Can Language Models Identify Corporate Peers from Annual Filings? An NLP Evaluation of Insider Trading Enforcement

Anonymous ACL submission

Abstract

When regulators apply a legal standard to real people, a basic fairness requirement is that individuals can determine in advance what conduct is prohibited. This paper examines whether that requirement is met for a recently litigated theory of insider trading, in which liability may attach to trades in any firm deemed "economically linked" to the source of inside information. We ask a concrete NLP question: can the set of economically linked firms be identified *before the fact* from publicly mandated disclosures, the same information available to any market participant? Using a two-stage LLM pipeline applied to Item 7 (Management's Discussion and Analysis) sections of SEC 10-K filings, we score semantic similarity across 30 M&A events spanning five industries and test whether similarity predicts announcement-day peer stock reactions. Our pipeline succeeds on the case that established the enforcement theory, validating the methodology. Across the full dataset, however, the relationship is indistinguishable from chance: 14 of 30 events support the hypothesis, 12 contradict it, and 4 are ambiguous (binomial $p \approx 0.86$ against $H_0 = 0.5$). We conclude that NLP cannot reliably operationalize this legal standard from public filings—suggesting the standard may not provide the fair notice that due process requires, and that targeted NLP-based watchlists could be a more equitable alternative to mass market surveillance.

1 Introduction

On August 18, 2016, Matthew Panuwat, a business-development executive at Medivation, Inc., received an email from his CEO indicating that Pfizer would soon likely acquire the company. Seven minutes later, he purchased call options—not in Medivation, but in Incyte Corporation, a separate publicly traded biopharmaceutical company with no direct business relationship to Medivation. When the acquisition was announced the next day, Incyte's stock rose 7.7%, and Panuwat sold his options for approximately $107,000 in profit (United States District Court, N.D. Cal., 2023).

The SEC charged Panuwat with insider trading under a novel legal theory termed "shadow trading." Under this theory, Material Nonpublic Information (MNPI) concerning Firm A is deemed material to any "economically linked" peer firm, meaning that trading in Firm B's securities constitutes fraud (United States District Court, N.D. Cal., 2023).

The enforcement problem. Classical insider trading enforcement begins with a *known flagged trade*: regulators observe unusual volume in a specific stock, trace it to a specific person, and bring a case. Shadow trading breaks this pattern. If someone learns that Company A will be acquired and instead buys shares in Company B—a competitor they believe will also benefit—the connection between that trade and the inside information is only visible *after* the fact, once Company A's deal is announced and Company B's stock moves. Regulators cannot know in advance which security to watch. They must instead monitor *all* trading across the entire market and search backward after each corporate event to find correlated trades. The Consolidated Audit Trail (CAT)—a government database recording every order and execution across all U.S. exchanges for more than 100 million investors (New Civil Liberties Alliance, 2024a)—is the infrastructure that makes this possible. Critics have characterized CAT as a "dangerous dog" (Peirce, 2019) and challenged its constitutionality (New Civil Liberties Alliance, 2024b; Supreme Court of the United States, 2018); this paper's results bear directly on whether that surveillance is empirically necessary.

The NLP question. If economic linkage can be identified in advance from public disclosures, targeted watchlists could replace mass surveillance.

Because regulators require firms to disclose their risk-reward profiles in annual reports, Item 7 (Management's Discussion and Analysis) of the Form 10-K provides an ideal corpus for testing this hypothesis: if two firms genuinely occupy the same strategic niche, their narrative disclosures should exhibit high semantic similarity. Furthermore, if this enforcement hypothesis is empirically grounded, those same firms should experience positive abnormal stock returns when a peer acquisition is announced. We operationalize this test using a two-stage LLM pipeline: a peer discovery stage that extracts candidate companies from the SEC's EDGAR database (the central public repository for US corporate filings), followed by a document-grounded similarity ranking stage that assigns semantic similarity scores. We then compute announcement-day abnormal returns for each peer firm and assess whether textual similarity predicts peer stock reactions across 30 acquisition events in five industries.

Contributions.

1. A reproducible two-stage LLM pipeline for long-context multi-document financial similarity analysis, including principled handling of Foreign Private Issuers (FPIs).

2. A dataset of 30 M&A events (2013–2023) with LLM-derived similarity scores and manually computed abnormal returns, spanning biopharmaceuticals, consumer staples, technology, finance, and automotive industries.

3. Empirical evidence that textual similarity in SEC filings is an unreliable predictor of peer stock reactions ($14/30$ events supportive; binomial $p \approx 0.86$ against $H_0 = 0.5$), with striking cross-industry variation: the Technology sector yields zero supporting cases across five acquisitions, while Automotives yields zero contradictions, suggesting that document-recoverable economic linkage is industry-contingent rather than universal, directly challenging the predictability assumed by shadow trading enforcement.

Evaluating whether NLP can operationalize a legal standard is a concrete form of public-interest research: it asks not just whether a model can perform a task, but whether that task is tractable enough to justify imposing legal liability on those who fail at it. Our results suggest that text-based enforcement tools, deployed in a sector-stratified way, could make insider trading liability more predictable and equitable than the current mass-surveillance approach.

Table 1: Timeline of *SEC v. Panuwat*.

Date	Event
Aug. 2016	Panuwat trades Incyte options, 7 minutes after CEO email
Aug. 2021	SEC files civil complaint
Jan. 2022	Motion to dismiss denied
Nov. 2023	Motion for summary judgment denied
Apr. 2024	Jury finds Panuwat liable after 8-day trial
Sep. 2024	Civil penalty of $321,197 imposed
Nov. 2024	Panuwat appeals (9th Cir. No. 24-6882)
May 2025	U.S. Chamber of Commerce files amicus brief
May 2026	Case **pending** at Ninth Circuit

2 Background

2.1 Shadow Trading and *SEC v. Panuwat*

Under the classical theory of insider trading, liability attaches to a relationship between the trader and the *issuer* of the traded security. *United States v. O'Hagan* (Supreme Court of the United States, 1997) extended this to misappropriation: a person who trades on MNPI obtained from a source owed a duty of trust and confidence commits securities fraud even if they have no relationship with the issuer. *SEC v. Panuwat* (United States District Court, N.D. Cal., 2023) pushes this further still: the duty-source need not be connected to the security traded at all. Because Panuwat owed a duty to Medivation, and because the court found that Medivation and Incyte were "economically linked" mid-cap oncology companies, his MNPI about Medivation was deemed material to Incyte's securities. Table 1 shows the case timeline.

The key doctrinal problem is that neither firm mentioned the other in its annual report, and neither shared a direct business relationship. "Economic linkage" was established retroactively by the SEC's expert witness, who testified that Incyte's 7.7% announcement-day return was not caused by regular fluctuations. The "mid-cap oncology" category itself is not a recognized industry classification but a narrative label invoked after the fact. The economic stakes are large: Augustin et al. (2019) document positive abnormal options volume in approximately 25% of U.S. takeovers between 1996 and 2012, and Mehta et al. (2021) estimate single-event shadow-trading profits rang-

ing from $139,400 to $678,000. Since *Panuwat*, the only concluded enforcement action has been *Bechtolsheim*, which settled for a $923,740 civil penalty (U.S. Securities and Exchange Commission, 2024); even so, roughly one in five public companies has now formally prohibited shadow trading in its insider-trading policy (Jenkins, 2025).

2.2 Text-Based Firm Similarity in Financial Disclosure

The SEC's "economic linkage" standard presupposes that similarity between firms is readable from their public disclosures. The foundational evidence for this comes from the Hoberg-Phillips Text-based Network Industry Classification (TNIC) (Hoberg and Phillips, 2016), which demonstrates that pairwise firm similarity constructed from mandatory 10-K product descriptions captures competitive dynamics far better than static SIC codes. If TNIC-style analysis can classify industries, it should, in principle, identify the economically linked peers that shadow trading liability requires.

However, TNIC relies on bag-of-words (BoW) similarity over Item 1 (Business Description), which is semantically blind to synonymous terminology: "oncology" and "cancer treatment" would be treated as unrelated tokens. This creates the *administrability gap*: if two firms describe the same strategic niche using different technical vocabularies, BoW similarity fails to detect the linkage that shadow trading liability presupposes. Long-context generative LLMs, applied to the forward-looking narrative of Item 7 (MD&A) rather than the backward-looking Item 1, are the first class of models capable of resolving this gap at the document scale required by full 10-K sections. Whether closing the administrability gap is sufficient to make economic linkage reliably identifiable is the empirical question this paper addresses.

3 Methodology

3.1 Dataset Construction

We identified 30 major public acquisitions across five industries: biopharmaceuticals (10 events), consumer staples (5), technology (5), finance (5), and automotive (5), spanning 2013–2023.

Selection rationale. Biopharmaceuticals receives double the allocation for two reasons: it is the industry of origin for shadow trading doctrine, and it features tight, product-defined competitive peer structures, which is the most favorable setting for the SEC's theory and therefore the most critical test. The four remaining industries (5 events each) form a cross-industry control group, chosen to span structurally different competitive environments: vertically integrated consumer supply chains, platform-based technology ecosystems, regional banking, and automotive supply chains. Within each industry, we selected events by deal value subject to two constraints: (a) the target's Form 10-K must have been filed on EDGAR before the announcement date, and (b) at least five peer firms with active NYSE/NASDAQ listings must be identifiable from the filing alone. The 2013–2023 window predates widespread awareness of shadow trading liability, ensuring that trading behavior was not shaped by knowledge of the enforcement theory itself.

For each acquisition target (Firm A), we use Item 7 (MD&A) and Item 7A as the similarity corpus rather than Item 1 (Business Description). Item 7 is forward-looking, as it describes management's view of competitive risks, market opportunities, and strategic priorities, reflecting the perspective a company insider would actually hold. Prior bag-of-words approaches such as TNIC (Hoberg and Phillips, 2016) use Item 1; our shift to Item 7 is the key methodological extension. For Foreign Private Issuers (FPIs) filing Form 20-F, we substitute Item 5 ("Operating and Financial Review and Prospects") and Item 11 ("Quantitative and Qualitative Disclosures About Market Risk") accordingly.

3.2 Two-Stage LLM Pipeline

Figure 1 illustrates our **Augmented Hoberg-Phillips (HP) framework**, which extends the classical HP BoW approach (Hoberg and Phillips, 2016) by substituting a long-context LLM for cosine similarity over binary term vectors, and shifting from Item 1 (Business Description) to Item 7 (MD&A). We decompose the task into two sequential stages to avoid self-consistency bias: models anchored to prior numerical scores resist revision even when new evidence warrants it (Huang et al., 2024; Harshavardhan, 2026; Khot et al., 2023). Both stages are implemented in Google AI Studio using Gemini 3.1 Pro in isolated per-event sessions; full model configuration details are provided in Section 3.3.

Stage 1 — Peer Discovery. The target firm's Item 7 is uploaded to identify ten peer companies as of the acquisition date, with Google Search

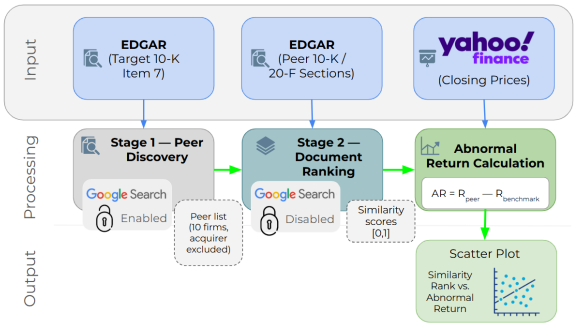

Figure 1: Augmented Hoberg-Phillips Framework

grounding enabled solely to verify historical public trading status. The model provides natural language dimensional justifications (Table 2) without numerical scores to mitigate anchoring bias. The three dimensions are grounded in the SEC's own theory of economic linkage as operationalized at the *Panuwat* trial: the expert witness established linkage through shared therapeutic mechanism (Tech Overlap), comparable commercialization stage (Strategic/Financial), and common market exposure (Risk). Using the SEC's own implicit rubric as our similarity metric ensures that a positive result would be directly probative of the enforcement theory, and that a negative result cannot be dismissed as measuring the wrong dimension.

Table 2: Stage 1 Dimensional Justification Rubric

Dimension	Focus	Assessment Criteria
Tech Overlap	Product	Similarity of platforms or pipelines.
Strategic	Financial	Matching growth stages and revenue models.
Risk	External	Geographic exposure and headwinds.

Stage 2 — Document-Grounded Ranking. Item 7 sections for all identified peers are uploaded as a single multi-document prompt. The model assigns scores in $[0, 1]$ based exclusively on the provided disclosures with Google Search disabled, mirroring the closed-corpus logic of HP (Hoberg and Phillips, 2016). Outputs are parsed into structured `.csv` tables. We weight Technical/Product Overlap most heavily (40%) to reflect the SEC's emphasis on shared therapeutic mechanisms in *Panuwat*; Financial (Strategic Stage) and Risk & Market (Exposure) each receive 30%, capturing the forward-looking narrative dimensions unique to Item 7 relative to Item 1.

Table 3: Stage 2 Similarity Score Rubric

Dimension	Weight	Assessment Criteria
Technical	40%	Core technology and product engines.
Financial	30%	Growth trajectories and margin profiles.
Risk/Market	30%	Geographic focus and specific headwinds.

Why automate? Manual expert evaluation of firm similarity at this scale—30 events, each requiring a rater to read up to 10 peer 10-K sections of 20,000+ tokens apiece—is prohibitively time-consuming and introduces subjective inconsistency across events. Our pipeline applies a fixed, documented rubric uniformly across all 30 events, enabling reproducible cross-industry comparison. Ground truth validation on Pfizer/Medivation (Section 4.1), where the pipeline independently identifies Incyte as the third most similar peer—matching the SEC's own post-hoc identification—provides external validation of the automated scores.

3.3 Model and Configuration

We select Gemini 3.1 Pro (Google DeepMind, 2026) for its 1M-token context window (sufficient to ingest multiple complete 10-K sections without chunking), robust multi-document cross-referencing, and state-of-the-art reasoning performance (77.1% on ARC-AGI-2; 94.3% on GPQA Diamond as of February 2026). The pipeline runs in Google AI Studio with Thinking Level set to High, enabling Chain-of-Thought reasoning over dense SEC disclosures (Wei et al., 2023). Each acquisition event runs in an isolated session to prevent cross-contamination. Top-P is slightly lowered from 0.95 to 0.90 to reduce uncertain token probability mass while preserving domain-specific financial vocabulary (Li et al., 2024); all other parameters remain at default.

3.4 Abnormal Return Calculation

For each peer firm, we compute the announcement-day abnormal return:

$$AR_i = R_i^{(0)} - R_{\text{benchmark}}^{(0)} \quad (1)$$

where $R_i^{(0)} = (P_i^{(0)} - P_i^{(-1)})/P_i^{(-1)}$ is the peer firm's raw return on announcement day (Day 0) relative to the prior trading day (Day -1), and $R_{\text{benchmark}}^{(0)}$ is the same-day return of a sector-level

ETF: XBI (biopharmaceuticals), IYK (consumer staples), XLK (technology), XLF (finance), and IYC (automotive). This directly mirrors the SEC economist's methodology at the *Panuwat* trial. We compute the Spearman rank correlation (ρ) between similarity rank and AR_i per event and generate industry-specific scatter plots (Appendix C).

Mid-cap validation. Across our 30 events, roughly 35% of identified peers fall within the standard \$2B–\$10B mid-cap band, and roughly half fall within the broader S&P 400 range (\$1.79B–\$35.5B). A comparable share holds within the biopharmaceutical subset specifically. This confirms that "mid-cap" is not a discriminative criterion: it captures only a minority of the peer universe and therefore provides no meaningful constraint on which firms an insider would know to trade, directly undermining the SEC's post-hoc invocation of the label in *Panuwat*.

Reproducibility. All prompts, per-event CSV outputs, and scatter plots are provided in the Appendix. Abnormal return calculations are implemented in Python using `yfinance` and `pandas`, with Stooq as a fallback for delisted tickers. Because Gemini 3.1 Pro is a closed-weight model, exact score reproduction is not guaranteed across API versions; we therefore treat LLM outputs as structured qualitative judgments and compute all financial metrics independently.

4 Results

4.1 Ground Truth: Pfizer/Medivation (2016)

We treat the Pfizer/Medivation acquisition as our ground truth case since it is the factual basis of *SEC v. Panuwat*. Figure 2 shows the scatter plot of similarity rank versus abnormal return. Our pipeline correctly identifies Tesaro (TSRO, Rank 1), Exelixis (EXEL, Rank 2), and Incyte (INCY, Rank 3) as the most similar peers; all three show positive abnormal returns (+1.60%, +1.33%, and +5.04%, respectively). Incyte—the actual stock Panuwat traded—is the third most similar firm in our analysis, consistent with the SEC's theory. The trendline is slightly negative, indicating a modest tendency for more similar firms to show higher positive abnormal returns. While the correlation is not strong, Incyte ranks third in our similarity analysis and shows the largest positive AR ($+5.04\%$), which is consistent with the SEC's theory for this specific case.

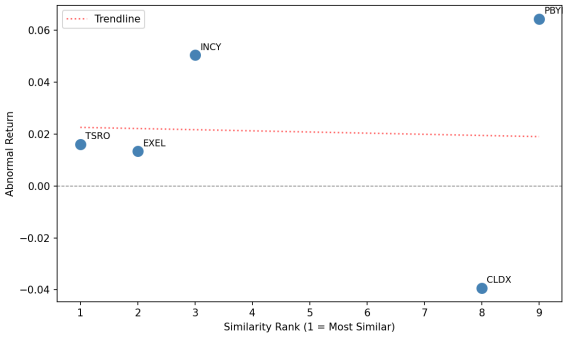

Figure 2: Pfizer/Medivation (2016): negative trendline confirms that more similar peers showed larger positive ARs, consistent with the shadow trading hypothesis.

4.2 Full Results: 30 Acquisition Events

Table 4 summarizes all 30 events. It is important to note that our final interpretation of whether a case supports or contradicts the SEC's theory is not a mechanical byproduct of the OLS trendline alone. Instead, we performed a holistic case-by-case analysis considering the trendline direction, the percentage of peers experiencing positive versus negative abnormal returns, and the absolute scale of those returns (See Appendix B).

A "negative" trendline (↓) indicates that higher similarity scores correlate with higher abnormal returns—the core requirement for the shadow trading theory. Conversely, a "positive" trendline (↑) indicates that higher similarity scores correlate with lower abnormal returns. However, if a trend is negative but the absolute returns for the top-ranked peers are negligible or negative, the case does not "support" a theory of predictable profit, and similarily vice versa. Peer counts per event vary below ten for several reasons: firms identified in Stage 1 were subsequently excluded when their EDGAR filings lacked sufficient Item 7 content, or when historical price data was unavailable due to post-announcement delisting.

4.3 Biopharmaceutical Results

Our analysis begins with the Biopharmaceutical sector, the very crucible of the SEC's shadow trading doctrine. In the Pfizer/Medivation acquisition, the algorithm successfully identified Incyte as a highly similar peer that experienced a positive abnormal return in Figure 2. This confirms that our semantic similarity metric effectively captures the SEC's conceptualization of "economic linkage."

However, when this methodology is expanded across a broader dataset of ten major biopharmaceu-

Table 4: Summary of all 30 acquisition events.
Trend: ↓ = negative (similarity predicts positive AR); ↑ = positive (similarity predicts negative AR).
†Factual basis of *SEC v. Panuwat*; treated as ground truth.

Industry	Acquisition Announcement	Year	Peers	Trend	Top-3 AR Direction	Interpretation
Biopharmaceuticals	Pfizer / Medivation†	2016	5	↓	$+, +, +$	Supports (ground truth)
	AbbVie / ImmunoGen	2023	8	↓	$+, +, +$	Strongly supports
	BMS / Karuna	2023	8	↓	$+, -, \approx 0$	Supports
	BMS / Mirati	2023	7	↓	$+, -, +$	Contradicts
	BMS / MyoKardia	2020	8	↓	$+, \approx 0, -$	Strongly contradicts
	BMS / Turning Point	2022	9	↓	$+, +, +$	Supports
	Gilead / Forty Seven	2020	9	↑	$-, +, \approx 0$	Contradicts
	Merck / Acceleron	2021	7	↓	$\approx 0, -, +$	Strongly contradicts
	Merck / Pandion	2021	9	↓	$+, +, +$	Strongly supports
	Pfizer / Seagen	2023	7	↑	$\approx 0, -, -$	Contradicts
Consumer Staples	Amazon / Whole Foods	2017	9	↓	$-, -, \approx 0$	Supports
	Danone / White Wave Foods	2016	6	↓	$+, \approx 0, \approx 0$	Supports
	J.M. Smucker / Hostess	2023	5	↓	$+, +, +$	Supports
	Kroger / Albertsons	2022	9	↓	$+, -, +$	Contradicts
	Reckitt / Mead-Johnson	2017	9	↑	$\approx 0, -, -$	Contradicts
Technology	AMD / Xilinx	2020	9	↑	$+, -, -$	Contradicts
	Avago / Broadcom	2015	5	↑	$-, -, +$	Contradicts
	Cisco / Splunk	2023	10	↑	$-, +, +$	Mixed
	Microchip / Microsemi	2018	6	↑	$-, -, -$	Mixed
	SS&C Technologies / DST Systems	2018	7	↑	$+, \approx 0, \approx 0$	Contradicts
Automotives	Apollo / Tenneco	2022	9	↑	$-, +, -$	Mixed
	Goodyear / Cooper	2021	9	↓	$+, +, +$	Supports
	Tenneco / FDML	2018	7	↓	$-, +, -$	Supports
	Warner / Delphi	2020	7	↑	$-, \approx 0, +$	Mixed
	Warner / Remy	2015	8	↓	$+, +, \approx 0$	Strongly supports
Finance	BB&T / SunTrust	2019	7	↑	$+, +, +$	Contradicts
	KeyCorp / FirstNiagara	2015	5	↓	$+, -, -$	Strongly supports
	People / United	2019	6	↑	$-, \approx 0, \approx 0$	Contradicts
	Sterling / Astoria	2017	5	↓	$+, -, -$	Supports
	WSFS / Beneficial	2018	6	↓	$+, +, +$	Supports
Summary				**Supports: 14**	**Contradicts: 12**	**Mixed: 4**

tical acquisitions, the predictability of the SEC's theory completely collapses. As summarized in Table 4, the results are essentially a coin flip: exactly half of the transactions (five out of ten) directly contradict the shadow trading hypothesis. In these contradictory cases, the firms with the highest 10-K semantic similarity to the target either experienced negligible price movement or negative abnormal returns on the announcement day.

For example, while transactions like AbbVie's acquisition of ImmunoGen generated the uniform sympathy rally the SEC expects (strong positive returns for the top three peers), deals like Pfizer's acquisition of Seagen and Gilead's acquisition of Forty Seven yielded fundamentally unaligned market reactions. In the Pfizer/Seagen transaction, the top three most semantically similar peers registered zero or negative abnormal returns. This directly refutes the premise that shared fundamental business traits in mandatory SEC filings reliably dictate correlated market behavior.

These findings carry direct implications for NLP-based enforcement heuristics. Even within this controlled cohort, a state-of-the-art LLM pipeline operating on the SEC's own mandated disclosures predicts peer stock reactions at chance level, suggesting that "economic linkage" as operationalized through disclosure text is an unreliable signal, not a consistent market phenomenon.

4.4 Cross-Industry Results

To determine whether the shadow trading phenomenon extends beyond the Biopharmaceutical sector, we expanded our NLP analysis to twenty additional acquisitions across four diverse industries: *Consumer Staples, Technology, Automotives*, and *Finance*. As detailed in Table 4, the expansion of the dataset further erodes the structural validity of the SEC's "economic linkage" theory, with results distributed across supporting, contradicting, and

Table 5: Results by industry.

Industry	Events	Supports	Contradicts	Mixed
Biopharmaceuticals	10	5	4	1
Consumer Staples	5	3	2	0
Technology	5	0	3	2
Automotives	5	3	0	2
Finance	5	3	2	0
Total	**30**	**14**	**12**	**4**

mixed outcomes.

The breakdown by sector reveals profound inconsistencies that undermine any notion of a universal, predictable market mechanism. For example, deals like AMD/Xilinx and Avago/Broadcom produced inverted trend lines where highly similar peers saw negligible or negative returns. Similarly, the Consumer Staples and Finance sectors exhibited near-random distributions of support and contradiction. In the BB&T/SunTrust banking merger, a robust sympathy rally lifted the entire cohort, yet the magnitude of those gains completely decoupled from semantic similarity, yielding a strongly contradictory trend line. As shown in Table 5, Technology is the only sector with zero supporting cases while Automotives yields zero contradictions, suggesting that document-recoverable economic linkage may be industry-contingent: tight supplier ecosystems in automotive manufacturing produce consistent spillover effects that dispersed platform-based technology firms do not.

Across the full 30-event dataset, semantic similarity successfully predicts peer stock movement direction in only $14/30$ cases (46.6%). A two-sided binomial test against $H_0 = 0.5$ yields $p \approx 0.86$: the signal is statistically indistinguishable from a coin flip. The data suggests that "economic linkage" is not a stable, document-recoverable feature of firm relationships but an idiosyncratic outcome that varies unpredictably across industries, deal structures, and market conditions—and that no NLP-based pipeline operating on public disclosures can reliably identify it beforehand.

5 Discussion

LLM similarity as an enforcement heuristic. Our pipeline demonstrates that LLM-based analysis of SEC filings can identify semantically similar firms with precision unavailable to BoW methods. In cases like AbbVie/ImmunoGen and BMS/Karuna, textual similarity correlates cleanly with announcement-day stock reactions. However, the overall hit rate of $14/30$ is too low to underpin a reliable enforcement standard. This matters for fairness: when enforcement depends on a similarity argument constructed after the fact, individuals cannot determine in advance which trades are prohibited. Our results suggest the SEC could deploy more equitable, sector-stratified watchlists in industries where the textual signal is strong, with oncology biotechnology being the clearest candidate. A recent graph-based surveillance framework for shadow trading (Stenfors et al., 2025) confirms the computational intensity of the problem, but better models alone will not solve an issue rooted in market structure variability rather than text quality.

Failure mode analysis. Qualitative inspection of contradicting cases reveals two recurring failure modes. The first is *sector-wide contagion*: in BB&T/SunTrust, all peers rallied regardless of similarity rank because the merger signaled regulatory approval of large-bank consolidation generally, swamping any firm-specific signal. The second is *outlier dominance*: in BMS/MyoKardia, a single firm (CYTK, $+14.3\%$ AR) drove an apparent negative trendline while the remaining peers showed negligible returns, making the case technically "contradicting" by trendline but reflecting one idiosyncratic event rather than a systematic pattern. These failure modes suggest that even where 10-K similarity is high, macro-structural events and single-stock noise can overwhelm the textual signal—a very difficult issue for an NLP pipeline to resolve.

Asymmetric identifiability. The fundamental issue is not that NLP fails to detect economic linkage—it clearly does in some cases—but that it fails *unpredictably*. An insider running our pipeline before trading would face genuine uncertainty about whether a given trade crosses the shadow trading line. The SEC, by contrast, can observe the announcement-day return first and construct a textual similarity argument after the fact. This is the *Panuwat* pattern exactly: "mid-cap oncology" was a category invoked *after* Incyte's stock moved, not before the trade was placed. When peer firms cannot be reliably identified beforehand from public disclosures, mass market surveillance becomes the only practical enforcement mechanism.

Implications for NLP-based enforcement tools. Our cross-industry results suggest that sector-stratified deployment could improve precision. In

biopharmaceuticals, our pipeline performs reliably and could support a targeted peer watchlist. In platform-based technology, the signal is consistently absent across all five events; no disclosure-based similarity measure is likely to recover it. Future work should investigate whether combining multiple filing sections (Items 1, 1A, and 7) or incorporating supplementary data (analyst peer groupings, institutional ownership overlap) improves predictive validity in low-signal sectors.

Related Work

Firm similarity from financial text. Prior work applying HP-style text similarity to financial disclosures has focused on market clustering, risk assessment, and portfolio optimization (Hanley and Hoberg, 2019; Chen and Sarkar, 2020), and on evaluating LLMs applied to 10-K sections for question answering and information extraction tasks (Araci, 2019; Wu et al., 2023). None of this work evaluates whether textual firm similarity predicts announcement-day stock reactions, nor does it examine the legal implications of that relationship. We extend this lineage by applying long-context generative LLM reasoning to operationalize the SEC's "economic linkage" standard and test it empirically against market outcomes.

Shadow trading in finance and law. Mehta et al. (2021) introduced "shadow trading" in the economics literature, estimating single-event profits of \$139,400–\$678,000; Augustin et al. (2019) documented positive abnormal options volume in ∼25% of U.S. takeovers between 1996–2012. The Brattle Group (2025) discuss enforcement implications from recent SEC cases. Stenfors et al. (2025) propose AMGIN, a graph-based surveillance framework that models the market as a spatio-temporal graph integrating sectoral ties and price co-movements; unlike our approach, AMGIN operates on market structure data rather than disclosure text and is designed as a regulator surveillance tool rather than an identifiability test beforehand. To our knowledge, this paper is the first to apply LLMs to mandatory SEC disclosures to operationalize economic linkage as a legal standard and evaluate whether that textual signal predicts peer stock reactions across industries.

6 Conclusion

We present the first NLP-grounded empirical evaluation of the SEC's shadow trading enforcement theory. Our two-stage LLM pipeline successfully replicates the factual basis of *SEC v. Panuwat* on its most analogous case, validating the methodology. Across 30 acquisition events, however, textual similarity in 10-K MD&A sections predicts positive peer stock reactions in only 14 of 30 cases (46.6%, binomial $p \approx 0.86$ against $H_0 = 0.5$), with 12 cases contradicting the hypothesis. The results are inconsistent both within Biopharmaceuticals and across industries, suggesting that "economic linkage" as a legal standard is not reliably identifiable from publicly mandated disclosures.

More broadly, our results suggest that peer identification from SEC disclosures is not a retrieval problem where better NLP would yield better results, but a fundamentally noisy prediction problem shaped by market structure, macro-level events, and deal-specific dynamics that disclosure text alone cannot resolve. This has implications for any NLP system designed to operationalize economic relationships from regulatory filings. Positively, our findings point toward a more tractable and equitable enforcement design: sector-stratified watchlists built from disclosure text in high-signal industries could replace indiscriminate mass surveillance, making the scope of liability legible to those subject to it before and not after the fact.

Limitations

Our study has several limitations. First, while the aggregate binomial test supports the headline result, the per-event Spearman correlations are computed over no more than 10 peers each and cannot themselves establish statistical significance. Additionally, Gemini 3.1 Pro is a closed-weight model whose outputs are structured qualitative judgments rather than fully reproducible quantitative measurements; all financial metrics are computed independently using open-source tools to mitigate this. Second, 10-K similarity captures only one dimension of economic linkage; shared institutional investors, common competitors, and supply-chain relationships are dimensions not captured by Item 7 text alone (Hoberg and Phillips, 2016). Third, stock price reactions on announcement day reflect not only information spillover but also market structure effects, liquidity constraints, and idiosyncratic noise, all of which are difficult to separate from the shadow trading signal with a one-day window.

Ethical Considerations

Data and privacy. All data used in this study consists of publicly available SEC filings retrieved from EDGAR and historical stock prices from public market feeds. No private, proprietary, or personally identifiable information was accessed at any stage.

Dual use. The primary ethical concern is dual use: a pipeline that tests whether peer firms are identifiable in advance could in principle be used by a would-be insider trader to assess whether a contemplated trade would be detectable under the shadow trading theory. We believe the public interest in evaluating the empirical basis of a contested enforcement doctrine outweighs this risk, particularly since our central finding is that the pipeline *fails* at the task in the majority of cases, providing little operational guidance to a potential bad actor that is not already available through standard financial analysis.

Positive social impact. Across law and regulation, NLP and ML systems are increasingly used to flag enforcement targets and support legal determinations—from recidivism scoring to financial surveillance—yet their reliability is rarely tested before deployment. The EU AI Act classifies such systems as high-risk precisely because errors in legal contexts cause asymmetric harm (European Parliament and Council of the European Union, 2024). This paper contributes a methodology for measuring that reliability before deployment, applied to a concrete case: whether the "economic linkage" standard underlying insider trading enforcement is recoverable from mandatory public disclosures. We believe this contributes positively in three ways. First, it provides empirical evidence bearing on whether a legal enforcement standard—one that currently triggers mass financial surveillance of over 100 million U.S. investors—is grounded in a tractable prediction task. Second, our cross-industry findings point toward a more targeted and equitable enforcement design: sector-stratified watchlists built from disclosure text in high-signal industries (e.g., oncology biotechnology) could make the scope of liability more legible to those subject to it, replacing indiscriminate surveillance with principled, text-grounded targeting. Third, this paper demonstrates a general methodology—using NLP to empirically evaluate whether a legally-defined relationship is text-recoverable from mandatory disclosures—that can be applied to other regulatory standards where similar questions arise. We hope this encourages further NLP work at the intersection of regulatory transparency, legal fairness, and computational social science.

References

Dogu Araci. 2019. Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Patrick Augustin, Menachem Brenner, and Marti G. Subrahmanyam. 2019. Informed options trading prior to takeover announcements: Insider trading? *Management Science*, 65(12):5697–5720.

Jiafeng Chen and Suproteem Sarkar. 2020. A semantic approach to financial fundamentals. In *Proceedings of the Second Workshop on Financial Technology and Natural Language Processing*, pages 22–26, Kyoto, Japan. -.

European Parliament and Council of the European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council... (Artificial Intelligence Act).

Google DeepMind. 2026. Gemini 3.1 Pro.

Kathleen Weiss Hanley and Gerard Hoberg. 2019. Dynamic interpretation of emerging risks in the financial sector. *The Review of Financial Studies*, 32(12):4543–4603.

Harshavardhan. 2026. Self-anchoring calibration drift in large language models: How multi-turn conversations reshape model confidence. *Preprint*, arXiv:2603.01239.

Gerard Hoberg and Gordon Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Jie Huang, Xinyun Chen, Swaroop Mishra, Huaixiu Steven Zheng, Adams Wei Yu, Xinying Song, and Denny Zhou. 2024. Large language models cannot self-correct reasoning yet. *Preprint*, arXiv:2310.01798.

John Jenkins. 2025. Insider trading policies: Shadow trading.

Tushar Khot, Harsh Trivedi, Matthew Finlayson, Yao Fu, Kyle Richardson, Peter Clark, and Ashish Sabharwal. 2023. Decomposed prompting: A modular approach for solving complex tasks. *Preprint*, arXiv:2210.02406.

Junyi Li, Jie Chen, Ruiyang Ren, Xiaoxue Cheng, Wayne Xin Zhao, Jian-Yun Nie, and Ji-Rong Wen. 2024. The dawn after the dark: An empirical study

on factuality hallucination in large language models. *Preprint*, arXiv:2401.03205.

Mihir N. Mehta, David M. Reeb, and Wanli Zhao. 2021. Shadow trading. *The Accounting Review*, 96(4):367–404.

New Civil Liberties Alliance. 2024a. Complaint at 2, Davidson, Restivo, and NCPPR v. Atkins, no. 6:24-cv-00197 (w.d. tex. apr. 16, 2024).

New Civil Liberties Alliance. 2024b. Complaint, Davidson, Restivo, and NCPPR v. Atkins, no. 6:24-cv-00197 (w.d. tex.).

Hester Peirce. 2019. This CAT is a dangerous dog. *RealClearPolicy*.

Alexis Stenfors, Ting Guo, Boyu Li, Kaveesha Hewage, Peter Mere, and Fang Chen. 2025. Shadow trading detection: A graph-based surveillance approach. *Finance Research Letters*, 86:108524.

Supreme Court of the United States. 1997. United States v. O'Hagan, 521 U.S. 642.

Supreme Court of the United States. 2018. Carpenter v. United States, 585 U.S. 296.

The Brattle Group. 2025. Shadow trading: Economic evidence and potential implications from recent SEC insider trading cases. Brattle Group White Paper.

United States District Court, N.D. Cal. 2023. SEC v. Panuwat, 702 F. Supp. 3d 883. No. 21-cv-06322-WHO.

U.S. Securities and Exchange Commission. 2024. *SEC v. Bechtolsheim*. Settled May 2024; civil penalty of $923,740 without admission of liability. Replace with the official docket number when finalizing.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. Chain-of-thought prompting elicits reasoning in large language models. *Preprint*, arXiv:2201.11903.

Shijie Wu, Ozan Irsoy, Steven Lu, Vadim Dabravolski, Mark Dredze, Sebastian Gehrmann, Prabhanjan Kambadur, David Rosenberg, and Gideon Mann. 2023. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

A Prompt Templates

We define our multi-stage prompting architecture in Table 6, which outlines the system instructions and the dimensional rubrics provided to the model.

Table 6: LLM Prompt Specifications.

Type	Prompt Content
System	You are a securities analyst and computational linguist specializing in the analysis of SEC regulatory filings. You assess economic similarity between public companies based solely on their formal disclosures, applying the standard a sophisticated reasonable investor would use when evaluating whether two firms occupy the same market niche.
Stage 1	I have uploaded the 10-K Item 7 (MD&A) for [TARGET_COMPANY], covering fiscal year [YEAR]. This was the primary disclosure available when [ACQUIRER] announced its acquisition of [TARGET_COMPANY] in [MONTH_YEAR]. Using only this document, identify 10 publicly traded companies that were most similar to [TARGET_COMPANY] as of [MONTH_YEAR]. Constraints: - Use No Hindsight: Base your assessment only on the state of the market as of [MONTH_YEAR] - Reasonable Insider Standard: Prioritize firms that an employee at [TARGET_COMPANY] would realistically view as a peer or benchmark company - Do not include [ACQUIRER] in your results - SEC Filing Requirement: All 10 companies must have an active SEC filing obligation (10-K or 20-F). - Price Data Availability: All 10 companies must have been actively traded on a major US exchange (NYSE, NASDAQ). For each, provide: A) Name/Ticker; B) Dimensional Justification (B1: Technical, B2: Strategic, B3: Risk); C) Validation Document.
Stage 2	Using only the content of these uploaded documents, assign each candidate a Total Similarity Score (0.00 to 1.00) based on: A) Technical/Product Overlap (40%): Core technology or product engines. B) Financial/Strategic Stage (30%): Revenue growth and commercialization milestones. C) Risk & Market Exposure (30%): Geographic focus and competitive headwinds. Additionally, apply dual-track analysis: D) Noun-Based Technical Comparison (HP Logic): Identify specific mechanism-of-action nouns. E) Semantic Bridging: Resolve instances where companies describe identical strategies using different terminology. This is critical to assessing the "Administrability Gap."

Table 7: Stage 1 Output: Pfizer/Medivation

#	Company (Ticker)	B1: Technical	B2: Strategic	B3: Risk
1	Incyte (INCY)	Targeted oncology and hematologic malignancies overlap (e.g., lymphoma, myeloma).	Single-product revenue model with major global collaboration (e.g., Jakafi/Novartis).	Exposure to gross-to-net pricing dynamics and reliance on ex-US partners.
2	Exelixis (EXEL)	Small-molecule oncology therapies overlapping prostate and breast cancer.	Transition to fully integrated commercial model with internal oncology sales force.	Inventory volatility and R&D expense tied to late-stage trials.
3	Tesaro (TSRO)	PARP inhibitor targeting ovarian and breast cancer; direct overlap with MDV3800.	Scaling SG&A for commercial launch; growth via in-licensing.	Manufacturing scale-up risks for PARP inhibitors.
4	Clovis Oncology (CLVS)	PARP inhibitor (rucaparib) for ovarian and prostate cancers.	Focus on genetically defined populations using companion diagnostics.	FDA setbacks and clinical/regulatory volatility.
5	Seattle Genetics (SGEN)	Antibody-based oncology therapies for lymphoma and breast cancer.	Single-product revenue base reinvested into label expansion; co-development partnerships.	Pricing pressure and complex collaboration revenue accounting.
6	Ariad Pharmaceuticals (ARIA)	Targeted therapies for hematologic and solid tumors.	Transition to independent commercialization with specialized sales teams.	High exposure to drug pricing pressure and advocacy scrutiny.
7	Puma Biotechnology (PBYI)	Targeted breast cancer therapy (neratinib) overlapping TNBC strategy.	Growth via in-licensed assets from large pharma.	Large contingent liabilities tied to milestone payments.
8	Merrimack (MACK)	Targeted oncology therapies for solid tumors (breast, pancreatic).	Early-stage commercialization with ex-US licensing partnerships.	Manufacturing scale-up and clinical trial cost expansion risks.
9	Immunomedics (IMMU)	Antibody therapies targeting TNBC; direct indication overlap.	Advancing assets into Phase 3 with companion diagnostics.	Regulatory volatility and clinical trial design risk.
10	Celldex Therapeutics (CLDX)	Antibody-based immuno-oncology therapies for breast cancer and lymphoma.	Mixed internal discovery and in-licensing strategy.	High IPR&D impairment risk and clinical failure exposure.

Table 8: Stage 2 Output: Pfizer/Medivation

Rank	Company	Tech	Stage	Risk	Total	Semantic Justification
1	TSRO	0.38	0.25	0.25	**0.88**	Identical PARP inhibitor terminology and breast/ovarian cancer focus; overlapping PD-1 immuno-oncology pipeline narrative.
2	EXEL	0.30	0.28	0.28	**0.86**	Matching U.S. profit-sharing and ex-U.S. royalty structures; identical mCRPC clinical focus.
3	INCY	0.25	0.29	0.28	**0.82**	Single-product revenue with ex-U.S. partner; identical Medicare Part D gross-to-net pricing dynamics.
4	CLVS	0.35	0.20	0.25	**0.80**	Direct PARP inhibitor overlap; mirrored IPR&D impairment and FDA-related clinical setbacks.
5	SGEN	0.25	0.28	0.25	**0.78**	Identical U.S./ex-U.S. commercialization split and healthcare pricing pressure exposure.
6	ARIA	0.20	0.25	0.25	**0.70**	Similar reliance on third-party data and exposure to global pricing and reimbursement pressures.
7	MACK	0.20	0.22	0.20	**0.62**	Parallel ex-U.S. licensing strategy used to fund internal oncology R&D.
8	CLDX	0.30	0.10	0.15	**0.55**	Immuno-oncology and TNBC overlap; weaker commercial-stage alignment.
9	PBYI	0.25	0.15	0.15	**0.55**	Breast cancer focus with similar build-out of internal oncology sales infrastructure.
10	IMMU	0.28	0.10	0.15	**0.53**	Strong TNBC overlap but penalized for pre-commercial financial stage.

Table 9: Stage 1 Output: Pfizer / Seagen (2023)

#	Company (Ticker)	B1: Technical	B2: Strategic	B3: Risk
1	Genmab (GMAB)	Direct overlap in ADCs and monoclonal antibodies; co-development of TIVDAK targeting tissue factor.	50/50 co-commercialization and profit-sharing structure identical to Seagen.	Identical product-specific risks (TIVDAK adoption, eye toxicity management, global rollout).
2	Exelixis (EXEL)	Overlap in targeted oncology; ADC pipeline targeting tissue factor (XB002).	Large-scale commercial oncology firm with $1B+ revenue and global partnerships.	Exposure to U.S. pricing pressure (Medicaid rebates, 340B) and oncology competition.
3	ADC Therapeutics (ADCT)	Pure-play ADC company targeting hematologic malignancies (e.g., CD30-expressing lymphomas).	Recently transitioned to commercial-stage biotech with FDA accelerated approvals.	Manufacturing complexity and geographic commercialization via partners.
4	ImmunoGen (IMGN)	Core ADC platform using monoclonal antibodies to deliver cytotoxic payloads.	Hybrid revenue model (product + licensing/royalties) similar to Seagen.	ADC-specific manufacturing, safety, and clinical execution risks.
5	MacroGenics (MGNX)	Monoclonal antibodies, bispecifics, and ADCs targeting HER2-positive cancers.	Early-commercial oncology firm with licensing and internal pipeline development.	Competitive pressure in HER2 oncology and commercialization cost burden.
6	Blueprint Medicines (BPMC)	Targeted oncology therapies (kinase inhibitors); overlap via precision oncology.	Commercial-stage biotech with expanding global infrastructure.	Medicare/340B pricing exposure and specialty distribution dependence.
7	Zymeworks (ZYME)	HER2-targeted ADCs and bispecific antibodies overlapping Seagen's pipeline.	Pre-commercial with heavy reliance on licensing deals and milestone funding.	Competitive dynamics in HER2 space and pipeline execution risk.
8	Incyte (INCY)	Broad oncology pipeline (JAK inhibitors, antibodies); limited ADC overlap.	Large commercial biotech with significant royalty revenue streams.	Pricing reform exposure (IRA), global partner reliance, and reimbursement risk.
9	Mersana Therapeutics (MRSN)	ADC platform using proprietary payload technologies (e.g., microtubule inhibitors).	Pre-commercial biotech relying on partnerships to fund R&D.	High clinical risk and early-stage financial uncertainty.
10	Sutro Biopharma (STRO)	Next-gen ADCs with linker-warhead architecture; strong platform similarity.	Early-stage firm focused on partnerships and pipeline development.	Manufacturing scale-up and pre-commercial execution risk.

Table 10: Stage 2 Output: Pfizer / Seagen (2023)

Rank	Company	Tech	Stage	Risk	Total	Semantic Justification
1	GMAB	0.38	0.26	0.28	**0.92**	Exact match via TIVDAK co-development; identical product, revenue structure, and geographic commercialization risk profile.
2	EXEL	0.30	0.28	0.28	**0.86**	Shared targeted oncology focus; identical tissue-factor ADC targeting and nearly identical gross-to-net pricing exposure.
3	ADCT	0.38	0.18	0.26	**0.82**	Pure-play ADC model targeting hematologic malignancies; strong payload-delivery semantic alignment with Seagen.
4	IMGN	0.38	0.15	0.25	**0.78**	Identical "antibody-drug conjugate" architecture; payload terminology bridges directly to Seagen's cell-killing mechanism.
5	MGNX	0.34	0.15	0.22	**0.71**	HER2 oncology overlap with similar commercialization dynamics and competitive pressures.
6	BPMC	0.25	0.20	0.25	**0.70**	Shared oral oncology commercialization risks (Medicare Part D, 340B) despite different molecular modality.
7	ZYME	0.35	0.12	0.20	**0.67**	Direct HER2-targeted ADC overlap but weaker commercial-stage alignment.
8	INCY	0.15	0.25	0.26	**0.66**	Strong financial structure similarity (royalties, partnerships) despite technical divergence.
9	MRSN	0.36	0.08	0.15	**0.59**	ADC payload equivalence but large gap due to pre-commercial financial profile.
10	STRO	0.35	0.08	0.15	**0.58**	Platform-level ADC similarity ("linker-warhead") but minimal commercialization overlap.

B Notable Case Studies

B.1 Ground Truth: Pfizer / Medivation (2016)

The Pfizer/Medivation acquisition serves as the factual basis for SEC v. Panuwat. Applying our pipeline ex ante, the model ranks Tesaro, Exelixis, and Incyte as the closest peers, reflecting similarities in oncology pipelines, commercialization approach, and pricing exposure.

The market response aligns with these rankings: all top three peers exhibit positive abnormal returns (+1.60%, +1.33%, +5.04%). Incyte, the firm actually traded, ranks among the most similar and shows the strongest return. This case demonstrates that textual similarity can, in some instances, identify economically linked firms consistent with the shadow trading theory.

Tables 7 and 8 show the two-stage LLM outputs. Tables 11 and 12 summarize the abnormal return calculations and the combined similarity-ranking analysis, respectively, while Figure 3 visualizes the relationship between similarity rank and abnormal returns.

Table 11: Abnormal Returns Medivation

Ticker	Raw Return	Abnormal Return	Direction
CLDX	-0.012658227848101266	-0.03933565933364627	Decreased
EXEL	0.039999961853027344	0.013322530367482339	Increased
INCY	0.07712524897527939	0.050447817489734384	Increased
PBYI	0.09103616209622678	0.06435873061068177	Increased
TSRO	0.04272332537878824	0.01604589389324324	Increased

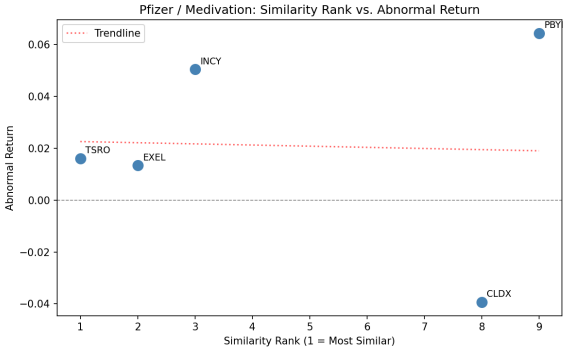

Figure 3: Pfizer / Medivation Similarity vs. Abnormal Returns

Table 12: Final Analysis Medivation

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	TSRO	0.88	0.0427233253787882	0.0160458938932432	Increased	0.014120386626054016
2	EXEL	0.86	0.0399999618530273	0.0133225303674823	Increased	0.011457376116034779
3	INCY	0.82	0.0771252489752793	0.0504478174897343	Increased	0.041367210341582125
8	CLDX	0.55	-0.0126582278481012	-0.0393356593336462	Decreased	-0.021634612633505412
9	PBYI	0.55	0.0910361620962267	0.0643587306106817	Increased	0.03539730183587494

B.2 Contradicting Case: Pfizer / Seagen (2023)

The Pfizer/Seagen acquisition provides a direct counterexample in the same industry. Using the identical pipeline, the model again identifies highly similar oncology peers with comparable technical and financial profiles.

However, the market response diverges: the most similar firms show negligible or negative abnormal returns ($\approx$ 0%, -, -) , and the overall relationship between similarity and returns reverses. Despite strong textual similarity, no clear trading signal emerges, contradicting the predictability assumed by the shadow trading theory.

Tables 9 and 10 show the two-stage LLM outputs. Tables 13 and 14 summarize the abnormal return calculations and the combined similarity-ranking analysis, respectively, while Figure 4 visualizes the relationship between similarity rank and abnormal returns.

Table 13: Abnormal Returns Seagen

Ticker	Raw Return	Abnormal Return	Direction
ADCT	-0.07299271787868707	-0.10207023020129924	Decreased
EXEL	0.0077937145316635875	-0.02128379779094858	Decreased
GMAB	0.03297011685681662	0.003892604534204449	Increased
INCY	0.01624771745919238	-0.012829794863419789	Decreased
MGNX	0.10526314374246708	0.0761856314198549	Increased
STRO	0.05847953303345698	0.02940202071084481	Increased
ZYME	0.040353112054028716	0.011275599731416546	Increased

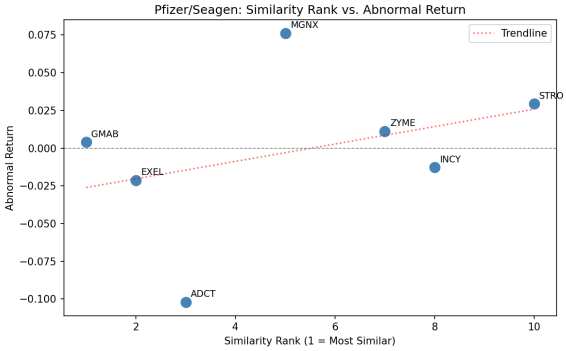

Figure 4: Pfizer / Seagen Similarity vs. Abnormal Returns

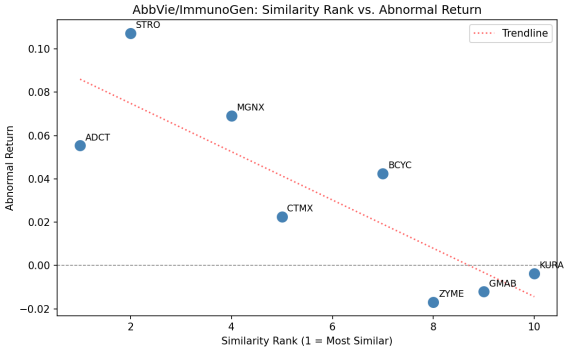

Figure 5: AbbVie / ImmunoGen Similarity vs. Abnormal Returns

Table 14: Final Analysis Seagen

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	GMAB	0.92	0.0329701168568166	0.0038926045342044	Increased	0.003581196171468048
2	EXEL	0.86	0.0077937145316635	-0.0212837977909485	Decreased	-0.01830406610021571
3	ADCT	0.82	-0.0729927178786870	-0.1020702302012992	Decreased	-0.08369758876506533
5	MGNX	0.71	0.1052631437424670	0.0761856314198549	Increased	0.05409179830809698
7	ZYME	0.67	0.0403531120540287	0.0112755997314165	Increased	0.00755465182004905 5
8	INCY	0.66	0.0162477174591923	-0.0128297948634197	Decreased	-0.008467664609857003
10	STRO	0.58	0.0584795330334569	0.0294020207108448	Increased	0.01705317201228998 3

Table 16: Final Analysis ImmunoGen

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	ADCT	0.93	0.0779220829486124	0.0553085158053836	Increased	0.05143691969900675
2	STRO	0.87	0.1298701277254752	0.1072565605822464	Increased	0.09331320770655437
4	MGNX	0.74	0.0917553269907650	0.0691417598475362	Increased	0.05116490228717679 4
5	CTMX	0.68	0.0451127374763250	0.0224991703330962	Increased	0.01529943582650541 6
7	BCYC	0.61	0.0650886931191834	0.0424751259759546	Increased	0.02590982684533230 5
8	ZYME	0.52	0.0056689560617790	-0.0169446110814497	Decreased	-0.00881119776235384 5
9	GMAB	0.40	0.0105498694521109	-0.0120636976911178	Decreased	-0.00482547907644712
10	KURA	0.35	0.0189673666509025	-0.0036462004923262	Decreased	-0.00127617017231417

C Remaining Case Studies

These subsections present the empirical case studies and results of the paper's analysis. These sections apply the NLP pipeline to real-world acquisition events and examine whether semantic similarity between firms (based on SEC filings) actually predicts stock price reactions—i.e., whether "shadow trading" is empirically supported.

C.1 Biopharmaceuticals

C.1.1 AbbVie / ImmunoGen

AbbVie's acquisition of ImmunoGen is one of the clearest supporting cases in the dataset. The pipeline identifies ADC-focused peers (ADCT, STRO, MGNX) that show uniformly positive abnormal returns on announcement day, with the top-ranked firm (ADCT, $+5.5\%$ AR) exhibiting the strongest reaction. The negative trendline and positive top-3 AR direction confirm that textual similarity in this case successfully predicts the direction and magnitude of peer stock reactions, consistent with the shadow trading hypothesis.

Table 15: Abnormal Returns ImmunoGen

Ticker	Raw Return	Abnormal Return	Direction
ADCT	0.077922082948612 43	0.05530851580538361	Increased
BCYC	0.06508869311918347	0.04247512597595465	Increased
CTMX	0.045112737476325084	0.022499170333096268	Increased
GMAB	0.010549869452110954	-0.01206369769111786	Decreased
KURA	0.018967366650902547	-0.0036462004923262688	Decreased
MGNX	0.09175532699076502	0.0691417598475362	Increased
SNDX	0.02116572005222625	-0.0014478470910025652	Decreased
STRO	0.12987012772547524	0.10725656058224642	Increased
ZYME	0.005668956061779067	-0.016944611081449748	Decreased

C.1.2 BMS / Karuna Therapeutics

The BMS/Karuna Therapeutics case is classified as supporting, driven primarily by the top-ranked firm RVPH ($+17.0\%$ AR), an extreme positive outlier. The remaining peers show mixed or near-zero reactions. This case illustrates the outlier dominance failure mode: the negative trendline is real but largely attributable to one firm rather than a systematic pattern across the cohort.

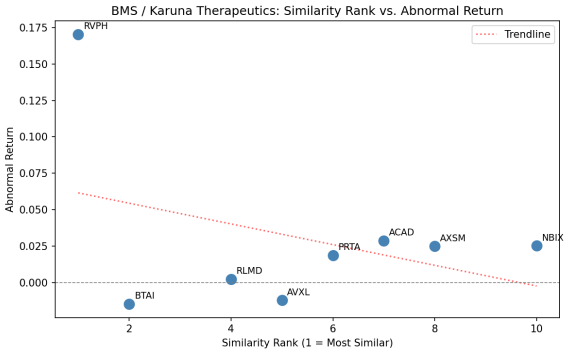

Figure 6: BMS / Karuna Therapeutics Similarity vs. Abnormal Returns

Table 17: Abnormal Returns Karuna Therapeutics

Ticker	Raw Return	Abnormal Return	Direction
ACAD	0.0636856074394867	0.02884434523941582	Increased
AVXL	0.022857121058887277	-0.011984141141198112	Decreased
AXSM	0.05980958724745921	0.02496832504738833	Increased
BTAI	0.02020200074126503	-0.014639261458805852	Decreased
NBIX	0.06016831516828872	0.025327052968217842	Increased
PRTA	0.053662462447954765	0.018821200247883883	Increased
RLMD	0.037288099429759344	0.002446837229688463	Increased
RVPH	0.20519484482802336	0.17035358262795247	Increased

Table 18: Final Analysis Karuna Therapeutics

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	RVPH	0.94	0.2051948448280233	0.1703535826279524	Increased	0.1601323676702752 2
2	BTAI	0.87	0.0202020007412650	-0.0146392614588058	Decreased	-0.0127361574691610 45
4	RLMD	0.84	0.0372880994297593	0.0024468372296884	Increased	0.0020553432729382 56
5	AVXL	0.81	0.0228571210588727	-0.011984141141981	Decreased	-0.0097071543243704 62
6	PRTA	0.76	0.0536624624479547	0.0188212002478838	Increased	0.0143041121883916 88
7	ACAD	0.73	0.0636856074394867	0.0288443452394158	Increased	0.0210563720247735 34
8	AXSM	0.72	0.0598095872474592	0.0249683250473883	Increased	0.0179771940341195 77
10	NBIX	0.63	0.0601683151682887	0.0253270529682178	Increased	0.0159560433699772 15

C.1.3 BMS / MyoKardia

Despite a negative trendline, BMS/MyoKardia is classified as strongly contradicting. The top-ranked firm CYTK (+14.3% AR) is a large positive outlier, while all remaining high-similarity peers show negative or negligible ARs. This is a textbook outlier dominance case: the trendline direction is mechanically driven by a single observation and does not reflect a consistent relationship between similarity and peer stock reactions.

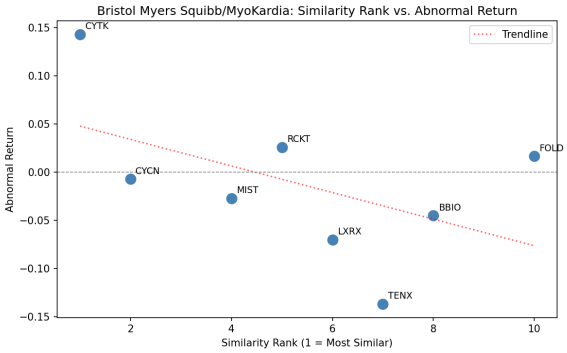

Figure 7: BMS / MyoKardia Similarity vs. Abnormal Returns

Table 19: Abnormal Returns Myokardia

Ticker	Raw Return	Abnormal Return	Direction
BBIO	0.004292979550332126	-0.04509094642333769	Decreased
CYCN	0.04213942069154834	-0.00724450528212147	Decreased
CYTK	0.19258027858739152	0.1431963526137217	Increased
FOLD	0.06607270764700231	0.016688781673332494	Increased
LXRX	-0.02068971697413619	-0.070073642947806	Decreased
MIST	0.02215658726588405	-0.027227338707785764	Decreased
RCKT	0.075305691387476	0.02592176541380619	Increased
TENX	-0.08759124087591241	-0.13697516684958222	Decreased

Table 20: Final Analysis Myokardia

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	CYTK	0.94	0.1925802785873915	0.1431963526137217	Increased	0.1346045714568984 1
2	CYCN	0.80	0.0421394206915483	-0.0072445052821214	Decreased	-0.0057956042256971 2
4	MIST	0.69	0.0221565872658840	-0.0272273387077857	Decreased	-0.0187868637083721 3
5	RCKT	0.63	0.075305691387476 0	0.0259217654138061	Increased	0.0163307122106978 42
6	LXRX	0.58	-0.0206897169741361	-0.0700736429478060	Decreased	-0.0406427129097274 8
7	TENX	0.50	-0.0875912408759124	-0.1369751668495822	Decreased	-0.0684875834247911
8	BBIO	0.42	0.0042929795503321	-0.0450909464233376	Decreased	-0.0189381974978017 93
10	FOLD	0.30	0.0660727076470023	0.0166887816733324	Increased	0.0050066345019997 2

C.1.4 BMS / Turning Point Therapeutics

BMS/Turning Point Therapeutics is a clean supporting case. The top three peers (RLAY, RVMD, NUVL) all show positive abnormal returns, and the negative trendline is consistent across the cohort. This case, alongside AbbVie/ImmunoGen and Merck/Pandion, represents the clearest evidence that the shadow trading hypothesis can be validated from 10-K text in targeted oncology contexts.

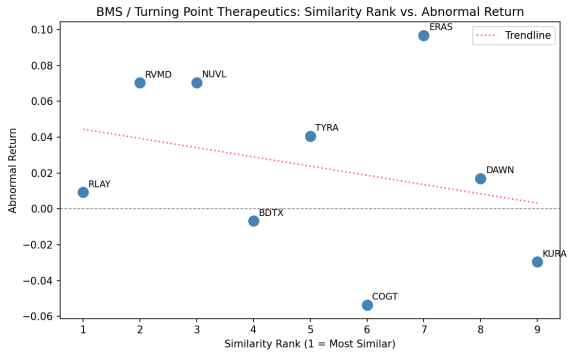

Figure 8: BMS / Turning Point Therapeutics Similarity vs. Abnormal Returns

Table 21: Abnormal Returns Turning Point Therapeutics

Ticker	Raw Return	Abnormal Return	Direction
BDTX	0.02873560462197694	-0.006745240120843793	Decreased
COGT	-0.018140572896618242	-0.05362141763943898	Decreased
DAWN	0.052473748220260734	0.01699290347744	Increased
ERAS	0.13226460010465574	0.096783755361835	Increased
KURA	0.005930313199658817	-0.029550531543161916	Decreased
NUVL	0.10589640389057788	0.07041555914775714	Increased
RLAY	0.0446960649171423	0.009215220174321567	Increased
RVMD	0.10578034707572571	0.07029950233290498	Increased
TYRA	0.07613469791143926	0.040653853168618526	Increased

Table 22: Final Analysis Turning Point Therapeutics

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	RLAY	0.92	0.0446960649171423	0.0092152201743215	Increased	0.0084780025603757 8
2	RVMD	0.91	0.1057803470757257	0.0702995023329049	Increased	0.0639725471229434 7
3	NUVL	0.90	0.1058964038905778	0.0704155591477571	Increased	0.0633740032329814
4	BDTX	0.87	0.0287356046219769	-0.0067452401208437	Decreased	-0.0058683588905134 019
5	TYRA	0.83	0.0761346979114392	0.0406538531686185	Increased	0.0337426981299533 5
6	COGT	0.82	-0.0181405728966182	-0.0536214176394389	Decreased	-0.0439695624643399
7	ERAS	0.81	0.1322646001046557	0.0967837553618350	Increased	0.0783948418430863 6
8	DAWN	0.77	0.0524737482202607	0.0169929034774400	Increased	0.0130845356776288 02
9	KURA	0.76	0.0059303131996588	-0.0295505315431619	Decreased	-0.0224584039728030 45

C.1.5 BMS / Mirati Therapeutics

BMS/Mirati Therapeutics is classified as contradicting despite a negative trendline. The top-3 AR direction is mixed $(+, -, +)$, and the second-ranked

firm (TNGX, -12.5% AR) shows a large negative reaction that disrupts any clean positive signal. The negative trendline is driven by the interaction between RVMD (high similarity, positive AR) and TNGX (high similarity, strongly negative AR), producing a noisy, uninterpretable pattern overall.

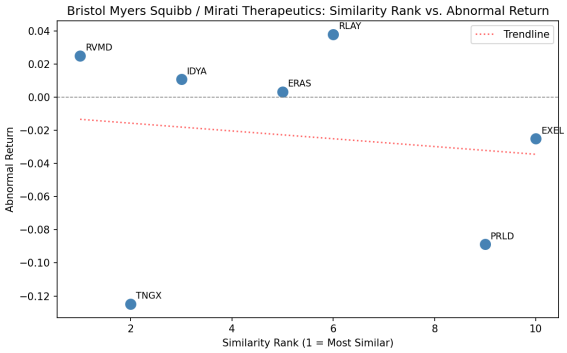

Figure 9: BMS / Mirati Therapeutics Similarity vs. Abnormal Returns

Table 23: Abnormal Returns Mirati Therapeutics

Ticker	Raw Return	Abnormal Return	Direction
ERAS	-0.004524882484455325	0.0033664856347955076	Increased
EXEL	-0.03285326574234348	-0.024961897623092644	Decreased
IDYA	0.00299962217762286	0.010890990296873692	Increased
PRLD	-0.09665426949169337	-0.08876290137244254	Decreased
RLAY	0.030156815579058048	0.03804818369830888	Increased
RVMD	0.017172459349252746	0.02506382746850358	Increased
TNGX	-0.13268894672203788	-0.12479757860278705	Decreased

Table 24: Final Analysis Mirati Therapeutics

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	RVMD	0.91	0.0171724593492527	0.0250638274685035	Increased	0.022808082996338186
2	TNGX	0.85	-0.132688946722037 8	-0.124797578602787 0	Decreased	-0.10607794181236894
3	IDYA	0.84	0.002999622177622 8	0.010890990296873 6	Increased	0.009148431849373823
5	ERAS	0.80	-0.004524882484455 3	0.003366485634795 5	Increased	0.002693188507836 4
6	RLAY	0.78	0.030156815579058 0	0.038048183698308 8	Increased	0.029677583284680863
9	PRLD	0.58	-0.096654269491693 3	-0.088762901372442 5	Decreased	-0.051482482796016645
10	EXEL	0.45	-0.032853265742343 4	-0.024961897623092 6	Decreased	-0.01123285393039167

C.1.6 Gilead / Forty Seven

Gilead/Forty Seven is classified as contradicting. The trendline is positive (higher similarity correlates with lower ARs), and the top-ranked firm shows a negative abnormal return. The case is notable because IPHA, ranked ninth, shows the largest positive AR ($+9.2\%$), inverting the expected pattern entirely.

Table 25: Abnormal Returns Forty Seven

Ticker	Raw Return	Abnormal Return	Direction
ALEC	0.03384280642895883	-0.005628199429125465	Decreased
CLDX	-0.015209200962377005	-0.0546802068204613	Decreased
FATE	0.06198628146063721	0.02251527560255292	Increased
IPHA	0.1309903617959911	0.0915193559379068	Increased
KURA	0.040596581373783934	0.0011255755156996394	Increased
SNDX	0.041445308145250374	0.00197430228716608	Increased
TGTX	0.11553783165163699	0.0760668257935527	Increased

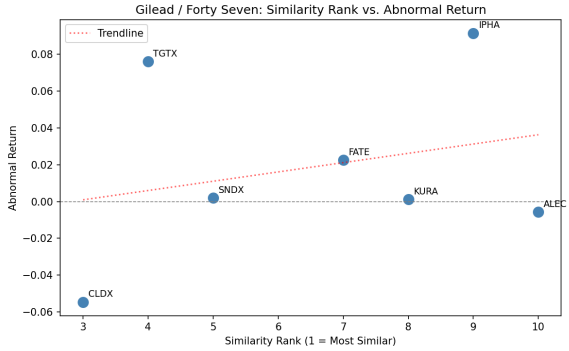

Figure 10: Gilead / Forty Seven Similarity vs. Abnormal Returns

Table 26: Final Analysis Forty Seven

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
3	CLDX	0.84	-0.0152092009623770	-0.0546802068204613	Decreased	-0.045931373729187486
4	TGTX	0.80	0.1155378316516369	0.0760668257935527	Increased	0.06085346063484216
5	SNDX	0.75	0.0414453081452503	0.0019743022871660	Increased	0.001480726715374 5
7	FATE	0.69	0.0619862814606372	0.0225152756025529	Increased	0.0155355401657615
8	KURA	0.59	0.0405965813737839	0.0011255755156996	Increased	0.0006640895542627639
9	IPHA	0.58	0.1309903617959911	0.0915193559379068	Increased	0.0530812264439859 4
10	ALEC	0.55	0.0338428064289588	-0.0056281994291254	Decreased	-0.00309550968601897

C.1.7 Merck / Acceleron Pharma

Merck/Acceleron Pharma is strongly contradicting. Despite a negative trendline, the top-ranked peers show near-zero or negative ARs, while the two firms with the highest and only positive ARs (SRRK, PTGX) rank in the middle of the similarity distribution. The announcement-day signal is weak across the cohort and does not concentrate in the most similar firms.

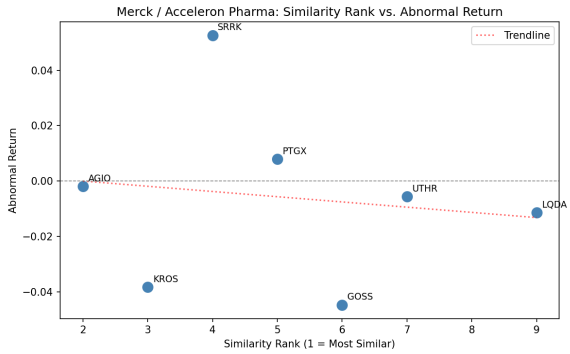

Figure 11: Merck / Acceleron Pharma Similarity vs. Abnormal Returns

Table 27: Abnormal Returns Acceleron Pharma

Ticker	Raw Return	Abnormal Return	Direction
AGIO	-0.0012983674482242133	-0.0019350482482927741	Decreased
GOSS	-0.04410645935599469	-0.04474314015606325	Decreased
KROS	-0.03770370260394145	-0.03834038340401001	Decreased
LQDA	-0.010752678064499305	-0.011389358864567866	Decreased
PTGX	0.008537257890474546	0.007900577090405985	Increased
SRRK	0.05326953926549374	0.05263285846542518	Increased
UTHR	-0.004959558862238881	-0.005596239662307442	Decreased

Table 28: Final Analysis Acceleron Pharma

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
2	AGIO	0.70	-0.0012983674482242	-0.0019350482482927	Decreased	-0.00135453377380489
3	KROS	0.68	-0.0377037026039414	-0.0383403834040100	Decreased	-0.0260714607147268
4	SRRK	0.65	0.0532695392654937	0.0526328584654251	Increased	0.034211358002526 32
5	PTGX	0.60	0.0085372578904745	0.0079005770904059	Increased	0.00474034625424354
6	GOSS	0.60	-0.0441064593559946	-0.0447431401560632	Decreased	-0.026845884093637917
7	UTHR	0.50	-0.0049595588622388	-0.0055962396623074	Decreased	-0.00279811983115 37
9	LQDA	0.40	-0.0107526780644993	-0.0113893588645678	Decreased	-0.00455574354582712

C.1.8 Merck / Pandion Therapeutics

Merck/Pandion Therapeutics is a strongly supporting case. All top-ranked peers show positive abnormal returns, and the pattern holds across the full cohort: six of seven firms show positive ARs, with the negative case (PTGX) ranking sixth. This is one of the most consistent results in the dataset and suggests that in immuno-oncology acquisitions with tightly defined peer groups, textual similarity can reliably predict announcement-day reactions.

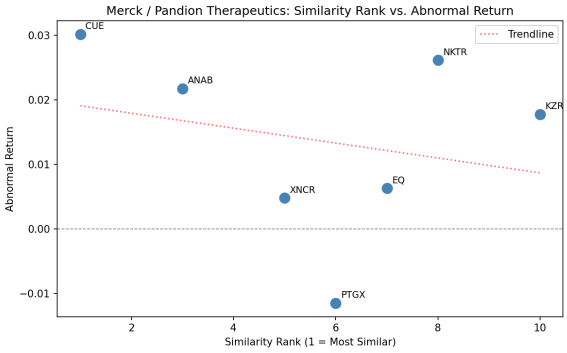

Figure 12: Merck / Pandion Therapeutics Similarity vs. Abnormal Returns

Table 29: Abnormal Returns Pandion Therapeutics

Ticker	Raw Return	Abnormal Return	Direction
ANAB	-0.01888217555300506	0.021744613828727936	Increased
CUE	-0.010471229956700903	0.03015555942503209	Increased
EQ	-0.03435114202800781	0.006275647353725183	Increased
KZR	-0.022887310819052452	0.017739478562680544	Increased
NKTR	-0.014492700531008559	0.026134088850724435	Increased
PTGX	-0.052153288134218864	-0.011526498752485868	Decreased
XNCR	-0.03587632326735664	0.004750466114376357	Increased

Table 30: Final Analysis Pandion Therapeutics

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	CUE	0.94	-0.0104712299567009	0.0301555594250320	Increased	0.028346225859530078
3	ANAB	0.83	-0.0188821755530050	0.0217446138287279	Increased	0.018048029477844157
5	XNCR	0.80	-0.0358763232673566	0.0047504661143763	Increased	0.003800372891501 0403
6	PTGX	0.79	-0.0521532881342188	-0.0115264987524858	Decreased	-0.009105934014463782
7	EQ	0.78	-0.0343511420280078	0.0062756473537251	Increased	0.004895004935905578
8	NKTR	0.76	-0.0144927005310085	0.0261340888507244	Increased	0.019861907526550544
10	KZR	0.72	-0.0228873108190524	0.0177394785626805	Increased	0.012772424565129 96

C.2 Consumer Staples

C.2.1 Amazon / Whole Foods

Amazon/Whole Foods is classified as supporting despite all peers showing negative ARs. The key is the trendline: higher-similarity grocery peers (SFM, NGVC) decline more than lower-similarity peers, which is consistent with the shadow trading mechanism in reverse—the acquisition of a major competitor signals competitive pressure rather than acquisition premium for peers. The negative trendline remains directionally consistent with the hypothesis.

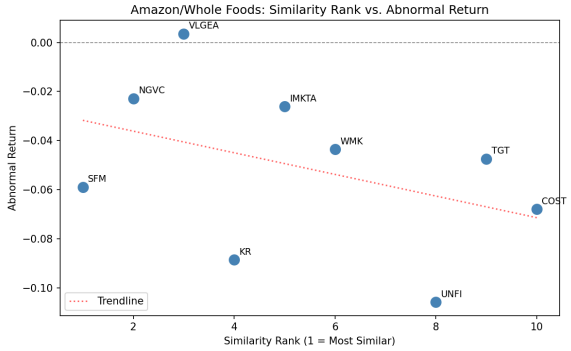

Figure 13: Amazon / Whole Foods Similarity vs. Abnormal Returns

Table 31: Abnormal Returns Whole Foods

Ticker	Raw Return	Abnormal Return	Direction
COST	-0.07192040374481146	-0.06796522838659612	Decreased
IMKTA	-0.03000031493783293	-0.026045139579617585	Decreased
KR	-0.09242663874926776	-0.08847146339105243	Decreased
NGVC	-0.026755852092367225	-0.022800676734151884	Decreased
SFM	-0.06289026951890986	-0.058935094160694514	Decreased
TGT	-0.05138819614229597	-0.047433020784080626	Decreased
UNFI	-0.10968554994595126	-0.10573037458773592	Decreased
VLGEA	-0.00039950769994962914	0.003555667658265714	Increased
WMK	-0.04752578168203486	-0.04357060632381952	Decreased

Table 32: Final Analysis Whole Foods

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	SFM	0.92	-0.0628902695189098	-0.0589350941606945	Decreased	-0.05422028662783894
2	NGVC	0.90	-0.0267558520923672	-0.0228006767341518	Decreased	-0.020520609060736623
3	VLGEA	0.83	-0.0003995076999496	0.0035556676582657	Increased	0.002951204156360 5307
4	KR	0.76	-0.0924266387492677	-0.0884714633910524	Decreased	-0.06723831217719982
5	IMKTA	0.65	-0.0300003149378329	-0.0260451395796175	Decreased	-0.016929340726751375
6	WMK	0.60	-0.0475257816820348	-0.0435706063238195	Decreased	-0.026142363794291697
8	UNFI	0.50	-0.1096855499459512	-0.1057303745877359	Decreased	-0.05286518729386795
9	TGT	0.45	-0.0513881961422959	-0.0474330207840806	Decreased	-0.021344859352836 27
10	COST	0.35	-0.0719204037448114	-0.0679652283865961	Decreased	-0.023787829935308636

C.2.2 Danone / WhiteWave Foods

Danone/WhiteWave Foods is a supporting case. The top-ranked firm (HAIN, +6.9% AR) shows a clear positive reaction, and the negative trendline holds moderately across the cohort. Several mid- and lower-ranked firms show slightly negative but negligible reactions, consistent with a weaker similarity-return relationship in consumer staples compared to biopharmaceuticals.

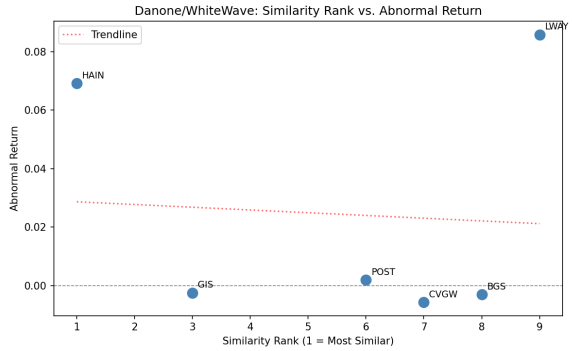

Figure 14: Danone / WhiteWave Foods Similarity vs. Abnormal Returns

Table 33: Abnormal Returns WhiteWave Foods

Ticker	Raw Return	Abnormal Return	Direction
BGS	-0.001464230372955497	-0.0031091979285966117	Decreased
CVGW	-0.004072587271339847	-0.0057175548269809 62	Decreased
DFP	0.00041192300467115514	-0.0012330445509699595	Decreased
GIS	-0.0008324021591602719	-0.0024773697148013866	Decreased
HAIN	0.07085453416477798	0.06920956660913687	Increased
LWAY	0.08752332905866998	0.08587836150302887	Increased
POST	0.0035473528668403583	0.0019023853111992436	Increased

Table 34: Final Analysis WhiteWave Foods

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	HAIN	0.87	0.0708545341647779	0.0692095666091368	Increased	0.060212322949949014
3	GIS	0.76	-0.0008324021591602	-0.0024773697148013	Decreased	-0.001882800983248 9879
6	POST	0.69	0.0035473528668403	0.0019023853111992	Increased	0.001312645864727448
7	CVGW	0.68	-0.0040725872713398	-0.0057175548269809	Decreased	-0.003887937282347 0123
8	BGS	0.64	-0.0014642303729554	-0.0031091979285966	Decreased	-0.001989886674301824
9	LWAY	0.63	0.0875233290586699	0.0858783615030288	Increased	0.054103367746908146

C.2.3 JM Smucker / Hostess

JM Smucker/Hostess is a clean supporting case. All five peers for which data was available show positive abnormal returns, and the top three (FLO, JJSF, THS) show the largest gains, consistent with the shadow trading hypothesis. This is one of the strongest results in the consumer staples subsample.

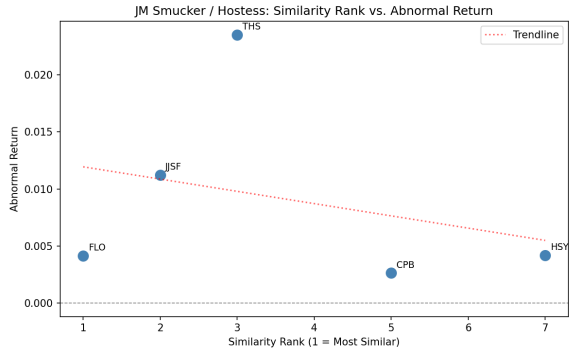

Figure 15: JM Smucker / Hostess Similarity vs. Abnormal Returns

Table 35: Abnormal Returns Hostess

Ticker	Raw Return	Abnormal Return	Direction
CPB	0.012802546888889169	0.00263678770741119	Increased
FLO	0.014310556528568134	0.004144797347090156	Increased
GIS	-0.003930505467713301	-0.01409626464919128	Decreased
HSY	0.01435192480057275	0.004186165619094772	Increased
JJSF	0.021397606391074405	0.011231847209596427	Increased
POST	0.004609905973885611	-0.005555853207592367	Decreased
THS	0.033685659526077175	0.023519900344599196	Increased

Table 36: Final Analysis Hostess

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	FLO	0.92	0.0143105565285681	0.0041447973470901	Increased	0.003813213559322 8924
2	JJSF	0.86	0.0213976063910744	0.0112318472095964	Increased	0.009659388600252904
3	THS	0.82	0.0336856595260771	0.0235199003445991	Increased	0.019286318282571262
5	CPB	0.78	0.0128025468888891	0.0026367877074111	Increased	0.002056694411780658
7	HSY	0.72	0.0143519248005727	0.0041861656190947	Increased	0.003014039245748184

C.2.4 Kroger / Albertsons

Kroger/Albertsons is classified as contradicting. Despite a negative trendline, the top-3 AR direction is mixed $(+, -, +)$, and the overall distribution shows no consistent relationship between similarity rank and return magnitude. Regulatory uncertainty surrounding the proposed merger likely introduced idiosyncratic noise that overwhelmed any fundamental peer signal.

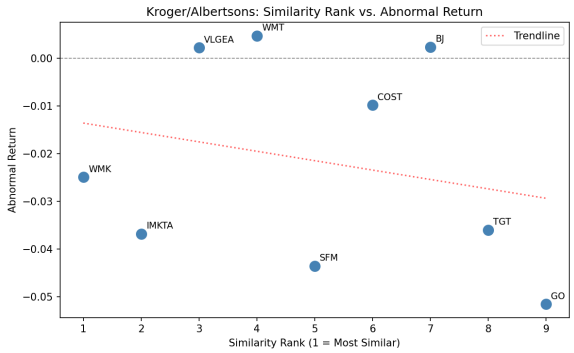

Figure 16: Kroger / Albertsons Similarity vs. Abnormal Returns

Table 37: Abnormal Returns Albertsons

Ticker	Raw Return	Abnormal Return	Direction
BJ	-0.016341231410284245	0.002392864995054235	Increased
COST	-0.02850488629098535	-0.00977078988564687	Decreased
GO	-0.07023213268824618	-0.05149803628290771	Decreased
IMKTA	-0.055493940265775386	-0.0367598438604369	Decreased
SFM	-0.06223925247850891	-0.04350515607317043	Decreased
TGT	-0.05469759121142679	-0.035963494806088306	Decreased
VLGEA	-0.01649676540765968	0.0022373309976788	Increased
WMK	-0.04363704397423403	-0.024902947568895547	Decreased
WMT	-0.013985591718419774	0.0047485046869187066	Increased

Table 38: Final Analysis Albertsons

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	WMK	0.94	-0.0436370439742340	-0.0249029475688955	Decreased	-0.023408770714761766
2	IMKTA	0.89	-0.0554939402657753	-0.0367598438604369	Decreased	-0.03271626103578884
3	VLGEA	0.88	-0.0164967654076596	0.0022373309976788	Increased	0.001968851277957344
4	WMT	0.76	-0.0139855917184197	0.0047485046869187	Increased	0.00360886356205821 2
5	SFM	0.68	-0.0622392524785089	-0.0435051560731704	Decreased	-0.029583506129755875
6	COST	0.65	-0.0285048862909853	-0.0097707898856468	Decreased	-0.00635101342567042
7	BJ	0.63	-0.0163412314102842	0.0023928649950542	Increased	0.001507504946884146
8	TGT	0.58	-0.0546975912114267	-0.0359634948060883	Decreased	-0.020858826987531214
9	GO	0.45	-0.0702321326882461	-0.0514980362829077	Decreased	-0.023174116327308467

C.2.5 Reckitt Benckiser / Mead Johnson Nutrition

Reckitt Benckiser/Mead Johnson is contradicting. The trendline is positive (primarily due to the outlier), but the top-ranked firm (ABT) shows essentially zero abnormal return. The acquisition of an infant nutrition company does not appear to have generated meaningful peer firm reactions, suggesting limited economic linkage within the consumer staples peer group as identified from 10-K text.

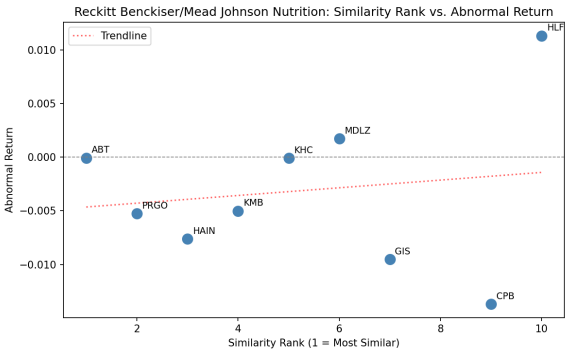

Figure 17: Reckitt Benckiser / Mead Johnson Nutrition Similarity vs. Abnormal Returns

Table 39: Abnormal Returns Mead Johnson Nutrition

Ticker	Raw Return	Abnormal Return	Direction
ABT	0.00470164327620931	-6.045667372484714e-05	Decreased
CPB	-0.008928184876190674	-0.013690284826124832	Decreased
GIS	-0.004757503745379787	-0.009519603695313943	Decreased
HAIN	-0.0028468067310845843	-0.007608906681018742	Decreased
HLF	0.016090798580756894	0.011328698630822736	Increased
KHC	0.004677904408272617	-8.419554166153986e-05	Decreased
KMB	-0.00024270730659042898	-0.005004807256524586	Decreased
MDLZ	0.006487801983811118	0.001725702033876961	Increased
PRGO	-0.0005037224269522213	-0.005265822376886378	Decreased

Table 40: Final Analysis Mead Johnson Nutrition

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	ABT	0.88	0.0047016432762093	-6.045667372484714e-05	Decreased	-5.32018728778654 9e-05
2	PRGO	0.80	-0.0005037224269522	-0.0052658223768863	Decreased	-0.0042126579015090 4
3	HAIN	0.60	-0.0028468067310845	-0.0076089066810187	Decreased	-0.0045653440086112 2
4	KMB	0.55	-0.0002427073065904	-0.0050048072565245	Decreased	-0.0027526439910884755
5	KHC	0.49	0.0046779044082726	-8.419554166153986e-05	Decreased	-4.125581541415453e-05
6	MDLZ	0.42	0.0064878019838111	0.0017257020338769	Increased	0.00072479485422829 8
7	GIS	0.38	-0.0047575037453797	-0.0095196036953139	Decreased	-0.0036174494042192 82
9	CPB	0.30	-0.0089281848761906	-0.0136902848261248	Decreased	-0.0041070854478374 4
10	HLF	0.27	0.0160907985807568	0.0113286986308227	Increased	0.0030587484630322129

C.3 Technology

C.3.1 AMD / Xilinx

AMD/Xilinx is contradicting. Most semiconductor peers show negative ARs and the top three peers are inconsistent. The positive trendline indicates that lower-similarity firms slightly outperformed higher-similarity ones, inverting the shadow trading prediction. Semiconductor M&A events may reflect competitive consolidation rather than acquisition premium contagion.

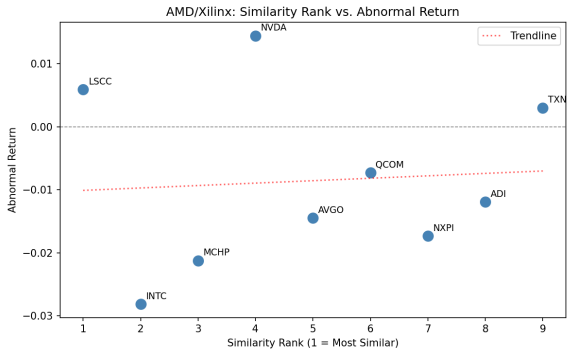

Figure 18: AMD / Xilinx Similarity vs. Abnormal Returns

Table 41: Abnormal Returns Xilinx

Ticker	Raw Return	Abnormal Return	Direction
ADI	-0.006934434890752879	-0.011939806696974281	Decreased
AVGO	-0.009475118768035064	-0.014480490574256464	Decreased
INTC	-0.02311685008217791	-0.028122221888399312	Decreased
LSCC	0.01090167240875663	0.005896300602535229	Increased
MCHP	-0.016221798231416912	-0.021227170037638313	Decreased
NVDA	0.019442558707294864	0.014437186901073464	Increased
NXPI	-0.012306568731181536	-0.017311940537402936	Decreased
QCOM	-0.0022979824548389656	-0.007303354261060367	Decreased
TXN	0.008016157780864525	0.0030107859746431242	Increased

Table 42: Final Analysis Xilinx

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	LSCC	0.86	0.010901672408756 6	0.005896300602535 2	Increased	0.005070818518180272
2	INTC	0.75	-0.023116850082177 9	-0.028122221888399 3	Decreased	-0.021091666416299476
3	MCHP	0.72	-0.016221792314169	-0.021227170037638 3	Decreased	-0.015283562427099575
4	NVDA	0.65	0.019442558707294 8	0.014437186901073 4	Increased	0.00938417148569771
5	AVGO	0.60	-0.009475118768035 0	-0.014480490574256 4	Decreased	-0.00868829434455384
6	QCOM	0.55	-0.002297982454838 9	-0.007303354261060 3	Decreased	-0.004016844843583 1654
7	NXPI	0.50	-0.012306568731181 5	-0.017311940537402 9	Decreased	-0.00865597026870145
8	ADI	0.48	-0.006934434890752 8	-0.011939806696974 2	Decreased	-0.005731107214547616
9	TXN	0.35	0.008016157780864 5	0.003010785974643 1	Increased	0.001053775091125085

C.3.2 Avago / Broadcom

Avago/Broadcom is contradicting. The trendline is positive, and the top two peers (QCOM, SWKS) both show negative ARs, while the fourth and fifth peers show extremely high ARs relative to the cohort. This case, along with AMD/Xilinx, suggests that in semiconductor acquisitions, the announcement-day signal does not propagate to textually similar peers in a manner consistent with the shadow trading hypothesis.

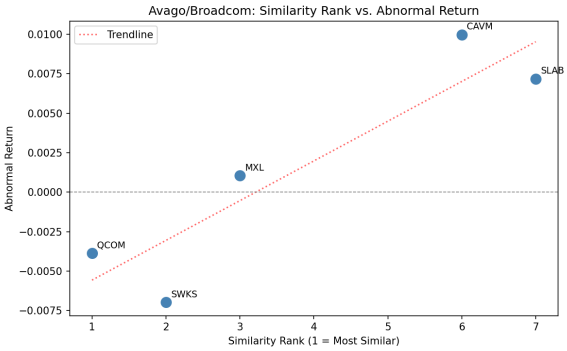

Figure 19: Avago / Broadcom Similarity vs. Abnormal Returns

Table 43: Abnormal Returns Broadcom

Ticker	Raw Return	Abnormal Return	Direction
CAVM	0.007932084987134312	0.009987993662528285	Increased
MXL	-0.0010019314294756854	0.001053977245918287	Increased
QCOM	-0.0059139995490219	-0.003858090873627927	Decreased
SLAB	0.005126374944520318	0.007182283619914291	Increased
SWKS	-0.009021781906765464	-0.006965873231371492	Decreased

Table 44: Final Analysis Broadcom

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	QCOM	0.91	-0.0059139995490219	-0.0038580908736279	Decreased	-0.00351086269500138 9
2	SWKS	0.83	-0.0090217819067654	-0.0069658732313714	Decreased	-0.005781674782038262
3	MXL	0.74	-0.0010019314294756	0.0010539772459182	Increased	0.00077994316197946 81
6	CAVM	0.60	0.0079320849871343	0.0099879936625282	Increased	0.005992796197516919
7	SLAB	0.58	0.0051263749445203	0.0071822836199142	Increased	0.004165724499550235

C.3.3 Cisco / Splunk

Cisco/Splunk is classified as mixed. The top-3 AR direction is $(-, +, +)$. The trendline is negative, and the distribution is bimodal: some peers show meaningful positive reactions (ESTC, DT) while others decline sharply (DDOG, SNOW). The mixed classification reflects genuine ambiguity rather than a clean contradicting signal.

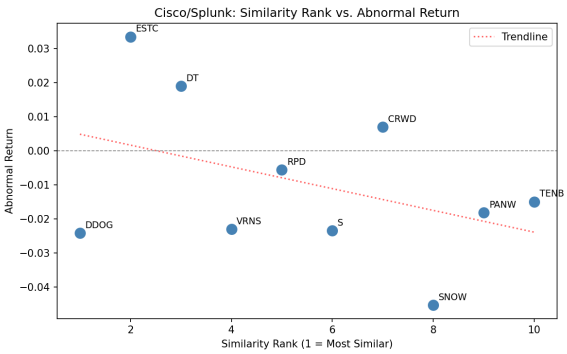

Figure 20: Cisco / Splunk Similarity vs. Abnormal Returns

Table 45: Abnormal Returns Splunk

Ticker	Raw Return	Abnormal Return	Direction
CRWD	-0.008298252580489437	0.006947863262951159	Increased
DDOG	-0.03943514136301994	-0.024189025519579344	Decreased
DT	0.003675636085304054	0.01892175192874465	Increased
ESTC	0.01819609112127831	0.033442206964718904	Increased
PANW	-0.03340416705656666	-0.018158051213126063	Decreased
RPD	-0.02088660657461326	-0.005640490731172662	Decreased
SNOW	-0.060518223589588895	-0.0452721077461483	Decreased
S	-0.038670659767480936	-0.02342454392404034	Decreased
TENB	-0.030199183741365668	-0.014953067897925072	Decreased
VRNS	-0.038279056262499216	-0.02303294041905862	Decreased

Table 46: Final Analysis Splunk

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	DDOG	0.92	-0.03943514136301 99	-0.024189025519579 3	Decreased	-0.022253903478012958
2	ESTC	0.89	0.01819609112127 83	0.033442206964718 9	Increased	0.029763564198599 82
3	DT	0.87	0.00367563608530 40	0.01892175192874 46	Increased	0.016461924178007 802
4	VRNS	0.76	-0.03827905626249 92	-0.02303294041905 86	Decreased	-0.017505034718484 535
5	RPD	0.74	-0.02088660657461 32	-0.00564049073117 26	Decreased	-0.004173963141067 724
6	S	0.73	-0.03867065976748 09	-0.02342454392404 03	Decreased	-0.01709991706454 942
7	CRWD	0.71	-0.00829825258048 94	0.00694786326295 11	Increased	0.004932982916695 281
8	SNOW	0.70	-0.06051822358958 88	-0.04527210774614 83	Decreased	-0.03169047542230 381
9	PANW	0.63	-0.03340416705656 66	-0.01815805121312 60	Decreased	-0.011439572264269 381
10	TENB	0.56	-0.03019918374136 56	-0.01495306789792 50	Decreased	-0.008373718022838001

C.3.4 Microchip / Microsemi

Microchip/Microsemi is also classified as mixed. All identified peers show small negative or positive ARs close to zero, and the trendline is positive but relatively flat when considering the scale. This case is notable for its uniformly low magnitude reactions, suggesting the acquisition was broadly anticipated by the market and generated no meaningful announcement-day peer spillover.

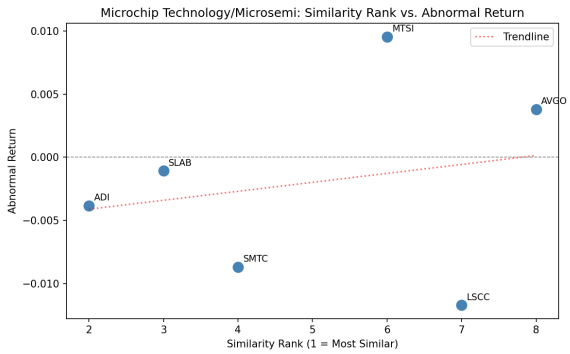

Figure 21: Microchip / Microsemi Similarity vs. Abnormal Returns

Table 47: Abnormal Returns Microsemi

Ticker	Raw Return	Abnormal Return	Direction
ADI	-0.020410672883296514	-0.003834261127835862	Decreased
AVGO	-0.012781175422850264	0.0037952363326103876	Increased
LSCC	-0.028286201301122735	-0.011709789545662083	Decreased
MTSI	-0.007029035494738882	0.00954737626072177	Increased
SLAB	-0.017647075143089906	-0.0010706633876292546	Decreased
SMTC	-0.025260096590624392	-0.008683684835163 74	Decreased

Table 48: Final Analysis Microsemi

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
2	ADI	0.87	-0.02041067288832965	-0.0038342611278358	Decreased	-0.003335807181217146
3	SLAB	0.84	-0.01764707514308899	-0.0010706633876292	Decreased	-0.000899357245608527 9
4	SMTC	0.81	-0.02526009659062 43	-0.0086836848351637	Decreased	-0.007033784716482597
6	MTSI	0.75	-0.0070290354947388	0.0095473762607217	Increased	0.007160532195541275
7	LSCC	0.72	-0.0282862013011227	-0.0117097895456620	Decreased	-0.008431048472876639
8	AVGO	0.68	-0.0127811754228502	0.0037952363326103	Increased	0.002580760706175004

C.3.5 SS&C / DST Systems

SS&C/DST Systems is contradicting. Despite most peers showing positive ARs, the trendline is inverted: lower-similarity peers (MMS, CNDT) show the largest positive reactions, while higher-similarity peers (SEIC, FIS) show negligible returns. This is a case where broad sector sentiment lifted the cohort but failed to concentrate in the most textually similar firms.

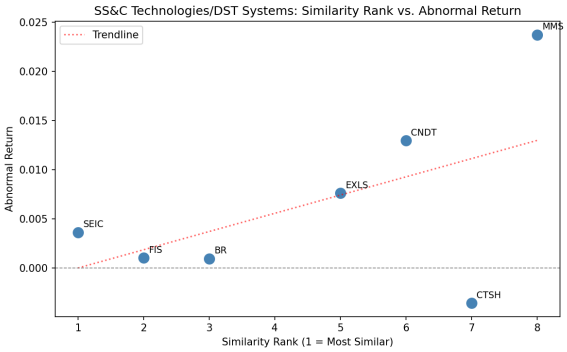

Figure 22: SS&C / DST Systems Similarity vs. Abnormal Returns

Table 49: Abnormal Returns DST Systems

Ticker	Raw Return	Abnormal Return	Direction
BR	0.005054569298770341	0.0009727415307561648	Increased
CNDT	0.017041921734393988	0.012960093966379812	Increased
CTSH	0.0005409325370 22394	-0.003540895230991782	Decreased
DXC	-0.0013830794533293202	-0.005464907221343496	Decreased
EXLS	0.011720661765818713	0.007638833997804536	Increased
FISV	0.005985018298631 03	0.0019031905306168542	Increased
FIS	0.0051402877822758035	0.0010584600142616273	Increased
MMS	0.02782137306097962	0.0237395452929654 44	Increased
SEIC	0.007733366430757337	0.003651538662743161	Increased

Table 50: Final Analysis DST Systems

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	SEIC	0.87	0.0077333664307573	0.0036515386627431	Increased	0.003176838636586497
2	FIS	0.79	0.0051402877822758	0.0010584600142616	Increased	0.000836183411266664
3	BR	0.77	0.0050545692987703	0.0009727415307561	Increased	0.000749010978682 1971
5	EXLS	0.64	0.011720661765818 7	0.0076388339978045	Increased	0.004888853758594 88
6	CNDT	0.54	0.017041921734393 9	0.0129600939663798	Increased	0.006998450741845092
7	CTSH	0.45	0.0005409325370223	-0.0035408952309917	Decreased	-0.001593402853946 2649
8	MMS	0.36	0.02782137306097 96	0.0237395452929654	Increased	0.008546236305467544

C.4 Automobiles

C.4.1 Apollo / Tenneco

Apollo/Tenneco is classified as mixed. The trendline is positive and the top-3 direction is $(-, +, -)$, reflecting the split between DAN (top-ranked, strongly negative AR) and BWA (second-ranked, strongly positive AR). The two most similar firms diverge sharply in their announcement-day reactions, making the case uninterpretable under either hypothesis.

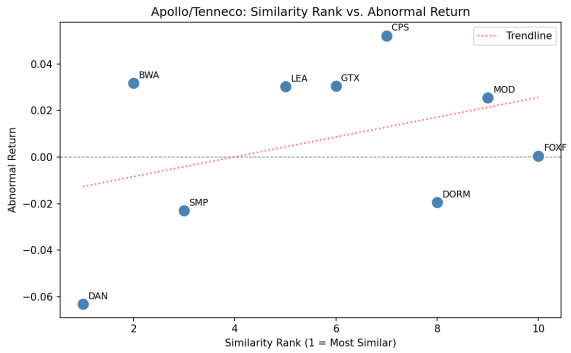

Figure 23: Apollo / Tenneco Similarity vs. Abnormal Returns

Table 51: Abnormal Returns Tenneco

Ticker	Raw Return	Abnormal Return	Direction
SMP	-0.05646922230037127	-0.023050145168253267	Decreased
GTX	-0.00291541343558791	0.03050366369653009	Increased
DCH	0.06953226312827365	0.10295134026039165	Increased
BWA	-0.001651781076413566	0.031767296055704435	Increased
MOD	-0.00795237316535015	0.025466703966767854	Increased
DORM	-0.052984267629254074	-0.01956519049713607	Decreased
FOXF	-0.032919795754358626	0.0004992813777593766	Increased
CPS	0.018749981839210065	0.052169058971328064	Increased
DAN	-0.0966093471696384	-0.0631902700375204	Decreased
LEA	-0.0031639338837505583	0.030255143248367443	Increased

Table 52: Final Analysis Tenneco

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	DAN	0.90	-0.0966093471696384	-0.0631902700375204	Decreased	-0.05687124303376836
2	BWA	0.90	-0.0016517810764135	0.0317672960557044	Increased	0.02859056645013396
3	SMP	0.84	-0.0564692223003712	-0.0230501451682532	Decreased	-0.01936212194133269
5	LEA	0.74	-0.0031639338837505	0.0302551432483674	Increased	0.022388806003791877
6	GTX	0.70	-0.0029154134355879	0.03050366369653	Increased	0.021352564587571
7	CPS	0.69	0.018749981839 21	0.0521690589713 28	Increased	0.035996650690216 32
8	DORM	0.62	-0.052984267629254	-0.019565190497136	Decreased	-0.012130418108224319
9	MOD	0.55	-0.0079523731653501	0.0254667039667678	Increased	0.014006687181722292
10	FOXF	0.50	-0.0329197957543586	0.0004992813777593	Increased	0.00024964068887965

C.4.2 Goodyear / Cooper

Goodyear/Cooper is a strongly supporting case. Eight of nine peers show positive abnormal returns, and the negative trendline holds clearly across the cohort. This is the most uniformly positive result in the Automotive subsample and suggests that in tire and rubber supply chains, textual similarity from 10-K disclosures reliably identifies economically linked firms.

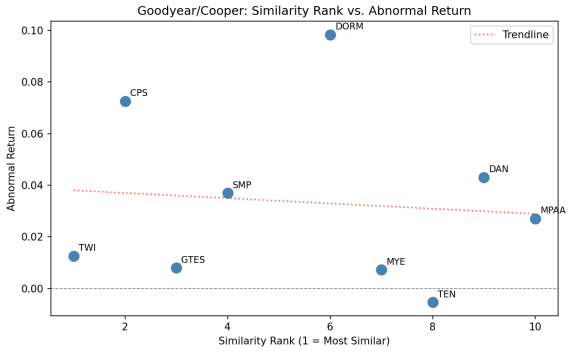

Figure 24: Goodyear / Cooper Similarity vs. Abnormal Returns

Table 53: Abnormal Returns Cooper

Ticker	Raw Return	Abnormal Return	Direction
DORM	0.10152553461907389	0.09833821671472792	Increased
TWI	0.01581509856664611	0.012627780662300143	Increased
DAN	0.04622995957741803	0.04304264167307206	Increased
CPS	0.07574440359430722	0.07255708568996125	Increased
GTES	0.01118308043656393	0.007995762532217964	Increased
MYE	0.010512060818193676	0.007324742913847709	Increased
TEN	-0.0020920732042427684	-0.005279391108588736	Decreased
MPAA	0.030302988601413512	0.027115670697067544	Increased
SMP	0.04016662089667492	0.036979302992328954	Increased

Table 54: Final Analysis Cooper

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	TWI	0.86	0.0158150985666461	0.0126277806623001	Increased	0.010859891369578085
2	CPS	0.77	0.0757444035943072	0.0725570856899612	Increased	0.05586895598127013
3	GTES	0.75	0.011183080436563 9	0.0079957625322179	Increased	0.00599682189916 3425
4	SMP	0.60	0.0401666208966749	0.0369793029923289	Increased	0.0221875817953 9734
6	DORM	0.56	0.1015255346190738	0.0983382167147279	Increased	0.05506940136024763
7	MYE	0.52	0.0105120608181936	0.0073247429138477	Increased	0.00380886631520 08044
8	TEN	0.50	-0.0020920732042427	-0.0052793911085887	Decreased	-0.00263969555429435
9	DAN	0.48	0.046229959577418	0.043042641673072	Increased	0.02066046800307456
10	MPAA	0.40	0.0303029886014135	0.0271156706970675	Increased	0.01084626827882 7

C.4.3 Tenneco / Federal-Mogul

Tenneco/Federal-Mogul is classified as supporting despite a mixed top-3 direction $(-, +, -)$. The classification reflects the overall pattern: the negative trendline is driven by the highest-similarity firms showing moderate positive reactions (BWA, SMP) while lower-similarity firms show larger negative reactions. The holistic analysis supports the hypothesis even though the top-ranked firm (DAN) shows a small negative AR.

Table 55: Abnormal Returns Federal

Ticker	Raw Return	Abnormal Return	Direction
BWA	0.027923368972255665	0.012623729497841695	Increased
CMI	0.03463282681896851	0.01933318734455454	Increased
DORM	0.01464721921591628	-0.000652420258497691	Decreased
LKQ	0.0005245657554827761	-0.014775073718931194	Decreased
SMP	0.017457329721931526	0.0021576902475175554	Increased
DAN	0.012209253473614552	-0.0030903860007994183	Decreased
DCH	0.0006250143051147461	-0.014674625169299224	Decreased
CPS	0.014450176569453304	-0.0008494629049606659	Decreased
MOD	0.004926127351136306	-0.010373512123277664	Decreased

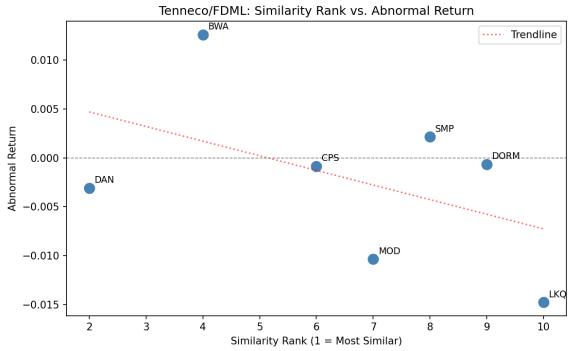

Figure 25: Tenneco / Federal-Mogul Similarity vs. Abnormal Returns

Table 56: Final Analysis Federal

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
2	DAN	0.91	0.0122092534736145	-0.0030903860007994	Decreased	-0.00281225126072 74544
4	BWA	0.85	0.0279233689722556	0.0126237294978416	Increased	0.01073017007316536
6	CPS	0.82	0.0144501765694533	-0.0008494629049606	Decreased	-0.00069655958206 76919
7	MOD	0.74	0.0049261273511363	-0.0103735121232776	Decreased	-0.00767639897122 5424
8	SMP	0.60	0.0174573297219315	0.0021576902475175	Increased	0.00129461414851 04998
9	DORM	0.51	0.0146472192159162	-0.0006524202584976	Decreased	-0.00033273433183 3776
10	LKQ	0.37	0.0005245657554827	-0.0147750737189311	Decreased	-0.00546677727600 4507

C.4.4 Warner / Delphi

Warner/Delphi is classified as mixed. The trendline is positive and the top-3 direction is $(-, \approx 0, +)$, with the top-ranked firm (TEN) showing a negative AR and the third-ranked (DAN) showing a positive one. The lack of a consistent pattern across either direction prevents a clear supporting or contradicting classification.

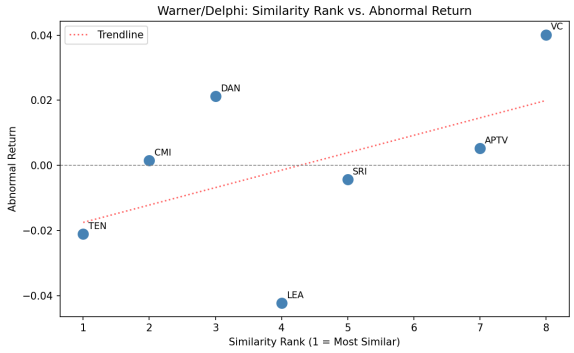

Figure 26: Warner / Delphi Similarity vs. Abnormal Returns

Table 57: Abnormal Returns Delphi

Ticker	Raw Return	Abnormal Return	Direction
VC	0.04662139122157009	0.04015559590695077	Increased
DCH	0.018104373810255786	0.011638578495636463	Increased
SMP	-0.0027740168695917916	-0.009239812184211115	Decreased
APTV	0.011738143800527612	0.00527234848590829	Increased
TEN	-0.014577270889797867	-0.02104306620441719	Decreased
ST	0.010699377862144956	0.004233582547525633	Increased
SRI	0.0021413085282714493	-0.004324486786347873	Decreased
DAN	0.027690593095965572	0.021224797781346248	Increased
CMI	0.007950362187841791	0.0014845668732224688	Increased
LEA	-0.03579390798828836	-0.04225970330290768	Decreased

Table 58: Final Analysis Delphi

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	TEN	0.94	-0.0145772708897978	-0.0210430662044171	Decreased	-0.019780482232152074
2	CMI	0.87	0.0079503621878417	0.0014845668732224	Increased	0.00129157317970348 8
3	DAN	0.85	0.0276905930959655	0.0212247977813462	Increased	0.018041078114144 27
4	LEA	0.81	-0.0357939079882883	-0.0422597033029076	Decreased	-0.034230359675355156
5	SRI	0.78	0.0021413085282714	-0.0043244867863478	Decreased	-0.003373099693351284
7	APTV	0.73	0.0117381438005276	0.0052723484859082	Increased	0.00384881439471298 57
8	VC	0.65	0.0466213912215 7	0.0401555959069507	Increased	0.026101137339517955

C.4.5 Warner / Remy

Warner/Remy is a strongly supporting case. Seven of eight peers show positive abnormal returns, and the negative trendline holds across the cohort. The top-ranked firm (MPAA, $+3.9\%$ AR) and second-ranked firm (SMP, $+1.0\%$ AR) both show positive reactions consistent with the shadow trading hypothesis.

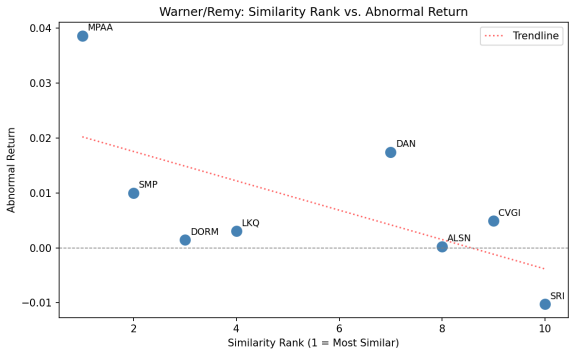

Figure 27: Warner / Remy Similarity vs. Abnormal Returns

Table 59: Abnormal Returns Remy

Ticker	Raw Return	Abnormal Return	Direction
ALSN	0.011316857070665666	0.00024048661547815438	Increased
DAN	0.02852791101708753	0.017451540561900018	Increased
LKQ	0.014177439899237226	0.003101069444049715	Increased
MPAA	0.04971055644378255	0.03863418598859504	Increased
SRI	0.0008417700997898144	-0.010234600355397697	Decreased
DORM	0.01250525832727829	0.0014288878720907783	Increased
CVGI	0.01597442507608828	0.004898054620900767	Increased
SMP	0.021061745727400483	0.009985375272212972	Increased

Table 60: Final Analysis Remy

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	MPAA	0.94	0.049710556443782 5	0.0386341859885950	Increased	0.0363161348292793
2	SMP	0.87	0.0210617457274004	0.0099853752722129	Increased	0.008687276486825221
3	DORM	0.71	0.0125052583272782	0.0014288878720907	Increased	0.00101451038918439 69
4	LKQ	0.70	0.0141774398992372	0.0031010694440497	Increased	0.00217074861083479
7	DAN	0.59	0.0285279110170875	0.0174515405619000	Increased	0.0102964089315210
8	ALSN	0.56	0.0113168570706656	0.0002404866154781	Increased	0.000134672504667736 03
9	CVGI	0.48	0.0159744250760882	0.0048980546209007	Increased	0.00235106621803233 63
10	SRI	0.48	0.0008417700997898	-0.0102346003553976	Decreased	-0.004912608170590848

D Finance

D.0.1 BB&T / SunTrust

BB&T/SunTrust presents a nuanced case. All peer firms show positive abnormal returns on announcement day, which is consistent with the weak form of the shadow trading hypothesis—that MNPI about one bank would affect similar banks generally. However, the positive trendline indicates that the magnitude of those gains completely decoupled from semantic similarity rank: lower-similarity peers outperformed higher-similarity peers. This is the sector-wide contagion failure mode: the merger signaled regulatory approval of large-bank consolidation broadly, lifting all regional bank peers regardless of their textual similarity to SunTrust. Because our test evaluates the strong form of the hypothesis—whether similarity rank predicts return magnitude, which is what the SEC needs to identify specific trading targets—this case is classified as contradicting, even though all peers moved in the predicted direction.

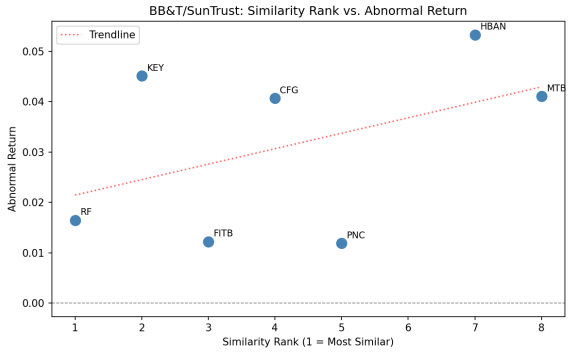

Figure 28: BB&T / SunTrust Similarity vs. Abnormal Returns

Table 61: Abnormal Returns SunTrust

Ticker	Raw Return	Abnormal Return	Direction
MTB	0.031114360274174226	0.041091210736052076	Increased
RF	0.006489173391255033	0.016466023853132884	Increased
FITB	0.002245720098404834	0.012222570560282685	Increased
CFG	0.030707060645199617	0.04068391110707747	Increased
KEY	0.03517316578329104	0.04515001624516889	Increased
PNC	0.0018759089574348529	0.011852759419312704	Increased
PNFP	0.02699211079993078	0.03696896126180863	Increased
FHN	0.02747994594638716	0.037456796408265006	Increased
HBAN	0.0433482958726529	0.053325146334530754	Increased

Table 62: Final Analysis SunTrust

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	RF	0.91	0.0064891733912550	0.0164660238531328	Increased	0.014984081706350849
2	KEY	0.88	0.0351731657832910	0.0451500162451688	Increased	0.03973201429574855
3	FITB	0.87	0.0022457200984048	0.0122225705602826	Increased	0.010633636387445862
4	CFG	0.86	0.0307070606451996	0.0406839111070774	Increased	0.03498816355208656
5	PNC	0.85	0.0018759089574348	0.0118527594193127	Increased	0.010074845506415795
7	HBAN	0.78	0.0433482958726529	0.0533251463345307	Increased	0.041593614140933946
8	MTB	0.74	0.0311143602741742	0.0410912107360520	Increased	0.03040749594467848

D.0.2 KeyCorp / First Niagara

KeyCorp/First Niagara is classified as strongly supporting. The top-ranked firm (CFG) shows a positive AR while all lower-ranked firms show negative reactions, producing a clear negative trendline that

aligns with the shadow trading hypothesis. This is the cleanest supporting case in the Finance subsample.

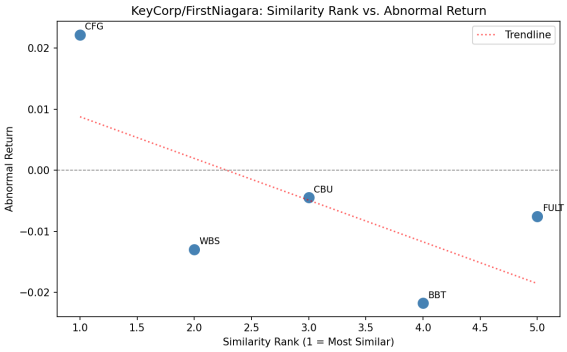

Figure 29: KeyCorp / First Niagara Similarity vs. Abnormal Returns

Table 63: Abnormal Returns First Niagara

Ticker	Raw Return	Abnormal Return	Direction
BBT	-0.03606357708124112	-0.02173759961246282	Decreased
CFG	0.007880570014119051	0.02220654748289735	Increased
WBS	-0.027267811913015836	-0.012941834444237537	Decreased
FNB	-0.023912975220883568	-0.009586997752105269	Decreased
CBU	-0.0187772386177642	-0.0044512611489859	Decreased
VLY	-0.018691419065318966	-0.004365441596540667	Decreased
FULT	-0.021865289671713632	-0.007539312202935333	Decreased
MTB	-0.025768358038338268	-0.011442380569559969	Decreased

Table 64: Final Analysis First Niagara

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
nheight1	CFG	0.92	0.007880570014119 0	0.0222065474828973	Increased	0.020430023684265514
n2	WBS	0.85	-0.0272678119130158	-0.012941834444237 5	Decreased	-0.011000559277601876
n3	CBU	0.79	-0.0187772386177642	-0.0044512611489859	Decreased	-0.0035164963076988608
n4	BBT	0.76	-0.0360635770812411	-0.0217375996124628	Decreased	-0.01652057570547173
n5	FULT	0.71	-0.021865289671713 6	-0.0075393122029353	Decreased	-0.005352911664084063
nheight						

D.0.3 People's / United

People's/United is classified as contradicting. The trendline is positive and most peers show negative or near-zero ARs. The one strongly positive outlier (MRBK, $+4.4\%$ AR) ranks last in similarity, directly inverting the expected pattern.

Table 65: Abnormal Returns United

Ticker	Raw Return	Abnormal Return	Direction
NBTB	-0.019291550895895102	-0.013641987240980431	Decreased
WASH	-0.007836123577671349	-0.0021865599227566777	Decreased
BBT	-0.0287973922743519	-0.02314782861943723	Decreased
MRBK	0.038690462622618854	0.04434002627753353	Increased
WNEB	-0.0032019992606314137	0.0024475643942832572	Increased
FFIC	-0.010700167840185815	-0.005050604185271144	Decreased
INDB	-0.014858426109558339	-0.009208862454643668	Decreased

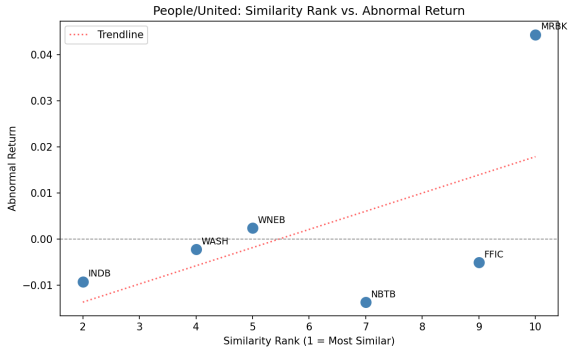

Figure 30: People's / United Similarity vs. Abnormal Returns

Table 66: Final Analysis United

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
2	INDB	0.88	-0.0148584261095583	-0.0092088624546436	Decreased	-0.008103798960086369
4	WASH	0.81	-0.0078361235776713	-0.0021865599227566	Decreased	-0.0017711135374328462
5	WNEB	0.73	-0.0032019992606314	0.0024475643942832	Increased	0.001786722007826736
7	NBTB	0.66	-0.0192915508958951	-0.0136419872409804	Decreased	-0.009003711579047065
9	FFIC	0.52	-0.0107001678401858	-0.0050506041852711	Decreased	-0.002626314176340972
10	MRBK	0.42	0.0386904626226188	0.0443400262775335	Increased	0.01862281103656407

D.0.4 Sterling / Astoria

Sterling/Astoria is classified as supporting. The top-ranked firm (FFIC, $+0.79\%$ AR) shows a positive reaction while lower-ranked firms show negative reactions, producing a modest negative trendline. The magnitude of reactions is small throughout, reflecting the limited market impact of a regional bank acquisition.

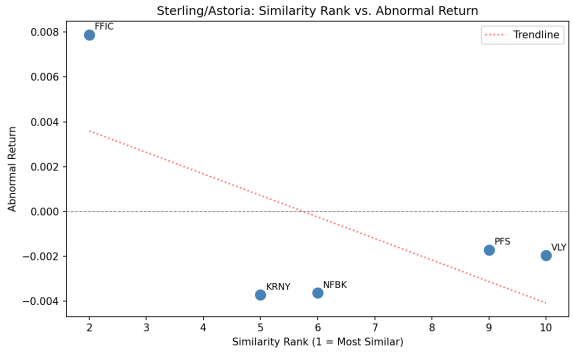

Figure 31: Sterling / Astoria Similarity vs. Abnormal Returns

Table 67: Abnormal Returns Astoria

Ticker	Raw Return	Abnormal Return	Direction
VLY	-0.004796628053415628	-0.0019719103377265033	Decreased
CNOB	-0.008080841000498168	-0.005256123284809043	Decreased
FFIC	0.005063261483582687	0.00788797919927181	Increased
FLG	-0.007487942790061821	-0.004663225074372696	Decreased
NFBK	-0.006468928281572891	-0.003644210565883767	Decreased
CFG	-0.007839123578677706	-0.005014405862988582	Decreased
DCOM	0.002816869865432588	0.005641587581121712	Increased
PFS	-0.004540318237324739	-0.0017156005216356143	Decreased
KRNY	-0.006557784922778603	-0.003733067207089479	Decreased

Table 68: Final Analysis Astoria

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
2	FFIC	0.93	0.0050632614835826	0.0078879791992718	Increased	0.007335820655322775
5	KRNY	0.83	-0.0065577849227786	-0.0037330672070894	Decreased	-0.003098445781884202
6	NFBK	0.82	-0.0064689282815728	-0.0036442105658837	Decreased	-0.002988252664024633 7
9	PFS	0.77	-0.0045403182373247	-0.0017156005216356	Decreased	-0.001321012401659412
10	VLY	0.70	-0.0047966280534156	-0.0019719103377265	Decreased	-0.001380337236408549 8

D.0.5 WSFS / Beneficial

WSFS/Beneficial is a supporting case. All peers show positive abnormal returns, and the two top-ranked firms (UVSP and PGC) show about the largest positive ARs. The negative trendline holds across the cohort, making this one of the more consistent supporting cases in the Finance subsample.

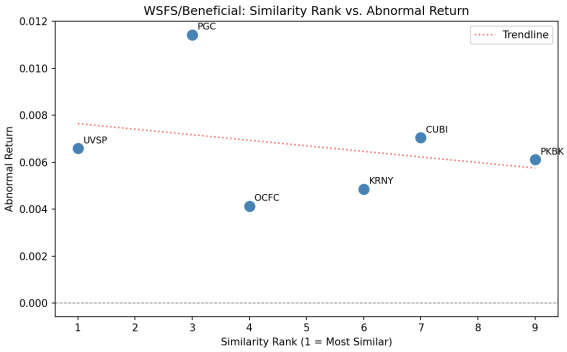

Figure 32: WSFS / Beneficial Similarity vs. Abnormal Returns

Table 69: Abnormal Returns Beneficial

Ticker	Raw Return	Abnormal Return	Direction
KRNY	0.007326169006966355	0.004861293886339441	Increased
UVSP	0.009057941329852869	0.006593066209225955	Increased
FRBA	0.017985763611004544	0.015520888490377631	Increased
PKBK	0.00858361590995938	0.006118740789332466	Increased
OCFC	0.006599403266681618	0.004134528146054704	Increased
CUBI	0.009516558646504194	0.00705168352587728	Increased
FRBKQ	0.0	-0.002464875120626914	Decreased
FULT	0.017192299737995555	0.014727424617368642	Increased
WSFS	-0.08130069889002943	-0.08376557401065635	Decreased
PGC	0.013901568834827368	0.011436693714200453	Increased

Table 70: Final Analysis Beneficial

Rank	Ticker	Sim. Score	Raw Return	Abnormal Return	Direction	Shadow Signal
1	UVSP	0.94	0.0090579413298528	0.0065930662092259	Increased	0.006197482236672346
3	PGC	0.87	0.0139015688348273	0.0114366937142004	Increased	0.009949923531354347
4	OCFC	0.84	0.0065994032666816	0.0041345281460547	Increased	0.003473003642685947 5
6	KRNY	0.70	0.0073261690069663	0.0048612938863394	Increased	0.00340290572043758
7	CUBI	0.65	0.0095165586465041	0.0070516835258772	Increased	0.00458359429182018
9	PKBK	0.00	0.0085836159099593	0.0061187407893324	Increased	0.0