The evidence expressly supports one-sided upward or participant-favorable authority and does not state downward authority.
`reduce`	The evidence expressly supports one-sided downward authority, including reduction, elimination, cancellation, forfeiture, or a lower payout, and does not state upward authority.
`two_sided`	The evidence explicitly supports both upward and downward discretion.
`not_mentioned`	Discretionary authority is present, but its direction is not stated or clearly implied.

Notes: Generic references to "Award" or "Awards" are coded as stock awards unless the surrounding language clearly identifies a cash incentive, an option-only provision, or another category, and an "award" described as a computed or cash-incentive payout is coded as bonus. Generic verbs such as "determine," "adjust," "modify," or "change" are not, by themselves, treated as evidence of two-sided authority, so `not_mentioned` means that direction is not established, not that discretion is absent. When the type is `others`, an open-text field retains a short non-harmonized description. The literal variable names retain the spelling used in the underlying dataset.

Each record retains a unique issuer-period identifier formed from CIK and a YYYY-MM year-month label, the CIK, the period label, the compensation-type classification, the open-text descriptor for the residual category, and the direction classification. The period label is a calendar filing-period label and is not assumed to be a fiscal-year end.

Classification sample. The resulting classification sample contains 24,225 unique CIK–year-month observations for 3,704 issuers from January 2002 through December 2024. All identifier values are unique, and the two primary categorical fields are populated for every row. The file is

conditional on observations that entered the committee-discretion classification stage, so its category shares describe the composition of classified discretion records, not the unconditional prevalence of discretion across all proxy statements or all compensation contracts.

Within that conditional file, bonus or cash-incentive discretion accounts for 9,758 observations, stock-award discretion for 9,269, the residual `others` category for 3,853, combined stock-award and bonus discretion for 1,155, and option-only discretion for 190. Direction is not stated in 15,533 observations; 5,441 identify reduction-only authority, 2,467 two-sided authority, and 784 increase-only authority. Among the 8,692 observations with an explicit direction, reduction-only authority is the most common at 62.6 percent, consistent with committees more often disclosing negative discretion over formulaic payouts than upward adjustments. These distributions are descriptive of the coded records and are not used as evidence that one type of authority is more common in the population of firms.

Stock-award-discretion indicator. For Relative-TSR awards, the principal indicator equals one when the classified compensation type is `stock_awards` or `stock_awards_and_bonus`, and direction remains a separate categorical variable. The indicator captures whether the disclosed contract permits committee authority over stock-award outcomes; it is not evidence that the committee exercised that authority in a particular award.

A filing or period without identified stock-award discretion is coded as zero in the binary indicator. This convention is deliberately conservative for disclosure-based analysis, but the interpretation is narrow: zero means that the classified evidence does not identify qualifying stock-award discretion, not that no authority existed outside the observed disclosure. The native type and direction fields are retained so that analyses can distinguish explicit reductions, explicit increases, two-sided discretion, and direction not established.

Interpretation and validation limits. The measure is a text classification of disclosed authority, not a measure of ex post intervention. Category shares are conditional on observations that entered the discretion-classification sample and should not be interpreted as unconditional prevalence among all proxy statements. The dataset does not contain an independently re-coded filing sample from which to estimate classification error. Empirical applications should therefore retain the underlying evidence where available and report sensitivity to alternative treatments of unmentioned direction and disclosure-based zeros.

IA.3.2 CEO Share Pledging, Pledging Restrictions, and Anti-Hedging Policies

This subsection separates actual CEO behavior from firm policy. A filing can report that the CEO has shares pledged even when the firm restricts future pledging, and a firm can disclose a pledging or anti-hedging policy even when the CEO has no identified pledged shares. The variables are therefore constructed and interpreted separately.

Actual CEO share pledging. Each DEF 14A filing is processed independently by GPT-5 nano through the OpenAI Batch API, with one request per filing and low reasoning effort. The filing text is normalized for encoding and whitespace, and the full filing is retained whenever it fits within the model's context budget. The classification prompt is restricted to the current CEO and directs the model to examine beneficial-ownership tables, the CEO's ownership row, and related footnotes.

A positive classification requires filing evidence that the current CEO actually pledged company shares. Information about another director, officer, beneficial owner, founder, former CEO, or an undifferentiated executive group is not transferred to the CEO, and the prompt distinguishes actual pledging from a policy restriction, permission to pledge, hypothetical risk language, and pledging by another person. Table IA.9 reports the retained variables.

Table IA.9. Filing-Level CEO Share-Pledging Variables

Field	Interpretation
`ceo_pledging_status`	CEO-specific classification: `yes`, `no`, or `Info_not_Found`. A positive classification requires evidence of actual current-CEO pledging.
`shares_pledged`	Total number of CEO shares explicitly identified as pledged; unresolved amounts retain the unresolved state rather than a fabricated number.
`total_beneficial_shares`	CEO beneficial ownership as of the relevant ownership measurement date.
`options_or_shares_acquirable_within_60_days`	Shares included in beneficial ownership through options, warrants, rights, or similar securities exercisable or acquirable within 60 days.
`record_date`	Beneficial-ownership measurement date, preferably standardized to YYYY-MM-DD.
`pledged_share_class`	Stock or ownership class associated with the pledge when disclosed.
`pledge_context`	`margin_loan`, `margin_account`, `collateral_loan`, `other_pledge`, `NONE`, or `Info_not_Found`.

Notes: This table reports the substantive variables retained from the CEO share-pledging extraction. Model evidence, section labels, confidence scores, and API metadata are produced during processing but excluded from the analytical file.

The filing-level sample contains 5,506 observations for 379 firms over 2002–2024. It contains 5,424 unique CIK-year combinations; 80 firm-years appear in more than one source month and

produce 82 additional observations relative to a strict firm-year panel, and those repeated firm-years should be resolved under a prespecified filing-selection or aggregation rule rather than dropped mechanically. The classification distribution contains 230 positive pledging observations, 4,000 negative observations, and 1,276 `Info_not_Found` observations. Among the 230 positive observations, collateralized borrowing dominates the disclosed context: 76.5 percent identify a collateral loan and 10.0 percent a margin account.

Numeric pledge amounts should be interpreted together with the classification and the ownership denominator. Among positive observations with numeric pledged shares and positive beneficial ownership, the median pledged-to-beneficial-ownership ratio is 21.95 percent, but two raw ratios exceed 100 percent, indicating that extreme ratios require source-level review before empirical use. Share-class labels are retained close to the filing language and are not fully harmonized; labels such as "Common Stock," "Common stock," and "Common Shares" should therefore be normalized in a separate, documented cleaning step rather than overwritten in the raw extraction.

Annual and event-time aggregation. The native unit of the pledging measure is the filing observation. Annual or event-time variables should be formed only after repeated CIK-year observations have been resolved under a fixed filing-selection or aggregation rule. An indicator for any identified CEO pledge can then be constructed from positive annual classifications, and pledged-share amounts can be averaged or summed over a stated event window before applying transformations such as $\ln(1+x)$.

The native dataset preserves the distinction between `no` and `Info_not_Found`. Collapsing the unresolved category into zero changes the economic interpretation from "the filing states that the CEO has no pledged shares" to "no pledge was identified in the available filing evidence." Any event-time aggregation that adopts the latter convention should state it explicitly and should leave years without source coverage missing.

Pledging restrictions and anti-hedging policies. The same DEF 14A filings are separately classified for pledging restrictions and anti-hedging policies. Each filing is assessed independently by GPT-5 nano through the OpenAI Batch API under a purpose-built prompt limited to these two governance provisions, and the model returns strict JSON output containing two binary indicators, the filing evidence for each policy, and a concise explanation of each decision. Because the prompt separates the two policies explicitly, a filing can receive any of the four combinations of the two indicators.

The pledging-policy indicator equals one when the filing discloses a meaningful ban or restriction on pledging company securities, including an outright prohibition, a limit, an approval

requirement, or another substantive constraint. The anti-hedging indicator equals one when the filing discloses a meaningful restriction on hedging the economic exposure associated with company securities; it does not capture the firm's corporate hedging of operating risks such as interest-rate, foreign-exchange, commodity, or fuel exposure. Table IA.10 summarizes the distinction between the behavior and policy measures.

Table IA.10. Distinguishing Pledging Behavior from Governance Policies

Variable	Positive classification means	A zero or unresolved value does not mean
Actual CEO pledging	The current CEO is explicitly identified as having company shares pledged.	That the firm permits pledging generally, or that another officer's pledge belongs to the CEO.
Pledging restriction	The filing discloses a meaningful ban, limit, approval requirement, or other restriction on pledging company securities.	That no actual pledge exists; policy and behavior are separate objects. A zero means no qualifying restriction was found in that filing, not that no policy existed elsewhere.
Anti-hedging policy	The filing discloses a meaningful restriction on hedging exposure to company securities.	That the firm does not hedge interest-rate, currency, commodity, or other operating risks.

Notes: This table contrasts the three pledging-related variables. Actual CEO pledging is a behavior measure extracted from beneficial-ownership disclosures; the pledging-restriction and anti-hedging indicators are policy-disclosure measures extracted from governance language in the same filings.

The policy dataset contains the same 5,506 filing observations and 379 firms over 2002–2024. Of the filing observations, 3,022, or 54.89 percent, are classified as containing a pledging ban or restriction, and 3,356, or 60.95 percent, as containing an anti-hedging policy. Both are present in 2,660 filings, or 48.31 percent, and neither is identified in 1,788 filings, or 32.47 percent. Policy disclosure rises strongly over time: pooled over 2002–2005, 18.39 percent of filings disclose a pledging restriction and 0.49 percent an anti-hedging policy, whereas pooled over 2021–2024 the corresponding rates are 90.43 and 98.23 percent. These are descriptive rates within the available filings, so empirical specifications using these variables should account for calendar time and should not interpret a filing-level zero as proof that a policy never existed.

The evidence fields make the binary indicators auditable, but the retention is asymmetric by construction. For positive classifications, the evidence field records the filing language that supports the indicator and the explanation summarizes why it qualifies; for zero classifications, the evidence field is often blank because no qualifying passage was identified, and the explanation records the basis for the zero. This asymmetry is expected and should not be interpreted as weaker documentation for positive cases.

IA.3.3 Managerial-Myopia Measure

Construct and source text. Managerial myopia is measured from the Management's Discussion and Analysis (MD&A) section of Form 10-K filings. The measure is intended to capture management's expressed time horizon, not the number of references to quarters, months, years, or other temporal words. A higher score indicates stronger emphasis on near-term earnings, margins, cash flow, or market objectives when the disclosure suggests that this emphasis constrains or sacrifices longer-horizon investment, capabilities, or value creation.

The classification centers on an intertemporal tradeoff. Long-horizon objects include R&D, technology, productive capacity, employees, brands, infrastructure, and market development. Routine discussion of current-year performance is not sufficient evidence of myopia, and cost reduction is treated as myopic only when the text indicates that valuable long-term activity is delayed, constrained, or sacrificed primarily to support near-term performance.

Scoring procedure. Each usable MD&A section is submitted to GPT-5 nano under a structured classification prompt with low reasoning effort. For each firm-year, the model returns an integer score from one to five, a concise rationale, and one to three pieces of supporting textual evidence, and the score with its supporting fields is retained at the CIK-year level. Table IA.11 reports the scoring rubric.

Table IA.11. Managerial-Myopia Scoring Rubric

Score	Classification	Interpretation
1	Very low myopia	Strong long-term orientation, including willingness to incur near-term costs or weaker current results to preserve valuable long-term investments.
2	Low myopia	Predominantly long-term oriented; ordinary attention to current financial performance without convincing evidence that long-term value is being sacrificed.
3	Moderate, mixed, or unclear	Meaningful short- and long-horizon considerations coexist, or the disclosure does not reveal a clear intertemporal tradeoff.
4	High myopia	Strong near-term orientation with credible indications that longer-horizon investments or capabilities are constrained, delayed, or subordinated to current performance.
5	Very high myopia	Explicit or exceptionally strong evidence that long-term corporate value or investment is sacrificed to achieve immediate financial or market objectives.

Notes: The endpoint categories are deliberately demanding. A score of one requires unusually clear willingness to accept near-term costs for long-term value, and a score of five requires explicit or exceptionally strong evidence of sacrifice, so most filings fall in the middle-low categories rather than at either endpoint.

Two source routes and duplicate resolution. The CIK-year panel combines two text-extraction routes that use the same rubric. In the first, the MD&A section is isolated from the 10-K before classification. In the second, Item 7 is already available as a structured filing section from an SEC crawler and is classified directly. The two sets of classifications are combined by CIK and fiscal or report year, and when the same CIK-year appears in both routes, the observation based on the structured Item 7 section is retained. This source-priority rule produces one observation per CIK-year and avoids averaging two classifications of the same filing-year. The retained fields are `cik`, `year`, `myopia_score`, `rationale`, and `key_evidence`.

Sample, distribution, and empirical alignment. The resulting panel contains 3,463 firm-year observations for 390 firms and spans fiscal or report years 2001–2026. The mean score is 2.455, the standard deviation is 0.628, and the median is two. Score two accounts for 61.57 percent of observations, score three for 31.24 percent, and scores four and five together for 7.13 percent. These proportions reflect the demanding endpoint definitions and should not be interpreted as externally validated prevalence estimates of managerial short-termism.

For event-time applications, the score can be aligned to the fiscal or report year preceding the relevant event so that the disclosure predates the outcome under study. The five categories are ordered rather than cardinal, so the score can enter a regression as an ordered index, a high-myopia indicator, or a moderator, but the chosen treatment should be stated explicitly, and robustness checks can use ordered categories or alternative thresholds rather than a linear one-unit interpretation.

The rationale and evidence fields form part of the measure and preserve the textual basis for each score, permitting qualitative review and manual reassessment of difficult cases. The panel does not contain an external human-coded benchmark or an inter-rater agreement statistic. The score should therefore be interpreted as an LLM-based measure of the horizon expressed in MD&A, which, as a disclosure, can differ from the manager's unobserved private horizon, rather than as a direct observation of the manager's preferences or actions.

IA.4 CEO Delta and Vega from ISS Incentive Lab

The CEO portfolio measures quantify two related but different exposures using standard Black–Scholes–Merton sensitivities (Black and Scholes, 1973; Merton, 1973). *Delta* is the dollar change in the CEO's firm-related wealth for a one-percent increase in the firm's stock price. *Vega* is the dollar change in option wealth for a 0.01 increase in annualized stock-return volatility, for example from 0.30 to 0.31. Stock has price sensitivity but no option-model vega; option tranches contribute to both measures.

In defining and calculating these sensitivities we imitate the methodology of Coles, Daniel, and Naveen (2013) and the related literature it codifies (Core and Guay, 2002; Coles, Daniel, and Naveen, 2006). Delta and vega carry their one-percent-price and 0.01-volatility conventions, each outstanding option tranche is valued with the dividend-adjusted Black–Scholes–Merton model at fiscal-year end from its own count, exercise price, and expiration date, and performance-unearned awards are excluded from the sensitivity calculations, as in their treatment of the post-2006 disclosure format. The main difference is the data source: we build the portfolio from Incentive Lab holdings tables and per-tranche records rather than from ExecuComp.

The construction uses ISS Incentive Lab, CRSP, Compustat, and Federal Reserve H.15 Treasury rates; no ExecuComp fields and no language-model input enter the measure. The unit of observation is the CEO–firm–calendar year. Repeated threshold, target, and maximum rows are collapsed before valuation, so a CEO-year is calculated once even when an award contributes several contractual cutoffs. The resulting data contain 1,551 economic events, 10,857 event-year bridge observations, and 5,026 unique CEO–firm–calendar-year calculations. Portfolio snapshots are requested for calendar years $t-3$ through $t+3$, where t is the calendar year in which the contractual performance period ends.

IA.4.1 Source Tables and Economic Scope

Table IA.12 summarizes the inputs. The measures cover reported outstanding holdings rather than all contractual compensation exposure. Current-year grant flows are not added to the outstanding portfolio, and performance-unearned stock and options are excluded. This scope prevents a new grant from being counted once as a grant flow and again in an outstanding-holdings table; it also means that the resulting delta is not a measure of every possible future payout under all contingent awards.

Table IA.12. Primary Inputs to CEO Delta and Vega

Source	Main fields	Role in construction
Relative-TSR event file	Firm CIK, participant identifier, performance-end date, grant/objective/period identifiers	Defines the economic event, the exact executive, and the seven-year event window. Repeated cutoff rows are collapsed before valuation.
Incentive Lab CompanyFY	Fiscal year, fiscal month, fiscal-year-end date, ticker, CUSIP	Supplies the authoritative valuation date and the native fiscal keys used to join other Incentive Lab tables.
Incentive Lab ParticipantFY	Participant identifier, current-CEO flag, title, role codes	Confirms that the exact event participant exists on the selected fiscal keys. Another executive is never substituted.
Incentive Lab BeneOwn	Beneficially owned shares, ownership percentages, separately reported Class B shares	Supplies common shares used for owned-stock delta; Class B shares are retained as a diagnostic but excluded without a known conversion ratio.
Incentive Lab OEStock	Time-unvested share count and provider value; separate performance-vesting fields	Supplies ordinary time-unvested stock. Performance-contingent unearned fields are excluded.
Incentive Lab OEOption	Exercisable and unexercisable counts, unearned count, exercise price, expiration date	Supplies each outstanding option tranche. Unearned options are excluded.
CRSP and historical links	Date-valid PERMNO, daily price, monthly total returns	Supplies the fiscal-year-end stock price and pre-period volatility.
Compustat	Exact-date fiscal-year-end price fallback, dividends per share	Supplies a restricted price fallback and the dividend yield.
Federal Reserve H.15	Annual Treasury constant-maturity yields at one-, two-, three-, five-, seven-, and ten-year maturities	Supplies the maturity-matched risk-free rate used in option sensitivity calculations.

Notes: This table reports the data sources for the CEO portfolio sensitivity calculations. All inputs are structured data; no language-model interpretation enters this construction.

IA.4.2 Event Time, Fiscal Time, and Executive Identity

The construction keeps three clocks distinct. The performance-end date is the exact contractual date on which the Relative-TSR performance period ends. The event calendar year is $t = \text{year}(\text{performance-end date})$, and portfolio snapshots are requested for $t+k$ with $k \in \{-3, \ldots, 3\}$. The valuation date is the Incentive Lab CompanyFY fiscal-year end that falls in the requested calendar year; the provider fields `fiscalYear` and `fiscalMonth` are native join keys and are not assumed to equal the event calendar year.

For firm i and requested calendar year y, we examine every CompanyFY row satisfying $\text{year}(\text{fiscalYearEnd}) = y$. A unique candidate is selected directly. When a fiscal transition creates multiple candidates, the firm's modal fiscal-year-end month defines its normal annual month; the candidate closest to that month is selected, and unresolved ties remain ambiguous. The selected native fiscal year and month are then used to join ParticipantFY, BeneOwn, OEStock, and OEOption.

The event participant identifier is required in every executive-level join. The current-CEO flag and title are retained as validation fields, but we do not replace a missing event participant with another CEO or senior officer. Calculation requests are deduplicated to CIK $\times$ participant identifier $\times$ calendar year, so one CEO-year calculation can be reused across several events.

IA.4.3 Market Inputs

Five market inputs enter the option formulas, and each is constructed so that it cannot embed information from after the valuation date. Table IA.13 summarizes the inputs and their look-ahead safeguards.

Stock price. The stock price S is the last positive CRSP closing price on or before the fiscal-year end, taken from a date-valid historical PERMNO, so a nontrading fiscal-year end uses the immediately preceding trading day and future prices are prohibited. If CRSP is unavailable, the absolute value of Compustat's fiscal-year-end price is used only when Compustat's data date exactly equals the selected fiscal-year end. The construction retains the price source, the price date, and the calendar-day distance from the fiscal-year end.

Return volatility. Annualized volatility σ is the sample standard deviation of up to 60 monthly CRSP total returns, multiplied by the square root of twelve, using only months that end strictly before the fiscal period being valued. At least 12 valid monthly returns are required for a firm-specific estimate; when fewer are available, the firm receives the same-calendar-year mean volatility of a broad, date-valid CRSP–Compustat reference universe. The completed value is clipped at the

reference universe's annual fifth and ninety-fifth percentiles, and a missing reference distribution never suppresses a firm-specific volatility that can be calculated directly.

Dividend yield. The decimal dividend yield q averages Compustat dividends per share divided by the fiscal-year-end price over the current and prior two data-date years, requiring nonnegative dividends and a positive price in each contributing year. The number of contributing years is retained, and the resulting yield is clipped at annual fifth and ninety-fifth percentiles. A reported zero dividend is a valid zero; a missing dividend field is not recoded as zero.

Risk-free rate and remaining maturity. Each option tranche's remaining maturity T_j is the number of calendar days from the fiscal-year end to the tranche's contractual expiration date, divided by 365.25, and must be positive; grant dates, vesting dates, and expected-life haircuts are never substituted for expiration. Federal Reserve H.15 annual averages supply Treasury constant-maturity rates at one-, two-, three-, five-, seven-, and ten-year maturities, with the missing four-, six-, eight-, and nine-year rates filled by linear interpolation between the adjacent published maturities. The tranche's maturity, rounded to the nearest whole year and clamped between one and ten years, selects the Treasury rate r_j; the reported percentage is converted to a decimal, and the unrounded maturity, not the rounded one, enters the Black–Scholes–Merton formulas.

Table IA.13. Market Inputs and Look-Ahead Safeguards

Input	Construction	Safeguard
S	Last positive CRSP price on or before fiscal-year end; exact-date Compustat fallback	No future trading date and no stale Compustat fallback from a different date.
σ	Up to 60 monthly returns strictly before the current fiscal period; annualized sample standard deviation	At least 12 observations for firm-specific volatility; otherwise a same-year reference mean, with annual tail clipping.
q	Mean of available decimal Compustat dividend yields over the current and prior two calendar data-date years	Zero is preserved as a real dividend outcome; missing is not zero.
r	Annual H.15 Treasury rate matched to rounded remaining contractual maturity	Percentage-to-decimal conversion is explicit; unrounded maturity remains in the option formula.
T	Contractual expiration minus fiscal-year-end valuation date	Must be positive; no vesting-date or expected-life substitute.

Notes: This table summarizes the five market inputs to the Black–Scholes–Merton sensitivities and the rule that prevents each input from embedding future information relative to the valuation date.

IA.4.4 Holdings Included in the Portfolio

The portfolio includes beneficially owned common stock, ordinary time-unvested stock, exercisable options, and unexercisable options; it excludes performance-unearned stock, unearned options, and current-year grant flows. The exclusions are explicit exclusions with zero contribution, not missing-data assumptions.

The owned-stock component uses Incentive Lab BeneOwn `sharesOwned` after exact matching on firm, native fiscal keys, and participant. Conflicting duplicate values are not added together; the component becomes missing. Separately reported Class B shares are retained as a diagnostic but are not added to priced shares without a known economic conversion ratio. Because a share moves one for one with the stock, the owned-stock delta is simply one percent of the stock price multiplied by the number of beneficially owned shares.

Time-unvested stock uses OEStock `numUnvested`. Every delivered source row is retained, including rows whose OEStock identifier is null; those rows receive stable synthetic identifiers. Performance-vesting fields `PVnumUnvested` and `PVvalueUnvested` are excluded. The component's delta is one percent of the stock price multiplied by the total time-unvested share count; a reported count of zero establishes zero delta, and a positive count requires a valid positive stock price.

Each OEOption row is treated as a separate valuation tranche. Exercisable and unexercisable option counts are included; `numUnearned` is always excluded. Rows with a null option identifier receive stable synthetic identifiers rather than being dropped. Current-year `optionAward` in the grant table is a flow and is never added to the outstanding option stock.

IA.4.5 Black–Scholes–Merton Option Sensitivities

For tranche j, let S be the stock price, K_j the exercise price, T_j remaining maturity, σ annualized volatility, q the dividend yield, and r_j the maturity-matched Treasury rate. Define

$$\begin{aligned} d_{1j} &= \frac{\ln(S/K_j) + (r_j - q + \frac{1}{2}\sigma^2)T_j}{\sigma\sqrt{T_j}}, \\ d_{2j} &= d_{1j} - \sigma\sqrt{T_j}. \end{aligned} \tag{IA.5}$$

The per-option price delta is

$$\delta_j = e^{-qT_j}\Phi(d_{1j}), \tag{IA.6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, and the per-option vega for a one-unit change in annualized volatility is

$$\nu_j = Se^{-qT_j}\phi(d_{1j})\sqrt{T_j}, \tag{IA.7}$$

where $\phi(\cdot)$ is the standard normal density.

Converting these per-option sensitivities into dollars follows the conventions of Coles, Daniel, and Naveen (2013). A tranche's dollar delta is one percent of the stock price, multiplied by the option count and by δ_j, computed separately for the exercisable and unexercisable counts; its dollar vega is 0.01 multiplied by the option count and by ν_j. Thus, a vega of \$25,000 means that the reported option portfolio increases in value by approximately \$25,000 when annualized volatility rises from, for example, 0.30 to 0.31, holding the other pricing inputs fixed. Stock contributes zero vega.

As an illustration, consider a hypothetical tranche with $S = 50$, $K = 40$, $T = 3$ years, $\sigma = 0.30$, $q = 0.02$, $r = 0.04$, and 100,000 options. The formulas give $d_1 = 0.8047$, per-option delta $\delta = 0.743531$, and unit-volatility vega $\nu = 23.53775$. The tranche's dollar delta for a one-percent stock-price increase is $0.01 \times 50 \times 100{,}000 \times 0.743531 = \$37{,}176.56$, and its dollar vega for a 0.01 volatility increase is $0.01 \times 100{,}000 \times 23.53775 = \$23{,}537.75$. This example is illustrative only; the calculations use each CEO-year's actual price, strike, contractual expiration, volatility, dividend yield, Treasury rate, and tranche counts.

IA.4.6 Tranche Aggregation and Portfolio Measures

Option delta and option vega are first summed across every delivered tranche, combining the exercisable and unexercisable contributions. The construction then reports three delta definitions, because different research questions require different components: an awards delta that adds time-unvested stock to the option portfolio, an owned-plus-options delta that adds beneficially owned stock to the option portfolio, and a headline full-reported delta that sums all three components. Each measure has its own validity flag. An unavailable owned-stock component does not erase an independently calculable option vega, and an unavailable OEStock component does not invalidate an owned-stock delta. The headline full-reported measure requires all three delta components. The portfolio vega is the sum of exercisable and unexercisable option vega; common stock and restricted stock contribute zero vega under this construction.

IA.4.7 Vested, Owned, and Unvested Portfolio Decompositions

For options, exercisable status is the observable vested component and unexercisable status the observable unvested component. OEStock supplies time-unvested stock but does not supply an exact count of vested restricted shares. The construction therefore uses beneficially owned shares as an explicitly labeled liquid/vested proxy; it does not claim that those shares are provider-reported vested award stock.

Two decompositions and two ratios follow. The vested-proxy delta adds owned stock to exercisable-option delta, and the unvested delta adds time-unvested stock to unexercisable-option delta. The option-composition ratio divides exercisable-option delta by unexercisable-option delta, and the broader vested-to-unvested ratio divides the vested-proxy delta by the unvested delta. A ratio is defined only when its denominator is strictly positive; a zero denominator remains missing rather than being replaced by an arbitrary large value. In event-time analyses, these ratios can be measured before performance realization so that portfolio composition is predetermined relative to the event. Any sample split or nonlinear transformation is an empirical choice and is not part of the measure's definition.

Component shares are also available. When full-reported delta is positive, owned stock, time-unvested stock, exercisable options, and unexercisable options are divided by full-reported delta, and the four shares sum to one. When portfolio vega is positive, exercisable- and unexercisable-option vega shares sum to one.

IA.4.8 Zeros, Missing Values, and Validity Rules

The construction treats zero and missing as different economic states. A reported option or stock count of exactly zero establishes a valid zero sensitivity for that component and does not require market inputs, whereas a missing count never becomes zero. A positive option count requires $S > 0$, $K > 0$, $T > 0$, $\sigma > 0$, $q \geq 0$, and an observed risk-free rate. A CEO-year option component is treated as valid only when every delivered tranche for that exercisability component is observed and either priced or affirmatively zero, and joint option delta and vega require both exercisable and unexercisable components across the delivered tranches. Conflicting beneficial-ownership duplicates are not summed; the owned-stock component remains missing. Missing future event years, invalid positive-count inputs, and unresolved fiscal-year mappings remain missing.

This component-specific logic explains why coverage counts are not a mechanically nested waterfall. For example, a CEO-year with exactly zero option counts has a valid zero option sensitivity even if no Treasury rate is available, so the number of valid option delta observations can

exceed the number of CEO-years with a retrieved risk-free rate. Table IA.14 reports coverage by stage and component.

Table IA.14. Coverage of the CEO Portfolio Construction

Stage or component	CEO-years	Percent of 5,026 requests
CompanyFY mapping	4,306	85.67%
Exact event participant found	3,373	67.11%
Beneficial-ownership source available	3,213	63.93%
OEStock source available	2,994	59.57%
OEOption source available	2,817	56.05%
CRSP mapping available	4,167	82.91%
Stock price available	4,179	83.15%
At least 12 months of firm-specific volatility	4,068	80.94%
Volatility available after reference fallback	5,023	99.94%
Dividend yield available	4,693	93.37%
Risk-free rate available	1,189	23.66%
Option delta valid	2,423	48.21%
Portfolio vega valid	2,423	48.21%
Time-unvested-stock delta valid	2,644	52.61%
Awards delta valid	2,345	46.66%
Full-reported delta valid	2,279	45.34%

Notes: Percentages use all 5,026 requested CEO–firm–calendar-year calculations. Coverage is component-specific rather than a nested waterfall: a component with an affirmative zero count is valid without market inputs. Valid option delta and vega equal 2,423 of the 2,817 CEO-years with OEOption rows, or 86.01 percent.

IA.4.9 Reproducibility and Interpretive Limits

Each option and stock source row has a stable identifier. A reported identifier is used when available; otherwise, a deterministic synthetic identifier preserves the row rather than dropping it. Tranche-level records retain the option counts, strike price, expiration date, remaining maturity, risk-free rate, market-input validity, d_1, per-option sensitivity, and dollar contribution. CEO-year records retain source availability, fiscal-date mapping, price and volatility provenance, duplicate and overlap indicators, component-specific validity flags, and quality classifications. These fields permit every aggregate sensitivity measure to be traced to its underlying holdings and market inputs.

Three interpretive limits are central. First, the full-reported delta can contain unresolved overlap between beneficial ownership and near-term exercisable option exposure, because proxy beneficial-ownership definitions can include securities acquirable within a short horizon. The overlap is recorded as a warning; it does not invalidate option vega, but the full-reported delta should be described as reported portfolio exposure rather than a guaranteed non-overlapping total. Second, OEStock does not provide an exact vested-stock count, so the stock vested/unvested diagnostic uses

owned shares as an explicit liquid/vested proxy and must not be described as an exact restricted-stock vesting ratio. Third, the portfolio measures are stocks of outstanding exposure measured at fiscal-year end. They are not the same economic object as the quarterly flow of equity scheduled to vest in Edmans, Fang, and Lewellen (2017). The core option-pricing logic is compatible, since each option quantity is converted into share-equivalent delta and multiplied by the stock price, but the observation time, option quantities, and economic interpretation differ. The delta measure captures contemporaneous outstanding portfolio price sensitivity around Relative-TSR event time; it is not, by itself, a vesting-flow instrument or a direct measure of short-termism.

In empirical files, dollar delta always refers to the wealth effect of a one-percent stock-price increase, and dollar vega always refers to the wealth effect of a 0.01 absolute increase in annualized volatility. The dollar measures may be reported in thousands and transformed as $\ln(1+x)$; these rescalings are mechanical and do not change the underlying economic definition.

Summary of Measurement Conventions

The constructions share five principles. First, the economic unit is defined before aggregation: grant–objective–period–cutoff for Relative TSR, application assignee for patent identity, filing or filing section for text-based measures, and CEO–firm–calendar year for portfolio sensitivities. Second, artificial intelligence is used for interpretation and identity resolution, while numerical calculations remain deterministic. Third, evidence and provenance are retained for every record that requires judgment, whether the judgment is an agent's method interpretation, an agent's identity adjudication, or a batch classification. Fourth, structural zeros, missing values, and assumption-sensitive observations remain distinct. Fifth, the numerical transformations are reconciled to the underlying dates, securities, holdings, and return paths through independent validation.

These safeguards do not eliminate measurement error, but they identify its source. Relative-TSR observations indicate whether the cutoff is point identified, interval identified, or dependent on a disclosed assumption. Patent links distinguish direct name matches from agent-reviewed links and leave unsupported identities unmatched. Filing-based measures retain their classification basis and preserve unresolved states. Delta and vega carry component-specific validity and overlap flags. The measures can therefore be restricted or reclassified in sensitivity analyses without changing their underlying economic definitions.

References

Bettis, J. C., J. Bizjak, J. L. Coles, and S. Kalpathy. 2018. "Performance-Vesting Provisions in Executive Compensation." *Journal of Accounting and Economics* 66(1): 194–221.

Black, F., and M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81(3): 637–654.

Coles, J. L., N. D. Daniel, and L. Naveen. 2006. "Managerial Incentives and Risk-Taking." *Journal of Financial Economics* 79(2): 431–468.

Coles, J. L., N. D. Daniel, and L. Naveen. 2013. "Calculation of Compensation Incentives and Firm-Related Wealth Using Execucomp: Data, Program, and Explanation." Working paper.

Core, J., and W. Guay. 2002. "Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility." *Journal of Accounting Research* 40(3): 613–630.

De Angelis, D., and Y. Grinstein. 2015. "Performance Terms in CEO Compensation Contracts." *Review of Finance* 19(2): 619–651.

Edmans, A., V. W. Fang, and K. A. Lewellen. 2017. "Equity Vesting and Investment." *Review of Financial Studies* 30(7): 2229–2271.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *Quarterly Journal of Economics* 132(2): 665–712.

Merton, R. C. 1973. "Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science* 4(1): 141–183.

Documentation of Human and AI Contributions for Taking Stock: Do Equity-Based Managerial Incentives Drive Corporate Innovation?

Mark A. Chen, Erol Oytun Ercan, Yaswanth Pothuru

AFA 2027 Special Session of Papers Written with Generative AI Workflows

This appendix describes how agentic AI and generative AI were used across the project. We begin by outlining the main research tasks in which AI played a role, followed by the agentic workflows and sub-agent structure developed for those tasks, the LLM prompts embedded in the analysis code, and the available conversation and execution records. We then report the documented time spent by us (human authors) on the project and summarize the relative human and AI contributions to the project's computational code, writing, and documentation. Together, these sections provide a detailed record of how AI was incorporated into the research process, how we interacted with those systems, and how the work underlying the final paper was divided between human and AI contributions.

1 Overview

We designed the workflow for this research project with the aim of using AI iteratively and interactively rather than treating the AI as a one-shot tool that received instructions and returned a final output. AI systems participated systematically across the project, from data acquisition and document interpretation to entity resolution, variable construction, validation, empirical analysis, literature review, manuscript preparation, and research documentation. Much of the task execution within these stages was delegated to AI agents and LLM workflows, while human effort was concentrated on a smaller set of tasks for which economic judgment was most valuable, such as defining the underlying economic concepts, resolving genuinely ambiguous cases, evaluating whether intermediate outputs were economically sensible, and deciding when the research design itself required revision. The resulting workflow therefore sought to maximize research quality per unit of human effort rather than either minimizing AI involvement or eliminating human judgment altogether.

Our iterative and interactive approach contrasts with both a fully autonomous one-shot approach and a multi-shot, non-interactive research process. Recent work by other researchers focuses on non-interactive approaches. For example, Lopez-Lira (2026) demonstrates through ZeroPaper that a multi-agent system can progress from an initial research domain to a completed manuscript without human intervention between launch and completion. Galiani, López, and Sosa (2026) provide controlled evidence for a multi-shot AI approach in econometric coding. When AI agents are allowed to execute their work, inspect the resulting output, and revise their implementation, output correctness increases from 67.3% to 94.9% under zero-shot prompting and from 81.6% to 96.8% under few-shot prompting. Our approach differs from this type of non-interactive iteration by incorporating human evaluation and judgment at key stages, allowing researchers to assess ambiguous outputs, refine instructions, and guide subsequent iterations of the AI workflow.

The resulting division of labor was heavily weighted toward AI in the production of observable research outputs. Based on our contribution records, agentic AI systems generated 97.4% of the project code and 83.1% of the writing and documentation. Human contributions were concentrated primarily in the conceptual and organizational dimensions of the project, including formulating the research question, defining the relevant economic concepts, determining the empirical strategy, evaluating ambiguous cases, and deciding when the research procedures required revision. Thus, although we retained responsibility for the principal intellectual and organizational decisions, AI systems performed the large majority of the coding, writing, documentation, data-processing, and implementation work.

The models helped us collect, parse, organize, and collate information from large numbers of filings, corporate records, web sources, and other documents; apply common definitions across those materials; identify difficult or unusual observations; and produce intermediate outputs that we could inspect. This was particularly useful when the relevant information was publicly available but dispersed across many documents, or could not be converted into a research variable through a simple rule. Difficult cases often showed us where an initial definition, prompt, or procedure had been incomplete. We reviewed these difficult cases, clarified the instructions, and applied the revised procedure to the relevant observations. In this way, what we learned from the AI outputs at each stage affected how we subsequently used the AI.

Our approach deliberately retained researcher involvement throughout the process. To assist with the judgement-intensive phases of the project, AI provided scale, contextual interpretation, document processing, and repeated scrutiny, while we defined the economic concepts, examined ambiguous cases, revised the procedures, and determined how the resulting measures entered the research. The human–AI interaction was therefore not simply a final validation step; it was present in many intermediate steps throughout the entire project development process.

We uses human–AI approach in several critical phases of the project. For example, agentic AI was used to resolve difficult USPTO-to-public-firm name matches, reconstruct the economic structure of Relative-TSR incentive contracts, build and validate research datasets from scratch,

and execute the empirical analysis. Purpose-built Agentic + LLM workflows helped to extract and interpret information from DEF 14A proxy statements and 10-K MD&A disclosures, including compensation-contract provisions, compensation-committee discretion, CEO ownership requirements, share pledging, pledging restrictions, anti-hedging policies, and managerial myopia. We also used agentic workflows in the literature review, manuscript preparation, and research documentation. Table 1 summarizes the main research tasks in which we used AI, and Figure 1 shows the detailed workflow schema.

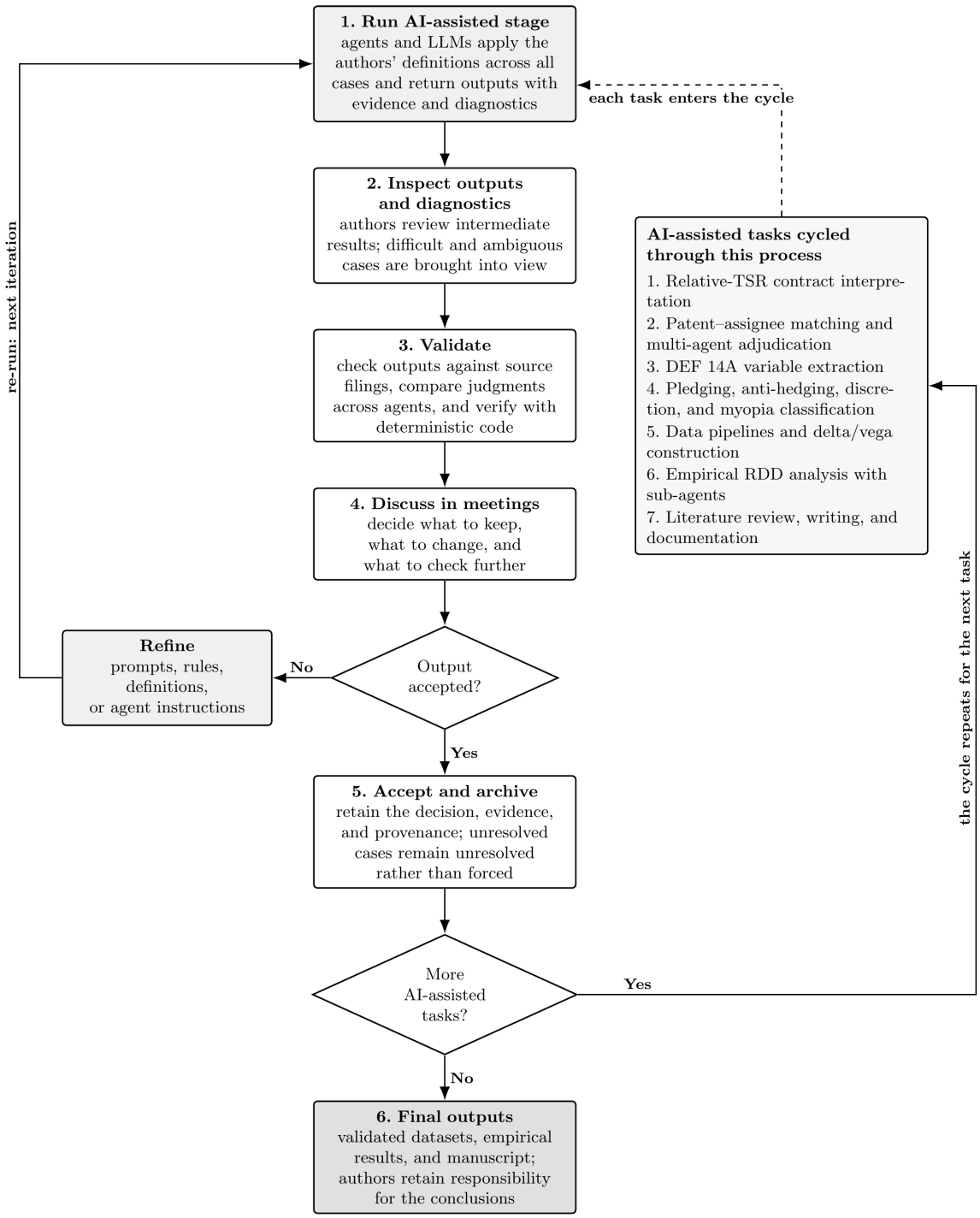


Figure 1: The iterative process applied to each AI-assisted research task.

Figure 1 summarizes the interactive and iterative process we used across the AI-assisted parts of the project. In broad terms, the process had three stages: (1) an AI-assisted task was run to collect, organize, extract, classify, match, or calculate the required information; (2) we inspected the intermediate outputs, validated them against the underlying source material and deterministic checks where appropriate, and discussed difficult or ambiguous cases; and (3) we either refined the definitions, prompts, or procedures and reran the affected stage, or accepted the output and retained the supporting evidence and provenance. This process was repeated across tasks rather than treating the first model response as the final research output.

The scale of both the human and AI contributions was also evident in the quantitative record of the project. AI-assisted workflows produced 28,615 lines of computational code, compared with 752 lines written directly by us. The patent-matching workflow involved automatic checking of 112,852 webpages while investigating company identities and corporate relationships that could not be established reliably from names alone.

The Scale of the Project

The scale of the underlying data made these tasks difficult to address through manual review alone. The Incentive Lab data contained 778,858 awards across 4,802 firms and covered fiscal years 1998–2024, giving us more than two decades of compensation information from which the Relative-TSR sample was constructed. Our filing-based work was similarly large. We assembled and processed 47,270 DEF 14A proxy statements, from which the LLM workflows recovered 10 CEO-specific contractual and governance variables. The initial extraction identified 24,225 observations involving compensation-committee discretion, which were then subjected to a separate classification of the type and direction of that discretion. We also processed 4,654 10-K MD&A disclosures for the managerial-myopia analysis.

The scale was important, but document volume alone did not explain why AI was useful. The harder problem was that the information we needed was often dispersed across long filings, tables, footnotes, analyst notes, corporate websites, and historical records, and firms frequently described

economically similar arrangements in different ways. The task therefore required both processing a large volume of information and interpreting that information consistently enough to turn it into research variables.

Table 1: Main research tasks in which we used AI

Research task	Data / source material	AI approach	Main challenges addressed
USPTO–Compustat name matching	USPTO patent-application assignees; Revelio Labs corporate hierarchy; CRSP/Compustat; public corporate information	Fuzzy matching was followed by agentic review, web search, and additional review for difficult cases.	Subsidiaries, acquisitions, former names, and other corporate relationships often could not be identified from string similarity alone. Manual investigation across the assignee universe would have required extensive firm-by-firm research.
Information extracted from DEF 14A filings • Relative-TSR contract structure and calculation rules • Contractual gates and caps • Compensation-committee discretion • CEO ownership guidelines and requirements • CEO share pledging • Pledging restrictions • Anti-hedging policies	DEF 14A proxy statements; Incentive Lab award information and analyst notes	Purpose-built LLM extraction and classification prompts were used, with additional classification when the first-stage extraction required further interpretation.	Important contractual and governance information was dispersed across lengthy filings and described differently across firms. Several required contract terms and governance variables were incomplete or unavailable in standardized Incentive Lab fields, so the underlying disclosures had to be interpreted directly.
Managerial myopia	Item 7, MD&A, from Form 10-K filings	LLM-based interpretation and classification used the surrounding MD&A context together with supporting evidence.	Short-term words or discussion of current performance did not by themselves establish managerial myopia. The classification required distinguishing ordinary operating commentary from evidence that management prioritized near-term results at the expense of longer-term investment.

Continued on the next page.

Table 1 (continued)

Research task	Data / source material	AI approach	Main challenges addressed
Sample construction and empirical analysis	Compensation, patent, return, governance, CEO, filing, and other project datasets; intermediate regression outputs	Agentic coding was used for data construction, merging, validation, regression implementation, and repeated empirical checks.	The project combined large datasets with different identifiers, units of observation, timing conventions, and measurement rules. Repeated construction and validation would have been difficult to perform consistently through manual coding and inspection alone.
Literature review	Academic papers, working papers, project materials, and the research question	Multiple agents were used for literature search, organization, synthesis, and citation checking.	Several related literatures had to be searched and connected to the paper while preserving relevance to the research question and checking the underlying citations.
Manuscript and appendix preparation	Research results, tables, figures, notes, existing drafts, and supporting documentation	LLM- and agent-assisted drafting, restructuring, comparison, and iterative revision were used.	The manuscript and appendix changed repeatedly as the data and empirical analysis developed, while numerical results, variable definitions, tables, and interpretation had to remain consistent across sections.

Matching Patent Assignees to Public Firms

Patent-to-firm matching illustrated a particularly difficult research problem. As Lerner and Seru (2022) noted, patents could be assigned to subsidiaries or other legal entities whose names differed substantially from the publicly traded parent. The difficulty was easy to see in familiar corporate relationships. A historical assignee such as LinkedIn could ultimately map to Microsoft following acquisition, while Whole Foods Market could map to Amazon. In both examples, the subsidiary or acquired-company name could bear little resemblance to the current public parent. A procedure that relied only on comparing the two names could therefore miss a valid economic link. The reverse problem also arose: two companies could have similar names while having no meaningful corporate relationship. The matching problem was therefore one of identifying the company behind the assignee rather than simply calculating similarity between two strings.

We addressed this problem in two broad stages. First, conventional matching methods cleaned and normalized the assignee names, searched the company universe, and produced plausible candidate firms. This stage was efficient at dealing with spelling differences, punctuation, legal suffixes, and other ordinary name variation. Second, cases that could not be established confidently from the name alone were sent to agentic review. The agents investigated the corporate relationship using subsidiary information, acquisitions, historical names, company websites, SEC filings, and other public evidence. The web-based matching process consulted 112,852 webpages as part of this investigation. The purpose of this second stage was not simply to accept the candidate with the highest fuzzy score. It was to determine whether the available evidence supported a genuine economic relationship and to leave the case unresolved when the evidence was insufficient.

The conventional fuzzy-matching stage linked 29,043 U.S. assignees to public firms, and the agentic review added another 6,962 links that the initial procedure did not resolve. The value of the agentic stage therefore came from addressing a limitation of conventional name matching: the economically relevant relationship between two corporate entities was often contained in their history or organizational structure rather than in the similarity of their names.

Understanding the Structure of Incentive Contracts

A critical part of our research design depended on reconstructing the exact compensation contracts that generated the Relative-TSR payout cutoffs. These contracts were highly heterogeneous and were not available in a fully standardized form in Incentive Lab. Incentive Lab provided important award information and analyst notes, but several details needed for our empirical design were incomplete or unavailable in structured fields. The missing or unstandardized information could include the exact benchmark, performance-period convention, averaging windows, return calculation, dividend treatment, rank or percentile method, payout interpolation, and additional gates or caps. Recovering these terms was therefore necessary before the contractual cutoff and the running variable used in the RD design could be calculated.

The scale of this problem was substantial. We processed 47,270 DEF 14A proxy statements, and the relevant information could appear in compensation tables, footnotes, plan descriptions, analyst notes, or other parts of the underlying filing. What we needed from these filings was not simply a sentence containing "TSR," "peer group," or "threshold." We needed to reconstruct the economic rule that determined how realized performance was converted into shares.

We again separated the task into two stages. First, the LLM interpreted the available filing and Incentive Lab evidence and converted the contractual language into structured fields describing how the award operated. This included identifying the performance metric, benchmark, measurement window, TSR convention, rank or percentile rule, payout landmarks, and other contractual provisions. Second, once the contract had been represented in a structured form, deterministic code implemented the rule. The code calculated issuer and benchmark returns, applied the relevant contractual convention, mapped realized performance to the threshold, target, or maximum, and constructed the running variable used in the empirical analysis. In other words, AI was used for the part of the problem that required interpreting an unstructured contract, while deterministic computation was used for the numerical calculation implied by that interpretation.

The same problem arose for the other DEF 14A variables. A filing could disclose that a compensation committee had the authority to exercise discretion without showing that the committee had actually used that discretion. A CEO could have pledged shares even when the firm had no disclosed pledging restriction, while a firm could prohibit pledging even when no shares were currently pledged. Corporate hedging of interest-rate, foreign-exchange, or commodity risks was economically different from an anti-hedging policy restricting executives from hedging company stock. These distinctions depended on the meaning of the surrounding disclosure rather than the presence of a particular word. The LLM prompts were therefore used to classify the underlying economic concept and retain supporting filing evidence rather than simply extract sentences.

Managerial myopia presented the same issue in the 10-K setting. We processed 4,654 MD&A disclosures, but a mechanical count of words such as "quarter," "current," "short-term," or "long-term" would not have identified managerial short-termism. A 10-K was naturally required to discuss

current earnings, costs, margins, and operating conditions. What mattered for our measure was whether the disclosure indicated an intertemporal tradeoff in which management appeared to protect near-term financial performance at the expense of longer-term investment or capability. The LLM therefore evaluated the MD&A in context and returned the classification together with the rationale and supporting evidence.

How the Agentic Approach Helped

The agentic approach was useful because the principal problems in this project could not be reduced to a single prediction from a single input. Each problem contained a part that conventional computation handled well and another part that required contextual judgment.

For patent name matching, the first stage was computational: conventional code normalized names and searched a very large company universe for plausible matches. That approach worked well when the difference between the assignee and public-company names was largely textual. It did not solve the harder question of whether an assignee was a subsidiary, former name, acquired entity, affiliate, or unrelated company with a similar name. The second stage therefore used agents to investigate the underlying corporate identity and evaluate public evidence. This division allowed deterministic methods to narrow a very large search problem efficiently while reserving agentic review for the observations that required judgment.

For Relative-TSR contracts, the two stages served a different purpose. The first stage required interpretation: the LLM read the unstructured contract information and reconstructed the rule that governed the award. The second stage required calculation: deterministic code applied that rule to the issuer and benchmark returns and produced the contractual performance measure. A purely text-based procedure could have extracted sentences without reconstructing the economic rule, while a purely deterministic calculation could not have been run until the contractual terms had first been understood. Combining the two allowed the interpretation and calculation stages to be handled by the method better suited to each.

For managerial myopia and the other filing-based measures, traditional keyword searches could have located passages containing words associated with short-termism, pledging, hedging, discretion, or ownership requirements, but they could not reliably determine what those passages meant. The agentic and LLM workflows instead evaluated the surrounding disclosure, applied the economic definitions we specified, retained the evidence supporting the classification, and allowed unusual cases to be reviewed separately. The same logic extended to sample construction and empirical analysis, where agents helped write and validate code, compare intermediate outputs, and identify inconsistencies that could then be corrected and rerun.

This division of tasks also made the process iterative. Some problems became visible only after the procedure had been applied to many observations. A subsidiary relationship could reveal that a name-matching rule was too restrictive; an unusual Relative-TSR contract could expose a missing contractual category; or an MD&A could show that a myopia prompt was treating ordinary short-term operating discussion too mechanically. When these cases appeared, we returned to the original evidence, revised the relevant definition or procedure, and reran the affected observations. The purpose of iteration was therefore not to continue prompting until the model produced a preferred answer. It was to use the difficult cases to improve the rule that was then applied across the data.

Our approach was related to the growing use of AI in academic research, including AI-powered finance scholarship and autonomous multi-agent research systems (Novy-Marx and Velikov, 2026; Lopez-Lira, 2026; Mo and Ouyang, 2025). The feature we emphasized was continued researcher interaction. AI allowed us to process much larger amounts of source material, investigate cases that conventional methods could not resolve, and apply complex definitions repeatedly. We remained responsible for defining the economic objects, examining ambiguous observations, refining the procedures, and determining how the resulting measures entered the research. The contribution of AI to the project therefore came from the combination of scale, contextual interpretation, deterministic computation, and continued human–AI interaction.

2 AI-Assisted Research Workflows

AI entered the project through two related but distinct types of research workflows. For tasks that required several linked decisions, such as resolving the identity of a patent assignee, reconstructing a Relative-TSR contract, constructing CEO incentive measures, or implementing and checking the empirical design, we used agentic workflows. These workflows separated a larger research problem into stages that could be constructed, reviewed, validated, and revised independently. For tasks in which the main challenge was interpreting unstructured disclosure, we used purpose-built LLM prompts. Those prompts translated the underlying economic concept into explicit instructions about the relevant unit of analysis, the evidence that could be used, the distinctions the model had to preserve, the treatment of missing information, and the format of the resulting research variable.

The two approaches were complementary. Agentic workflows helped organize multi-stage research tasks in which different forms of evidence and computation had to be combined, while the LLM prompts provided a common measurement rule that could be applied repeatedly across thousands of filings and observations. In both cases, the resulting outputs were retained with intermediate evidence or diagnostics where available and were reviewed before they entered the final research datasets or empirical analysis.

2.1 Agentic Workflows

We used agentic workflows when a research task could not be handled reliably as a single model call or a single deterministic program. Each workflow was built around a specific research object and divided the work into a sequence of smaller tasks. Depending on the application, those tasks included identifying candidate observations, interpreting source material, writing or executing code, checking intermediate results, comparing alternative classifications, and conducting independent validation. The purpose of the agentic structure was therefore not simply to increase automation. It allowed different parts of a difficult research problem to be examined separately and allowed a particular stage to be revised without treating the entire preceding workflow as a black box.

The structure differed across applications. In the patent-matching workflow, deterministic matching first narrowed a very large company universe to plausible candidates, after which agents investigated corporate identity and relationships using public evidence. In the Relative-TSR benchmark workflow, agents helped reconstruct the contract-specific performance rule and validate the resulting calculation structure. The Incentive Lab and CEO Delta/Vega pipelines focused more heavily on data construction, measurement consistency, and validation, while the empirical pipeline implemented the prespecified RD analysis and subjected the resulting outputs to repeated checks. Separate agentic workflows were also used to organize the literature review and to develop and revise the manuscript and supporting appendix.

Where a pipeline specification formally named specialized sub-agents, we reported those roles in Table 2. These labels described the division of responsibilities within the workflow rather than the number of times an agent was called. Where the available execution records showed additional review activity but no formal named roster had been specified, we reported the roster as not separately specified rather than assigning role names after the fact. Across the pipelines with explicit rosters, the documented roles covered data and schema review, contract interpretation, empirical construction, validation, finance review, literature search, synthesis, writing, and citation checking.

Table 2: Agentic workflows used in the project

#	Agent / Pipeline	Main agent purpose	Declared sub-agents	Execution record
1	Patent-Assignee Name-Matching and Entity Resolution Pipeline	Linked USPTO patent-application assignees to the economically relevant public firm. Deterministic matching generated candidate firms, while agentic review investigated subsidiaries, acquisitions, former names, and other corporate relationships that could not be resolved from name similarity alone.	Not separately specified	click here

Continued on the next page.

Table 2 (continued)

#	Agent / Pipeline	Main agent purpose	Declared sub-agents	Execution record
2	Incentive Lab Data Construction and Descriptive Analysis Pipeline	Constructed and reconciled the Incentive Lab award universe, defined the relevant award and objective structures, produced the research sample and descriptive analysis, and subjected the resulting data and report to independent review.	*6 declared* schema_cartographer contract_taxonomist publication_designer analysis_engineer qa_auditor finance_referee	click here
3	Relative-TSR Contract and Benchmark Reconstruction Pipeline	Reconstructed the contract-specific Relative-TSR calculation, including the performance period, benchmark or peer group, return convention, ranking method, and contractual payout cutoffs, and validated the resulting performance measures.	*3 declared* schema_engineer method_engineer validator	click here
4	CEO Delta/Vega and Incentive Duration Measurement Pipeline	Constructed CEO portfolio delta, vega, vesting, and incentive-duration measures from the underlying holdings and award data and checked the resulting measures against a common set of economic and timing rules.	*1 or more declared* validator	click here
5	Relative-TSR Regression-Discontinuity Empirical Analysis Pipeline	Implemented the specified sharp RD design, constructed the analysis samples, produced the empirical tables and figures, and applied reproducible quality-control checks to the reported results.	Not separately specified	click here
6	Literature Review Pipeline	Searched and organized the related literature, assigned specialized reviews of the relevant research areas, synthesized the resulting evidence, drafted the literature discussion, and checked the supporting citations.	*7 declared* project_mapper incentives_innovation_scout performance_contracts_scout thresholds_identification_scout synthesis_referee literature_writer citation_auditor	click here
7	Manuscript and Appendix Development and Revision Pipeline	Used the completed empirical results, tables, figures, and supporting materials to develop and revise the manuscript and appendix while preserving the underlying numerical results and empirical analysis.	Not separately specified	click here

The pipelines were deliberately specialized rather than interchangeable. The patent-matching agent was designed around entity resolution and external corporate evidence; the Relative-TSR pipeline was designed around reconstructing contract-specific economic rules; the Delta/Vega

pipeline was designed around measurement and event-time alignment; and the empirical pipeline was designed around implementation and validation of the RD analysis. The literature and writing workflows operated on a different part of the research process altogether. This specialization allowed the instructions, validation rules, and intermediate outputs to be tailored to the research problem being addressed rather than imposing one general-purpose agent on every task.

The execution logs linked in Table 2 preserved the available record of these workflows. Depending on the pipeline, they documented source and schema review, intermediate artifacts, agent handoffs, validation passes, code-line counts, execution times, and final outputs. The level of detail differed across pipelines because the underlying workspaces retained different forms of execution history. We therefore treated the logs as records of the evidence actually preserved by each workflow rather than attempting to reconstruct missing calls, timestamps, or sub-agent activity after the fact.

2.2 LLM-Based Extraction and Classification

We used purpose-built LLM prompts when the research variable depended on the meaning of unstructured disclosure rather than on a standardized database field. The prompts were therefore treated as part of the measurement design. Each prompt specified the economic object being measured, the relevant observation or executive, the evidence the model was permitted to use, distinctions that had to be maintained, rules for missing or ambiguous information, and a structured output format.

This mattered because many of the filing-based variables in the project were easy to confuse with superficially similar concepts. Committee authority to exercise discretion was different from actual use of that discretion. A CEO's pledged shares were different from a corporate policy restricting pledging. Hedging company stock was different from the firm's use of derivatives to hedge operating risks. Discussion of quarterly earnings was not by itself evidence of managerial myopia. Similarly, reconstructing a Relative-TSR contract required identifying the economic rule

governing the award rather than merely retrieving a sentence containing "TSR" or "peer group." The prompts encoded these distinctions directly so that the same economic definition could be applied across filings even when firms used different language.

The prompt archive reproduced the instructions that were recoverable directly from the project source files. The principal filing-based prompt families covered CEO-specific DEF 14A variable extraction, pledging and anti-hedging policies, compensation-committee discretion, and managerial myopia from MD&A disclosures. The patent-assignee matching code also preserved a separate web-validation prompt that instructed the model to investigate corporate identity using public evidence rather than accepting a match on fuzzy similarity alone.

2.2.1 DEF 14A Variable Extraction

The DEF 14A extraction prompt was used to recover CEO-specific contractual and governance information from proxy statements. The workflow required the model first to identify the focal CEO and then to restrict every award-specific classification to that executive. It covered ten research variables, including absolute-TSR gates, negative-TSR caps, accounting gates, TSR measurement and averaging conventions, dividend treatment, committee discretion, actual positive or negative use of discretion, and CEO ownership requirements. A central feature of the prompt was its treatment of missing information. Silence in a filing was not automatically interpreted as evidence that a provision did not exist. Where the requested information could not be established from the filing or confidently linked to the focal CEO, the model was instructed to return Info_not_Found. The prompt also required short supporting evidence so that the resulting classification could be checked against the source disclosure.

Complete text of SYSTEM_PROMPT (system):

```
You extract CEO-only compensation variables from one DEF 14A proxy statement.
Use only the supplied filing. Do not use outside knowledge.

SCOPE RULES
1. First identify the focal CEO by full name and an explicit CEO title in the filing.
2. Every award-specific field must come from that CEO's performance-vesting award,
```

```
   or from a plan/provision explicitly governing that CEO award.
3. Never substitute a CFO, COO, President-only executive, another NEO, director,
   executive-group value, or another manager's award terms.
4. A universal executive policy may be used for the CEO only when the filing clearly
   establishes that the CEO is covered.
5. For flag variables use exactly one of: yes, no, Info_not_Found.
   - yes = the filing affirmatively establishes the condition.
   - no = the filing affirmatively establishes the opposite/absence.
   - Info_not_Found = the filing does not provide enough evidence to establish the answer,
     including cases where the relevant information is missing, ambiguous, conflicting,
     or cannot be tied confidently to the focal CEO.
   Do NOT convert silence into "no".
6. For text variables return a concise filing-based description, or exactly
   Info_not_Found when the requested information cannot be established from the filing.
7. For positive_discretion_used and negative_discretion_used, distinguish actual use
   of discretion from mere authority to use discretion. Authority alone is not use.
8. Keep evidence short and auditable: one concise quotation or near-verbatim snippet.
   If Info_not_Found, use an empty evidence string.
9. If multiple CEOs served during the year, match the CEO to the relevant award,
   performance period, certification date, or policy measurement date. If that cannot
   be resolved, use Info_not_Found rather than choosing another executive.

VARIABLE DEFINITIONS AND EXAMPLES

1. absolute_tsr_gate
   Whether absolute TSR must be positive or above a minimum before relative-TSR payout
   is allowed.
   Example: "No payout is made unless absolute TSR is positive over the performance period."

2. negative_tsr_cap
   Whether payout is capped, often at target, when absolute TSR is negative even if
   relative TSR rank is high.
   Example: "If absolute TSR is negative, payout for the TSR component is capped at 100% of target."

3. accounting_gate
   Whether an accounting, liquidity, safety, profitability, or similar condition must
   be met before payout can occur.
   Example: "No payout will be made unless adjusted EBITDA exceeds the threshold level."

4. tsr_measurement_method_and_averaging_window
   Extract the exact TSR/stock-price measurement method: cumulative TSR, beginning or
   ending average-price windows, trailing averages, base-date prices, tranche hurdles,
   repeated measurement dates, and any relative-index cutback overlay.
   Example: "The TSR PSUs will fund progressively as the 90-day trailing average stock
   price appreciation exceeds performance tranches ... above the 15-day trailing average
   stock price as of the Base Date."

5. dividend_reinvestment_or_tsr_calculation_rule
   Extract whether TSR includes dividends/reinvested dividends and whether it uses
   cumulative vs. annualized TSR or average-price windows.
   Example: "TSR is calculated assuming all dividends are reinvested."

6. committee_discretion_authority
   Whether the compensation committee reserves authority to adjust payout based on
   judgment, other factors, or extraordinary circumstances.
   Example: "The Committee may adjust actual payout amounts in its discretion."

7. positive_discretion_used
   Whether the board/committee actually increased payout relative to the otherwise
   applicable formula, revised targets downward, or allowed payout despite missing the
   original formula. Do not mark yes for a predefined formulaic adjustment that did not
   involve an actual discretionary decision.
   Example: "Without these adjustments, no payout would have been earned."

8. negative_discretion_used
   Whether the board/committee actually reduced payout relative to formula, capped it
```

```
   below the formulaic result, or applied a risk/compliance reduction.
   Example: "The Committee reduced the formulaic payout by 20%."

9. ceo_ownership_guideline_flag
   Whether the CEO is subject to a formal minimum stock-ownership guideline or
   requirement. Distinguish this from a nonbinding aspiration or a separate holding rule.
   Example: "The Chief Executive Officer is required to maintain stock ownership equal to 5x annual base
     salary."

10. ceo_ownership_requirement_level
    Extract the numeric ownership threshold AND its unit. Examples include a salary
    multiple or a required number of shares. Do not force conversion to a multiple when
    the filing gives a different unit.
    Example: "Chief Executive Officer - 6x base salary."

OUTPUT RULES
- Return one JSON object matching the supplied schema.
- No markdown, commentary, or text outside the JSON.
- CEO name/title are audit metadata, not additional research variables.
- STRICT MISSING-INFORMATION RULE: whenever the requested information is not found or
  cannot be established from the supplied DEF 14A, return exactly "Info_not_Found".
  Do not return NOT_DISCLOSED, UNCLEAR, N/A, NA, unknown, null, or any other substitute.
```

2.2.2 Share-Pledging and Anti-Hedging Policy Extraction

The pledging and anti-hedging prompt was designed to identify corporate policy restrictions rather than actual trading or pledging behavior. A disclosure showing that an executive had pledged shares did not itself establish that the company maintained a pledging restriction, and a discussion of the company's commodity, interest-rate, or foreign-exchange hedging did not establish an anti-hedging policy covering executives. The prompt therefore treated pledging restrictions and anti-hedging restrictions as separate research variables and required each to be classified independently from the full DEF 14A. It also provided explicit rules for exceptions, approval requirements, margin arrangements, derivative transactions, and cases in which the filing was silent. The model returned both the classification and the filing language supporting a positive determination.

Complete text of SYSTEM_PROMPT (system):

```
Determine whether one U.S. DEF 14A proxy statement discloses each of the
following two governance-policy restrictions:

A. A BAN OR RESTRICTION ON PLEDGING the company's securities.
B. An ANTI-HEDGING BAN OR RESTRICTION covering hedging of the company's
   securities.

These are FILING-LEVEL POLICY CLASSIFICATIONS. They are not tasks about
whether a particular CEO, director, executive, employee, or shareholder
actually pledged or hedged securities, and they are not tasks about the
number, value, class, or purpose of any shares involved.

Use only the supplied DEF 14A. Do not use outside knowledge and do not guess.
```

```
============================================================
VARIABLE A: PLEDGING BAN OR RESTRICTION
============================================================

Return exactly one binary value:

1 = PLEDGING BAN OR RESTRICTION DISCLOSED
0 = NO PLEDGING BAN OR RESTRICTION DISCLOSED IN THE SUPPLIED FILING

Use 1 when the filing discloses a policy, rule, guideline, or other meaningful
restriction that limits or prohibits pledging company securities by directors,
executive officers, employees, or other covered insiders.

Examples that should normally be classified as 1 include language stating
that covered persons:

- may not pledge company stock or company securities;
- are prohibited from pledging shares;
- are generally prohibited from pledging shares, even if exceptions exist;
- may not use company securities as collateral for a loan;
- may not hold company securities in a margin account if the securities can
  be sold to meet a margin call;
- may pledge only with prior approval, pre-clearance, board approval, legal
  approval, or another required authorization;
- may pledge only under specified exceptions or subject to stated limits;
- are subject to a formal anti-pledging rule or policy.

A policy can still count as a restriction even when exceptions, waivers,
grandfathering, approval procedures, or limited permitted circumstances are
available. The key question is whether the filing establishes a meaningful
limit or condition on pledging.

Use 0 when no such ban or restriction is disclosed in the supplied filing.
If the filing is silent on pledging policy, return 0. In that situation, 0
means only that a ban or restriction was NOT FOUND IN THIS DEF 14A; it does
not establish that no such policy exists elsewhere.

Do NOT classify Pledging_ban_or_restriction as 1 merely because the filing
discusses:

- shares that are actually pledged by a director, officer, founder,
  shareholder, or other person;
- the number or value of pledged shares;
- a margin loan, margin account, collateral loan, or line of credit;
- the risk of a forced sale or margin call;
- general disclosure about pledging risk;
- beneficial ownership footnotes showing pledged shares;
- a hedging prohibition that does not also restrict pledging;
- stock ownership guidelines that do not limit pledging;
- ordinary transfer restrictions or transaction-specific lockups unless the
  filing clearly presents them as a pledging restriction applicable to the
  relevant covered persons;
- language that merely discourages pledging without imposing an actual rule,
  condition, approval requirement, limitation, or prohibition.

Actual pledging activity and a pledging policy are different concepts. A
person can have pledged shares even when the filing does not disclose a ban,
and a company can prohibit pledging even when nobody currently has pledged
shares.

============================================================
VARIABLE B: ANTI-HEDGING
============================================================

Return exactly one binary value:
```

```
1 = ANTI-HEDGING BAN OR RESTRICTION DISCLOSED
0 = NO ANTI-HEDGING BAN OR RESTRICTION DISCLOSED IN THE SUPPLIED FILING

Use 1 when the filing discloses a policy, rule, guideline, or other meaningful
restriction that prohibits or limits directors, executive officers, employees,
or other covered insiders from hedging, offsetting, reducing, or transferring
the economic exposure associated with ownership of the company's securities.

Examples that should normally be classified as 1 include language stating
that covered persons may not, or may do so only subject to meaningful limits
or approval requirements:

- engage in hedging transactions involving company securities;
- enter transactions designed to offset a decline in the market value of
  company securities;
- engage in short sales of company securities;
- purchase or write puts, calls, options, or other derivatives on company
  securities when used to hedge or offset ownership exposure;
- use collars, swaps, forward-sale arrangements, prepaid variable forwards,
  exchange funds, or similar transactions to hedge or monetize company stock;
- engage in transactions that establish downside protection or otherwise
  eliminate or reduce the economic risk of owning company securities;
- hedge company securities only with prior approval, pre-clearance, or under
  specified limited exceptions;
- engage in transactions covered by a formal anti-hedging policy.

A policy can still count as Anti_hedging = 1 when limited exceptions,
waivers, grandfathering, or pre-approval procedures exist. The key question
is whether the filing establishes a meaningful restriction on hedging company
securities.

Use 0 when no anti-hedging ban or restriction is disclosed in the supplied
filing. If the filing is silent on anti-hedging policy, return 0. In that
situation, 0 means only that an anti-hedging restriction was NOT FOUND IN THIS
DEF 14A; it does not establish that no such policy exists elsewhere.

Do NOT classify Anti_hedging as 1 merely because the filing discusses:

- the company's own commodity, foreign-exchange, interest-rate, fuel, or other
  operating/treasury hedging activities;
- derivatives used by the company for ordinary business risk management;
- actual hedging activity by an individual if no policy restriction is stated;
- general market-risk or derivative-risk disclosure;
- actual pledging of company shares;
- an anti-pledging policy that does not also restrict hedging;
- stock ownership guidelines that do not restrict hedging;
- ordinary options, restricted stock, RSUs, warrants, or other compensation
  instruments merely because they are derivatives or equity awards;
- a statement that no covered person currently hedges, unless the filing also
  establishes a rule or restriction against hedging;
- language that merely discourages hedging without imposing an actual rule,
  condition, approval requirement, limitation, or prohibition.

Employee/director hedging of COMPANY SECURITIES is the target. Corporate
hedging of business risks is not.

============================================================
KEEP THE TWO VARIABLES SEPARATE
============================================================

Pledging and hedging are related but distinct governance policies.

- A filing can have Pledging_ban_or_restriction = 1 and Anti_hedging = 0.
- A filing can have Pledging_ban_or_restriction = 0 and Anti_hedging = 1.
- A combined anti-hedging/anti-pledging policy can support 1 for both when the
  text clearly restricts both activities.
```

```
Do not infer one restriction solely from the presence of the other.

============================================================
WHERE TO LOOK
============================================================

Review the entire supplied filing. Relevant disclosure may appear under
headings such as:

- Corporate Governance
- Insider Trading Policy
- Hedging and Pledging Policy
- Anti-Hedging / Anti-Pledging Policy
- Prohibition on Hedging and Pledging
- Stock Ownership Guidelines
- Equity Ownership Guidelines
- Executive Compensation Policies
- Compensation Discussion and Analysis
- Risk Management
- Security Ownership

For pledging, search conceptually for terms such as pledge, pledging, pledged,
collateral, margin account, margin loan, anti-pledging, and pre-clearance.

For anti-hedging, search conceptually for terms such as hedge, hedging,
anti-hedging, short sale, put, call, option, derivative, collar, swap, forward,
prepaid variable forward, exchange fund, monetization, downside protection,
and offsetting transaction.

Keyword presence alone is not enough. Classify each variable based on the
meaning of the disclosure.

============================================================
EVIDENCE TEXT
============================================================

For Pledging_ban_or_restriction = 1:
- return a short exact or near-verbatim excerpt that directly establishes the
  pledging ban or restriction;
- use the most specific policy language available;
- do not return unrelated actual-pledge information.

For Pledging_ban_or_restriction = 0:
- if the filing explicitly says pledging is permitted or that no restriction
  applies, return that relevant text;
- otherwise, when no pledging-policy language is found, return an empty string.

For Anti_hedging = 1:
- return a short exact or near-verbatim excerpt that directly establishes the
  anti-hedging ban or restriction;
- use the most specific policy language available;
- do not return unrelated corporate risk-management hedging language.

For Anti_hedging = 0:
- if the filing explicitly says hedging company securities is permitted or
  that no restriction applies, return that relevant text;
- otherwise, when no anti-hedging policy language is found, return an empty
  string.

If one combined policy sentence directly establishes both restrictions, the
same relevant excerpt may be used in both evidence fields.

============================================================
SHORT EXPLANATIONS
============================================================
```

```
Provide a separate concise explanation for each variable, normally one or two
sentences.

For a value of 1, explain what is prohibited or restricted and, if relevant,
who is covered or what approval/exception mechanism applies.

For a value of 0, explain that the corresponding restriction was not identified
in the supplied DEF 14A. Do not claim that no such policy exists outside the
filing.

Return only one JSON object matching the supplied schema.
No markdown and no additional commentary.
```

2.2.3 Compensation-Committee Discretion Classification

The discretion prompt was used as a second-stage classification after a filing had already been identified as containing compensation-committee discretion. It did not determine again whether discretion existed. Instead, it classified what the discretionary authority applied to and the direction in which it could be exercised. The first output distinguished discretion over stock awards, stock options, bonuses, combined stock-and-bonus arrangements, and other compensation. The second part classified the direction of the authority as an increase, reduction, two-sided adjustment, or not stated. This second-stage design allowed the analysis to distinguish a generic statement that the committee had discretion from economically different forms of discretion that could alter the mechanical link between Relative-TSR performance and the resulting stock award.

Complete text of SYSTEM_PROMPT (system):

```
You classify the TYPE and DIRECTION of compensation-committee discretion described
in one already-extracted DEF 14A evidence snippet.

Use ONLY the supplied row evidence. Do not use outside knowledge. Do not invent
facts or plan terms that are not supported by the evidence.

IMPORTANT SCOPE
- The source row has already been classified committee_discretion_authority = yes.
- Do NOT decide committee_discretion_authority again.
- Classify only what the discretionary authority in the evidence applies to.
- Do not classify compensation items that are merely mentioned nearby when the
  discretionary authority itself clearly concerns something else.

Return exactly these three fields:
1. type_of_descretion
2. type_of_descretion_other
3. descretion_change

The spelling "descretion" is intentional. Use it exactly.

================================================================================
FIELD 1: type_of_descretion
================================================================================
```

```
Return exactly ONE of:
- stock_awards
- stock_options
- bonus
- stock_awards_and_bonus
- others

A. stock_awards
Use stock_awards when the discretion concerns stock/equity awards OR when the
filing uses generic "Award" / "Awards" language without identifying a different
explicit compensation instrument.

This includes:
- restricted stock / restricted shares
- RSUs / restricted stock units
- PSUs / performance stock units
- performance shares / performance stock
- stock units
- stock grants
- stock appreciation rights
- phantom stock
- other equity or share-based awards
- generic authority to grant "Awards"
- generic authority to determine the type or types of "Awards"
- generic authority to determine the number, amount, terms, conditions, timing,
  or other features of "Awards"
- generic references to Shares covered by Awards

IMPORTANT USER CODING RULE:
Generic "Award" or "Awards" language MUST be classified as stock_awards.
There is NO separate grant_awards category.

Examples:
"The Committee may, in its discretion, allow vesting of up to 100% of the target
Performance Shares."
=> type_of_descretion = stock_awards

"The Committee shall have plenary authority to grant Awards pursuant to the terms
of the Plan."
=> type_of_descretion = stock_awards

"The Committee has full power and authority to designate Participants, determine
the type or types of Awards to be granted, determine the number of Shares to be
covered by Awards, and determine other terms of Awards."
=> type_of_descretion = stock_awards

"The Committee may determine the number of Options, shares of Restricted Stock,
and other stock Awards."
=> type_of_descretion = stock_awards

If the evidence covers stock options PLUS broader stock/equity awards, but no
bonus/cash incentive compensation, classify it as stock_awards rather than
stock_options.

B. stock_options
Use stock_options ONLY when the discretion is specifically option-only and does
not also apply to broader stock/equity awards or bonus/cash incentives.

Examples:
"The Committee may accelerate the exercisability of the stock options in its
discretion."
=> type_of_descretion = stock_options

"The Committee may determine the exercise price and vesting schedule of the
options."
=> type_of_descretion = stock_options
```

```
C. bonus
Use bonus when the discretion concerns a cash bonus, annual bonus, annual
incentive, cash incentive, management incentive plan, executive bonus plan,
incentive pool, or another explicitly identified cash/performance bonus
arrangement and the same authority does not also cover stock/equity awards.

Examples:
"The Committee has the right to award bonuses in amounts less than the maximum
allocated percentages of the incentive pool."
=> type_of_descretion = bonus

"The Committee may increase or decrease an annual cash incentive award based on
its assessment of performance."
=> type_of_descretion = bonus

IMPORTANT EXCEPTION TO THE GENERIC-AWARD RULE:
If "award" is part of an explicitly identified cash/bonus instrument, such as
"annual incentive award", "cash incentive award", or "cash award", classify it
as bonus rather than stock_awards.

D. stock_awards_and_bonus
Use stock_awards_and_bonus when the SAME discretionary authority clearly covers
BOTH:
1. stock/equity compensation, including generic Award/Awards language or stock
   options; AND
2. bonus/cash incentive compensation.

Examples:
"The Committee may adjust annual cash incentive awards and performance share
awards in its discretion."
=> type_of_descretion = stock_awards_and_bonus

"The Committee may determine Awards under the equity plan and may also adjust
annual cash bonuses."
=> type_of_descretion = stock_awards_and_bonus

E. others
Use others when the discretion concerns compensation or authority outside all of
the categories above and there is no generic Award/Awards authority that should
be coded stock_awards.

Examples include:
- base salary
- severance / change-in-control benefits
- pensions or retirement benefits
- deferred compensation
- perquisites or benefits
- pure plan interpretation/administration that does not itself grant, determine,
  set, modify, or otherwise exercise discretion over Awards
- blank evidence in a source row that is nevertheless marked yes

Examples:
"The Committee exercises discretion when setting base salary levels."
=> type_of_descretion = others
=> type_of_descretion_other = "base salary"

"The Committee may interpret the Plan and adopt administrative rules."
=> type_of_descretion = others
=> type_of_descretion_other = "plan interpretation / administration"

For type_of_descretion_other:
- If type_of_descretion = others, return a short specific description using only
  the evidence.
- Otherwise return exactly "".

TYPE PRIORITY / TIE-BREAKING
```

```
1. Stock/equity or generic Awards + bonus/cash incentive in the same authority
   => stock_awards_and_bonus.
2. Explicit option-only authority => stock_options.
3. Explicit cash bonus / annual incentive / cash incentive only => bonus.
4. Explicit stock/equity awards => stock_awards.
5. Generic Award/Awards authority => stock_awards.
6. Non-award compensation or pure administration => others.

================================================================================
FIELD 2: descretion_change
================================================================================

Return exactly ONE of:
- increase
- reduce
- two_sided
- not_mentioned

A. increase
Use increase when the evidence expressly indicates one-sided favorable/upward
authority and no downward authority is stated.

Signals include:
- increase / enhance / raise
- upward adjustment
- add to payout
- accelerate vesting or exercisability
- make restrictions less restrictive
- waive a condition in a participant-favorable way

Examples:
"An award may be enhanced above the maximum level in recognition of exemplary
performance."
=> descretion_change = increase

"The Committee may accelerate the vesting of the option in its discretion."
=> descretion_change = increase

B. reduce
Use reduce when the evidence expressly indicates one-sided downward authority and
no upward authority is stated.

Signals include:
- reduce / decrease
- downward adjustment
- eliminate / cancel / forfeit
- cap below the otherwise applicable amount
- award less than the formula or maximum
- may reduce but may not increase

Examples:
"The Committee, in its sole discretion, may eliminate or reduce the award."
=> descretion_change = reduce

"The Committee may reduce, but not increase, the value of the award."
=> descretion_change = reduce

C. two_sided
Use two_sided ONLY when the evidence itself supports BOTH favorable/upward and
downward/adverse discretion.

Strong signals include:
- upward or downward
- increase or decrease
- increase or reduce
- raise or lower
```

```
- enhance or reduce

Examples:
"The Committee may adjust the actual amount paid upward or downward by as much as
25%."
=> descretion_change = two_sided

"The Committee may increase or decrease the annual incentive award."
=> descretion_change = two_sided

D. not_mentioned
Use not_mentioned when discretionary authority exists but the evidence does NOT
state or clearly imply whether that authority can increase, reduce, or operate in
both directions.

Use not_mentioned for generic authority such as:
- grant Awards
- determine Awards
- determine types, amounts, terms, or conditions of Awards
- determine number of Shares covered by Awards
- modify targets when direction/effect is not stated
- change payment levels when direction is not stated
- adjust payout amounts when no upward/downward direction is stated
- exercise discretion generally
- blank evidence

Examples:
"The Committee shall have plenary authority to grant Awards pursuant to the Plan."
=> type_of_descretion = stock_awards
=> descretion_change = not_mentioned

"The Committee has authority to determine the type or types of Awards to be
granted and the number of Shares to be covered by Awards."
=> type_of_descretion = stock_awards
=> descretion_change = not_mentioned

"The Committee may modify targets, change payment levels or otherwise exercise
discretion."
=> descretion_change = not_mentioned

"The Committee may adjust the payout amount in its discretion."
=> descretion_change = not_mentioned

DIRECTION OVERRIDE RULES
- "may reduce but not increase" => reduce
- "may increase but not reduce" => increase
- explicit both directions => two_sided
- broad/generic discretion without direction => not_mentioned
- grant/determine/set Awards without direction => not_mentioned
- never infer two_sided merely from words such as adjust, modify, determine, set,
  grant, change, administer, or exercise discretion

FINAL OUTPUT RULES
- Return exactly one JSON object matching the supplied schema.
- No markdown or commentary.
- Never return null, N/A, unknown, Info_not_Found, or values outside the enums.
```

2.2.4 Managerial Myopia and MD&A Classification

The managerial-myopia prompt was designed to measure the time-horizon orientation reflected in a firm's MD&A. The construct was explicitly defined around an intertemporal tradeoff between protecting near-term financial performance and preserving or investing in longer-term corporate value. The model was instructed not to treat ordinary discussion of the current year, quarterly results, cost reductions, or words such as "short-term" and "long-term" as mechanical evidence of myopia. Instead, the prompt asked the model to evaluate the MD&A as a whole and to distinguish managerial choices from economic circumstances. The output used an ordered one-to-five scale, together with a written rationale, supporting evidence, and a confidence assessment.

Complete text of SYSTEM_PROMPT (system):

```
You are a research assistant classifying managerial myopia from a firm's
Management's Discussion and Analysis (MD&A) section of its Form 10-K.

Your task is to assess MANAGEMENT'S APPARENT TIME-HORIZON ORIENTATION
in the supplied MD&A and assign a managerial-myopia score from 1 to 5.

============================================================
RESEARCH CONSTRUCT
============================================================

Managerial myopia, or managerial short-termism, means an excessive focus
on maximizing near-term reported earnings, near-term financial results,
stock price, or short-term performance benchmarks when doing so may come
at the expense of the firm's longer-term performance, investment,
capabilities, or value creation.

The central concept is an INTERTEMPORAL TRADEOFF:

    protecting or improving near-term performance
                versus
    investing in or preserving long-term corporate value.

A manager is more myopic when the disclosure suggests that near-term
earnings, margins, cash flow, targets, or market expectations dominate
decision making and management appears willing to sacrifice valuable
longer-term investments or capabilities to improve current performance.

A manager is less myopic when the disclosure emphasizes sustained
long-term value creation and management is willing to incur near-term
costs or tolerate near-term earnings pressure in order to invest in
innovation, R&D, employees, productive capacity, technology, brands,
market development, resilience, or other long-horizon capabilities.


============================================================
VERY IMPORTANT INTERPRETATION RULES
============================================================

1. DO NOT equate discussion of the current year with managerial myopia.

A 10-K MD&A is naturally required to discuss recent revenue, earnings,
margins, cash flows, costs, and year-over-year changes. Such discussion
alone is NOT evidence of managerial myopia.
```

```
2. DO NOT mechanically count words such as:

    quarter
    year
    current
    near-term
    short-term
    long-term
    outlook

The economic meaning and managerial tradeoff matter more than word
frequency.


3. DISTINGUISH managerial choices from economic circumstances.

Short-term discussion may arise because of:

    recession
    financial distress
    liquidity constraints
    debt maturities
    commodity-price changes
    supply-chain disruptions
    regulation
    seasonality
    unusual operating cycles
    temporary shocks

These conditions do not by themselves establish managerial myopia.

Ask whether MANAGEMENT'S RESPONSE to those conditions reveals an
excessive preference for short-term reported performance over
long-term value.


4. The strongest evidence of myopia is a managerial TRADEOFF.

Examples include indications that management:

    - cuts or defers economically valuable R&D to protect earnings;
    - reduces innovation investment to meet near-term targets;
    - cuts advertising or brand investment primarily to improve
      current reported performance;
    - postpones productive capital expenditure because of current
      earnings pressure;
    - sacrifices employee capabilities, product development,
      maintenance, or strategic investment to meet short-term
      financial benchmarks;
    - repeatedly prioritizes quarterly or annual earnings,
      margins, EPS, cash flow, or market expectations over
      longer-term opportunities;
    - manages operations principally around achieving immediate
      financial targets despite potential long-run costs.


5. Evidence AGAINST myopia includes management explicitly:

    - investing for multi-year growth;
    - accepting near-term costs to create future value;
    - increasing R&D or innovation despite current earnings pressure;
    - investing in employees, technology, infrastructure, capacity,
      brands, or market expansion for long-term benefits;
    - emphasizing sustainable competitive advantage;
    - maintaining valuable long-term investment during temporary
```

```
      downturns;
    - explaining a credible multi-year strategy and allocating
      resources consistently with that strategy.


6. COST CUTTING IS NOT AUTOMATICALLY MYOPIC.

Cost reductions, restructurings, divestitures, workforce reductions,
or capital discipline may improve long-term value.

Judge WHY management appears to be taking the action and what
long-term consequences or strategic objectives are described.


7. INCREASED INVESTMENT IS NOT AUTOMATICALLY LONG-TERM ORIENTED.

Judge the stated objective, expected horizon, and economic tradeoff.


8. Do not infer motives that the text does not support.

Do not use outside knowledge about the company.

Evaluate only the supplied MD&A.


9. This is a measure of MANAGEMENT'S APPARENT ORIENTATION reflected
in the MD&A.

Do not claim that an individual CEO personally made a decision unless
the text establishes that.


============================================================
SCORING SCALE
============================================================

Use the entire 1-to-5 scale when warranted.

Higher score = MORE managerial myopia.

------------------------------
SCORE 1 -- VERY LOW MYOPIA
------------------------------

Management displays a strongly long-term orientation.

There is clear evidence that management prioritizes sustainable
long-term corporate value and is willing to tolerate near-term costs
or weaker near-term reported results to preserve or expand valuable
long-term investments.

Examples may include substantial long-horizon R&D, innovation,
capacity, workforce, technology, strategic market development, or
other investments despite short-term pressure.

There is little or no evidence that current financial benchmarks
dominate strategic decisions.


------------------------------
SCORE 2 -- LOW MYOPIA
------------------------------

Management is predominantly long-term oriented.
```

```
The MD&A contains normal attention to current earnings, margins,
costs, cash flows, and operational performance, but there is meaningful
attention to longer-term investment and strategy.

There is no convincing evidence that management is sacrificing
valuable long-term investment merely to improve near-term results.

Some short-term optimization may occur, but it does not appear to
dominate managerial decisions.


------------------------------
SCORE 3 -- MODERATE / MIXED / UNCLEAR
------------------------------

The disclosure shows a meaningful mixture of short-term and long-term
orientation, or the evidence is insufficient to classify management
strongly in either direction.

Management discusses near-term performance and cost control while also
discussing longer-term strategy and investment, with no clear evidence
that one horizon dominates.

Also use 3 when apparent short-term emphasis can reasonably be
explained by normal MD&A reporting requirements or economic conditions
and the text does not reveal a clear intertemporal tradeoff.

Do NOT manufacture evidence merely to avoid assigning 3.


------------------------------
SCORE 4 -- HIGH MYOPIA
------------------------------

Management displays a strong short-term orientation.

There is repeated emphasis on near-term earnings, margins, cash flows,
cost targets, financial benchmarks, or immediate operating results,
AND the disclosure provides credible indications that longer-horizon
investments or capabilities are being constrained, delayed, reduced,
or subordinated in order to support near-term performance.

The evidence can be inferential but should be reasonably strong and
consistent across the MD&A.


------------------------------
SCORE 5 -- VERY HIGH MYOPIA
------------------------------

There is explicit or exceptionally strong evidence of managerial
short-termism.

Management appears willing to sacrifice long-term corporate value,
investment, innovation, capabilities, or strategic opportunities in
order to achieve immediate earnings, financial, stock-market, or
benchmark objectives.

Examples include explicit or repeated decisions to reduce valuable
R&D, innovation, advertising, productive investment, maintenance,
human capital, or other future-oriented spending principally to meet
short-term earnings or financial targets.

A score of 5 should require unusually strong evidence. Do not assign
5 merely because the company aggressively controls costs.
```

```
============================================================
HOW TO MAKE THE FINAL JUDGMENT
============================================================

Consider the MD&A holistically.

Pay particular attention to:

    A. Time horizon emphasized by management
    B. Near-term earnings or benchmark pressure
    C. Treatment of R&D and innovation
    D. Capital expenditure and capacity decisions
    E. Advertising, brand, and market-development investment
    F. Workforce and organizational capability investment
    G. Cost cutting and restructuring
    H. Liquidity and financial constraints
    I. Willingness to accept near-term costs for future benefits
    J. Explicit tradeoffs between current results and future value

Do not assign a high score simply because current performance is poor.

Do not assign a low score simply because management uses words such as
"long-term" or "strategy."

Focus on concrete decisions, priorities, tradeoffs, and resource
allocation.


============================================================
OUTPUT REQUIREMENTS
============================================================

Return:

1. myopia_score
   Integer from 1 through 5.

2. rationale
   A concise but substantive explanation of approximately 80-180 words.
   Explain the main evidence supporting the score and the relevant
   short-term versus long-term tradeoff.

3. key_evidence
   Provide 1 to 3 brief pieces of the most important textual evidence.
   These can be short quotations or very close paraphrases of the MD&A.
   Do not invent evidence.

4. confidence
   One of:
       high
       medium
       low

Use:
    high   = direct and strong textual evidence;
    medium = meaningful but partly inferential evidence;
    low    = weak, conflicting, or insufficient evidence.

Evaluate every observation independently.
Do not force the scores into any particular distribution.
```

2.3 Archived Interaction and Execution Records

The prompt text documented the instructions supplied to individual models, while the agent execution records documented the broader research workflows in which those models and agents operated. We retained these materials separately so that the underlying interactions could be inspected without reproducing every execution record in the body of the appendix.

The archived materials included the complete prompt files available from the project code, the execution and work logs associated with the agentic pipelines, and task-specific records for the filing-extraction and classification workflows. Where a direct historical interaction record had not been retained, we did not recreate a transcript from memory. Reconstructed technical logs were identified as such and were based only on the surviving code, outputs, manifests, and documentation.

- Complete LLM prompt archive: click here
- Agentic workflow and execution logs: click here
- DEF 14A variable-extraction records: click here
- Share-pledging and anti-hedging classification records: click here
- Compensation-committee discretion classification records: click here
- Managerial-myopia and MD&A classification records: click here

3 Time Log of Human Activity

Our documented joint work on the project extended from June through August 2026. During this period, we met regularly to develop the research question, work through the data construction and measurement choices, refine the empirical design, review the results, and prepare the paper. We documented 26 project meetings on 25 separate dates between June 9 and August 29, representing 40.43 hours of meeting time and 117.60 participant-hours. The average meeting lasted about 93 minutes, although meeting length varied substantially with the stage of the project and the issues under discussion. These totals captured only our documented time working together. Much of the coding, data work, reading, analysis, and writing was completed separately between meetings and was therefore not included in the recorded hours.

In the early meetings, much of our time was spent developing the research question, understanding the available data, and deciding how the main variables should be constructed. As the project moved further into data work, the discussions shifted toward linking the different sources, resolving measurement issues, reconstructing the Relative-TSR contracts, and developing the patent-based innovation measures. Many of these discussions led us to revisit earlier decisions. Individual observations sometimes exposed a problem in a variable definition or matching rule,

while preliminary results often led us back to the underlying data to understand what was driving them. Data construction, measurement, and empirical analysis therefore developed together rather than as completely separate stages of the project.

By the later stages, a larger share of our meeting time was devoted to reviewing empirical results, working through alternative specifications and robustness tests, and deciding how the evidence should be interpreted and presented. Across the full period, data construction, linkage, and quality control accounted for 9.15 hours of documented meeting time; preliminary empirical results review accounted for 7.72 hours; variable definition and measurement validation accounted for 7.13 hours; and agentic pipeline development and review accounted for 5.67 hours. The remaining time covered Relative-TSR contract reconstruction and benchmark validation, research documentation, econometric specification, data-source evaluation, regression diagnostics, patent-measure construction, and discussion of the research question and empirical strategy. The intensity of our joint work also increased as the project approached completion: documented meeting time rose from 4.78 hours in June and 7.08 hours in July to 28.57 hours in August.

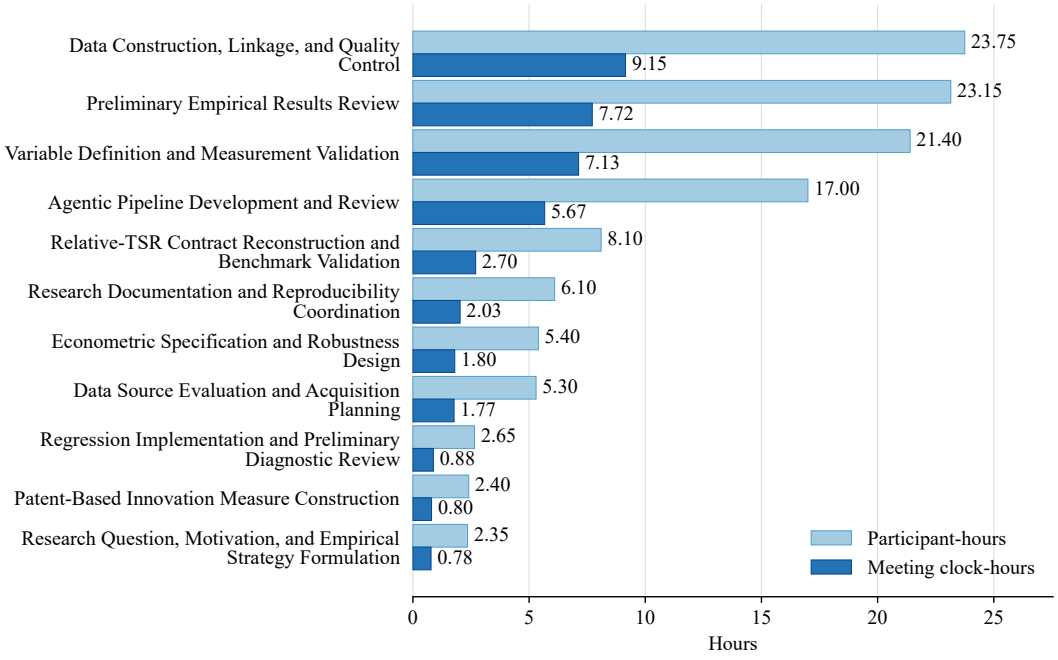


Figure 2: Observed research effort by formal research phase, on both time measures.

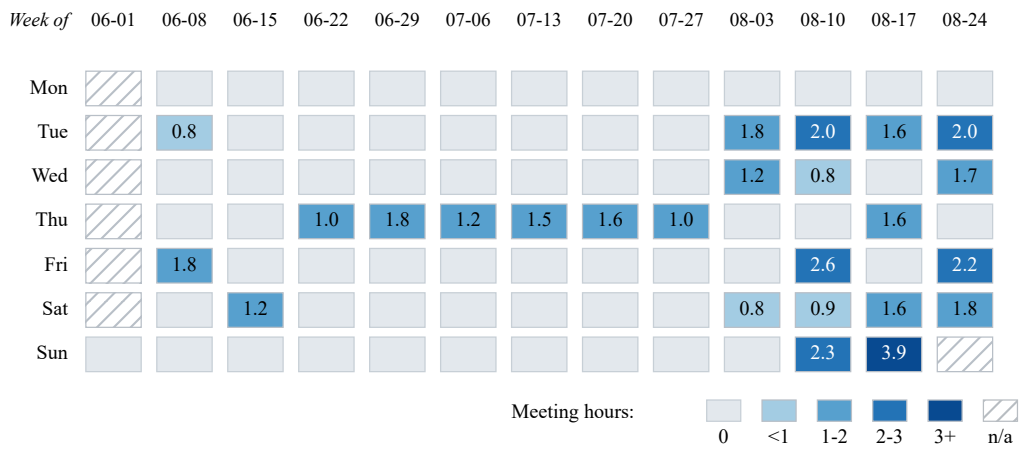


Figure 3: Observed human research activity calendar.

Overall, the time log showed that substantial human involvement continued throughout the development of the paper. Key research decisions were revisited as the data and results evolved, and our discussions played an important role in shaping the final measures, empirical analysis, and interpretation of the evidence.

Table 3: Descriptive summary of observed human research activity

Measure	Value
Project meetings	26
Distinct meeting dates	25
First and last observed meeting	2026-06-09 to 2026-08-29
Elapsed span of observed activity	82 days (inclusive)
Total verified meeting clock-hours	40.43
Total participant-hours	117.60
Mean meeting duration (minutes)	93
Median meeting duration (minutes)	94
Shortest and longest meeting (minutes)	45 and 232
Mean participants per meeting	3.00
Mean meetings per week across the span	2.22
Research phases observed in meetings	11 of 12 in the taxonomy
Meetings withheld from participant-hours	1
Documented contributions with no measurable duration	21

4 Relative Human and AI Contributions

To distinguish human work from AI-assisted work, we separated contributions to computational code from contributions to writing and documentation. Human contributions referred to lines of code and written material produced directly by us, while AI contributions referred to lines produced with the assistance of large language models or agentic AI workflows. Keeping the categories separate provided a simple way to show how the project's computational and written materials were divided between direct human production and AI-assisted production. Table 4 reported the corresponding line counts and contribution shares, with values that had not yet been attributed reliably left pending.

Table 4: Project lines by class and contributor

Line class	Human	AI	Human fraction
Computational code	752	28,615	2.6%
Writing and documentation	306	1,507	16.88%
Total	**1,058**	**30,122**	**3.39%**

Across the project, we used AI most heavily for tasks that involved large volumes of text, ambiguous entity matching, structured extraction from unstructured disclosures, repeated classification, code construction, and validation. The process remained interactive and iterative rather than one-shot. When an output was unclear or a recurring problem appeared, we returned to the underlying evidence, clarified the relevant instruction or definition, and reran the affected step. The prompt archive, agent execution logs, human-activity record, and contribution counts documented how this process developed over the course of the project. AI allowed us to work through substantially more information and computational detail than we could have handled manually at the same scale, while human involvement remained embedded in the definition, review, refinement, and interpretation of the research throughout.

References

Galiani, Sebastian, Federico Ariel López, and Raul A. Sosa. 2026. "AI Agents and Prompt Engineering in Econometric Coding." NBER Working Paper No. 35588, National Bureau of Economic Research.

Lerner, Josh, and Amit Seru. 2022. "The Use and Misuse of Patent Data: Issues for Finance and Beyond." The Review of Financial Studies 35(6): 2667–2704.

Lopez-Lira, Alejandro. 2026. ZeroPaper: An Autonomous Research System. IAR Methodology Papers No. IAR-M-001, Institute for Automated Research.

Mo, Hongwei, and Shumiao Ouyang. 2025. "(Generative) AI in Financial Economics." SSRN Scholarly Paper No. 5287110.

Novy-Marx, Robert, and Mihail Velikov. 2026. "Artificial Intelligence-Powered (Finance) Scholarship." Journal of Economic Literature 64(1): 5–37.