Report
Master Semester Project

Insider Trading Intensity and Abnormal Returns Around 8-K Filings

BY MATTHIAS WYSS (SCIPER 329884)
AND WILLIAM JALLOT (SCIPER 341540)
DATA SCIENCE, MA3
12 ECTS

SUPERVISED BY
PROF. PIERRE COLLIN-DUFRESNE

Contents

AUTUMN SEMESTER - 2025

1 Introduction

It is well established that some traders operate on material non-public information, notable cases include Martha Stewart and Albert H. Wiggin. However, detecting such informed trading remains a challenging task. Informed Trading Intensity (ITI) [1] introduces a novel machine-learning-based metric designed to quantify the presence of informed trading in financial markets.
In this project, we extend the analysis of ITI by focusing on its interaction with corporate disclosures through Form 8-K filings. We study the dynamics of ITI around both the report date and the public filing date, allowing us to separate pre-disclosure informed trading from post-disclosure market reactions. Furthermore, we decompose 8-K filings by item composition to assess whether specific types of corporate events are associated with distinct ITI and return patterns. Finally, we combine ITI with textual sentiment extracted from Item 8.01 disclosures using NLP techniques, enabling a joint analysis of trading-based and text-based information signals.

2 Related Work

2.1 Informative Trading Intensity (ITI)

Informative Trading Intensity (ITI) [1] is a data-driven measure of realized informed trading proposed by Bogousslavsky, Fos, and Muravyev (2024). The authors train a Gradient Boosted Trees model on observed informed trades, specifically Schedule 13D filings by activist investors between 1994 and 2018, to identify days characterized by informed trading activity. The model uses 41 daily variables related to liquidity, returns, volatility, and trading volume, enabling it to capture nonlinearities and interactions between market microstructure features that traditional econometric measures fail to represent. Importantly, the authors emphasize that ITI is designed to measure informed trading rather than being a liquidity measure.
Unlike classical theory-based proxies such as Kyle's λ [2], the bid-ask spread of Glosten and Milgrom [3], or the Probability of Informed Trading (PIN) of Easley and O'Hara [4], ITI relies on supervised machine learning to learn directly from empirical data on informed trading. Once trained on Schedule 13D data, the model is extrapolated to the full cross-section of U.S. common stocks from 1993 to 2019, computing for each stock-day a probability between 0 and 1 that reflects the intensity of informed trading.
Empirically, ITI increases prior to major informational events such as earnings announcements, mergers and acquisitions, and unscheduled news releases. Furthermore, higher ITI values are associated with subsequent abnormal returns, suggesting that realized informed trading contains predictive information for asset prices.
Overall, ITI provides a robust and interpretable empirical measure of informed trading intensity, demonstrating how machine learning techniques can uncover complex, nonlinear relations between trading activity, liquidity, and information asymmetry in financial markets.

2.2 Bloomberg-Reuters Dataset

The Bloomberg-Reuters dataset is a large-scale financial news corpus composed of 450,341 articles from Bloomberg and 109,110 articles from Reuters [5]. It primarily consists of firm-specific newswire articles covering a broad range of corporate events, including earnings announcements, mergers and acquisitions, management changes, guidance updates, and legal developments.

2.3 FNSPID Dataset

The FNSPID dataset [6] stands for *Financial News and Stock Price Integration Dataset.* It is a large-scale financial dataset that integrates quantitative and qualitative information to improve stock market prediction tasks. FNSPID contains approximately 29.7 million stock price records and 15.7 million financial news articles covering 4,775 S&P 500 companies from 1999 to 2023, collected from 4 stock financial news websites. In addition to raw news content and timestamps, the dataset also includes the stock symbol associated with the news, which facilitates the alignment of textual information with corresponding market data.

2.4 GDELT Project

The GDELT (*Global Database of Events, Language, and Tone*) project is a global dataset that collects news from around the world, including financial news. Since 2015, a new GKG (Global Knowledge Graph) file is uploaded every 15 minutes. Each entry includes the timestamp, source link, type of event, involved actors, locations, and tone, providing rich structured information for downstream analyses.
GDELT is particularly useful for financial applications because it enables:

- Tracking global events that may influence asset prices.
- Measuring media sentiment using the `tone` field.
- Constructing time-based features for predictive models, such as event-driven trading strategies or risk indices.

The frequent updates and structured nature of the dataset make it suitable for real-time or high-frequency analysis of news impact on markets.

3 Data Overview

To conduct our analysis, we require a dataset of corporate news with precise timestamps and firm-level coverage. We initially explored several existing news datasets, including Bloomberg-Reuters [5], FNSPID [6] and GDELT [7]. While these datasets offer broad coverage of financial and general news, they were ultimately not retained for our main analysis due to various limitations.
In addition, we explored the construction of a dataset based on Schedule 13D filings with the objective of replicating the ITI measure using more recent data. This approach was also not pursued further, as a full implementation of ITI requires access to intraday market data that are not available in our data environment. Details on the construction and limitations of these alternative data sources are provided in Appendix A.

3.1 ITI measure

Since our institution does not have access to the libraries used to compute this metric, we had to use the measure computed directly in the ITI paper [1], which spans from 1993 to 2019. We focus primarily on ITI(13D) because it:

- increases prior to earnings announcements,
- rises ahead of unscheduled news and merger-and-acquisition events,
- predicts the magnitude of announcement-day abnormal returns, and
- is associated with less return reversal, consistent with informed price discovery,

making it a particularly suitable measure for our analysis. The ITI dataset used in this study, which includes the date, the ITI(13D) measure, and the firm identifier (PERMNO), is publicly available in our GitHub repository [8] at `data/raw/ITIs.csv`.

3.2 8-K Filings

Based on our preliminary experiments with the FNSPID dataset and the GDELT project, we concluded that publicly available news datasets are often noisy and require substantial preprocessing, which can introduce additional noise into the analysis. To address these limitations, we rely on Form 8-K filings as a proxy for corporate news.

Form 8-K is a mandatory disclosure that publicly traded firms must file with the Securities and Exchange Commission (SEC) to report significant or material events that shareholders are required to be informed about. Each filing contains two relevant dates: the report date, corresponding to the occurrence of the event, and the filing date, which indicates when the disclosure becomes publicly available. Firms are required to file the report within four business days following the event. Filings are organized into standardized items, each representing a specific type of corporate event. A single 8-K filing may include one or multiple items. A complete list of all 8-K items and their descriptions can be found in Appendix B. Figure 1 provides an illustrative example of an Item 5.02 report within a Form 8-K filing.

Item 5.02 Departure of Directors or Certain Officers; Election of Directors; Appointment of Certain Officers; Compensatory Arrangements of Certain Officers

The Board of Directors (the "Board") of Apple Inc. (the "Company" or "Apple") previously adopted, subject to shareholder approval, the Apple Inc. 2022 Employee Stock Plan (the "2022 Plan"). Apple's shareholders approved the 2022 Plan at the Company's Annual Meeting of Shareholders held on March 4, 2022 (the "Annual Meeting"). Apple's grant authority under the Apple Inc. 2014 Employee Stock Plan (the "2014 Plan") will terminate after the 2022 Plan is registered on Form S-8. The 2022 Plan, which became effective upon shareholder approval, permits the granting of stock options, stock appreciation rights, stock grants and restricted stock units. Employees and consultants of Apple and its subsidiaries are eligible to participate in the 2022 Plan.

The maximum number of shares that may be issued or transferred pursuant to awards under the 2022 Plan will equal:

- 510 million shares, plus
- the number of shares available for new award grants under the 2014 Plan on the date of the Annual Meeting, plus
- the number of any shares subject to stock options granted under the 2014 Plan and outstanding as of the date of the Annual Meeting which expire or terminate after the Annual Meeting, plus
- two times the number of shares subject to restricted stock units ("RSUs") or restricted awards granted under the 2014 Plan that are outstanding as of the date of the Annual Meeting that are forfeited or terminated or with respect to which shares are withheld to satisfy tax withholding obligations after the date of the Annual Meeting.

The maximum number of shares that may be issued or transferred pursuant to awards under the 2022 Plan as a result of applying the share limit formula described above will not exceed 1,274,374,682 shares.

Shares issued with respect to full-value awards (RSUs or restricted stock awards) granted under the 2022 Plan are counted against the 2022 Plan's aggregate share limit as two shares for every one share actually issued in connection with the full-value award. The 2022 Plan also includes other rules for counting shares against the share limits.

The foregoing brief description is qualified in its entirety by the text of the 2022 Plan, a copy of which is filed as Exhibit 10.1 hereto and incorporated herein by reference.

Grants under the 2022 Plan may be evidenced by entry into the Restricted Stock Unit Award Agreement and the Performance Award Agreement under the 2022 Plan, forms of which are filed as Exhibits 10.2 and 10.3 hereto and incorporated herein by reference.

Figure 1: Illustrative example of an Item 5.02 report within a Form 8-K filing

For our analysis, we constructed a comprehensive dataset of all 8-K filings and merged it with the ITI dataset to produce the final 8-K ITI dataset. The main steps of our preprocessing pipeline were as follows:

1. Download the full set of corporate submissions from the SEC website.
2. Filter the dataset to retain only Form 8-K filings.
3. Map firm identifiers from CIK to PERMNO.
4. Merge the filings with the ITI dataset based on the report date.
5. Compute the time gap between the report date and the filing date.
6. Clean the sample by retaining only filings submitted after 2001, to ensure consistency with the revised 8-K reporting format. Outliers were further removed, including filings with more than

four days between the report and filing dates, as well as cases where the report date occurs after the filing date.

The resulting 8-K ITI dataset is publicly available in our GitHub repository [8] at `data/merged/8k_iti.parquet`. It contains 99,384 unique events. For each trading day, it includes the firm identifier (PERMNO), the ITI(13D) measure, and if an 8-K filing occurred on that day, the report date, the filing date, the filing URL, and the set of reported 8-K items.

One of the main challenges when working with Form 8-K filings is the prevalence of tables, cross-references, and exhibit mentions, which complicates text parsing and limits the applicability of standard NLP techniques. To address this issue, we restrict the remainder of our analysis to Item 8.01 disclosures, resulting in 13,359 samples. Item 8.01 corresponds to Other Events, which is used by firms to disclose material information not otherwise required to be reported under specific Form 8-K items. These disclosures primarily consist of free-form textual content and typically exclude highly structured elements, making them particularly well suited to sentiment analysis.
Focusing on Item 8.01 allows us to preserve the informational content most relevant for investors while ensuring reliable text preprocessing. We therefore parse the text of each Item 8.01 disclosure and compute sentiment scores using the FinBERT model [9] under three alternative approaches.
In the first approach, we apply FinBERT directly to the raw text of the disclosure. FinBERT outputs probabilities for three sentiment classes (positive, negative, and neutral), each ranging between 0 and 1. We define the sentiment score as:

$$sentiment\_score = positive\_score - negative\_score,$$

which lies in the interval $[-1, 1]$, where values close to 1 indicate strongly positive sentiment, values close to -1 indicate strongly negative sentiment, and values near zero correspond to neutral tone.
A limitation of this approach is that FinBERT processes inputs of at most 512 tokens, leading to truncation of longer disclosures and sentiment scores that reflect only the beginning of the text.
To mitigate this issue, our second approach divides each disclosure into consecutive 512-token chunks. FinBERT is applied to each chunk separately, and the final sentiment score is obtained by averaging the sentiment scores across all chunks.
Finally, we consider a third approach that leverages large language models for summarization. We first prompt the Mistral model [10] to generate a concise and factual summary of the full Item 8.01 disclosure. The summary is then split into 512-token segments, which are individually processed by FinBERT. The final sentiment score is computed as the mean of the segment-level sentiment scores. The prompt used for summarization is reported in Appendix C.
All datasets derived from the three sentiment extraction approaches are publicly available in our GitHub repository [8]. Specifically:

- Simple FinBERT: `data/preprocessed/item_8_01_finbert.parquet`
- FinBERT Mean (chunk-based): `data/preprocessed/item_8_01_finbert_mean_chunk.parquet`
- Mistral FinBERT: `data/preprocessed/item_8_01_mistral_summary_to_finbert.parquet`

3.3 Returns

We downloaded daily CRSP stock data for U.S. common stocks listed on NYSE, AMEX, or NASDAQ over 2001-2024, using Wharton Research Data Services API. Specifically, we query the CRSP daily stock file (`crsp.dsf`) and merge it with `crsp.stocknames` to attach firm identifiers (e.g., `comnam`, `ticker`) and listing information. We retrieve the required returns within the event study.

4 Analysis

4.1 Event Study Specification

To assess the impact of a Form 8-K filing, we conducted an event study structured as follows. Given an observation window of fixed length (500 observations in our case) and a specified gap between the event date and the estimation window, we estimated a linear regression model in which excess returns were regressed on the Fama–French factors and the momentum factor.

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{MKT}\,(R_{MKT,t} - R_{f,t}) + \beta_{SMB}\,SMB_t + \beta_{HML}\,HML_t + \beta_{MOM}\,MOM_t + \varepsilon_{i,t}.$$

Definition of each component.

- $R_{i,t}$: Return of asset i at time t.
- $R_{f,t}$: Risk-free rate at time t. The term $R_{i,t} - R_{f,t}$ denotes the asset's excess return.
- α_i: Asset-specific intercept capturing the average abnormal return unexplained by the factors.
- β_{MKT}: Loading on the market excess return factor $(R_{MKT,t} - R_{f,t})$.
- β_{SMB}: Loading on the size factor (*Small Minus Big*).
- β_{HML}: Loading on the value factor (*High Minus Low*).
- β_{MOM}: Loading on the momentum factor.
- $\varepsilon_{i,t}$: Error term capturing idiosyncratic risk.

Interpretation of the Factors.

- **SMB (Small Minus Big)**: It is constructed as the return of a portfolio long in small-cap firms and short in large-cap firms.
- **HML (High Minus Low)**: It corresponds to the return difference between firms with high book-to-market ratios and low book-to-market ratios.
- **MOM (Momentum)**: It represents the return of a portfolio long in past winners and short in past losers.

Abnormal Return (AR). Once the factor loadings are estimated over the observation window, the abnormal return at time t is defined as:

$$AR_{i,t} = (R_{i,t} - R_{f,t}) - \widehat{R}_{i,t},$$

where $\widehat{R}_{i,t}$ is the expected excess return predicted by the factor model:

$$\widehat{R}_{i,t} = \widehat{\alpha}_i + \widehat{\beta}_{MKT}\,(R_{MKT,t} - R_{f,t}) + \widehat{\beta}_{SMB}\,SMB_t + \widehat{\beta}_{HML}\,HML_t + \widehat{\beta}_{MOM}\,MOM_t.$$

Abnormal ITI (AI) All the same steps as in AR but returns are replaced by ITI measure, and we keep the same estimation model.

Cumulative Abnormal Return (CAR). For an event window from τ_1 to τ_2, the cumulative abnormal return is defined as:

$$CAR_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}.$$

A positive CAR indicates that the asset generated returns above what would be expected, suggesting a favorable market reaction to the event. Conversely, a negative CAR reflects underperformance relative to the model-implied benchmark, indicating a negative market reaction. The magnitude of the CAR provides an estimate of the economic impact of the event on the asset's value .

Cumulative Average Abnormal Return (CAAR) The Cumulative Average Abnormal Return (CAAR) measures the aggregated abnormal performance across all firms over the event window, providing a summary of the average market reaction to the event.

Placebo (Randomized) Events. To validate the robustness of our event study results, we also perform a placebo test based on randomly generated events. For each actual event, identified by a pair ($\text{permno}_i, \text{event\_date}_i$), we draw n_{rand} random trading dates for the *same* stock (same permno) from the CRSP daily database. In order to avoid contamination by true information releases, we exclude from the sampling universe any trading day that falls within a symmetric buffer window of $\pm B$ calendar days around any real event date of that permno. In our implementation, we set $B = 60$ days and $n_{\text{rand}} = 1$ placebo date per real event.
The resulting set of placebo events is then fed through the exact same event-study pipeline as the true events (estimation window, computation of abnormal returns, and CARs). If the model is well specified and our results are not driven by spurious patterns, the distribution of the abnormal metric of study around placebo dates should be noisy with no systematic patterns, providing a benchmark against which to interpret the CARs obtained at the true 8-K filing dates.

4.2 Experiments

This subsection presents the main experiments that produced significant results. Details of all other experiments can be found in Appendix D.

4.2.1 Initial Result

We first conducted an event study around the report dates of Form 8-K, without distinguishing between individual items, using the full sample of 99,384 events.
A challenge with this methodology is that averaging the CAR across both upward and downward trends tends to wash out meaningful patterns. As shown in Figure 2 we see a small scale and some counterintuitive results:

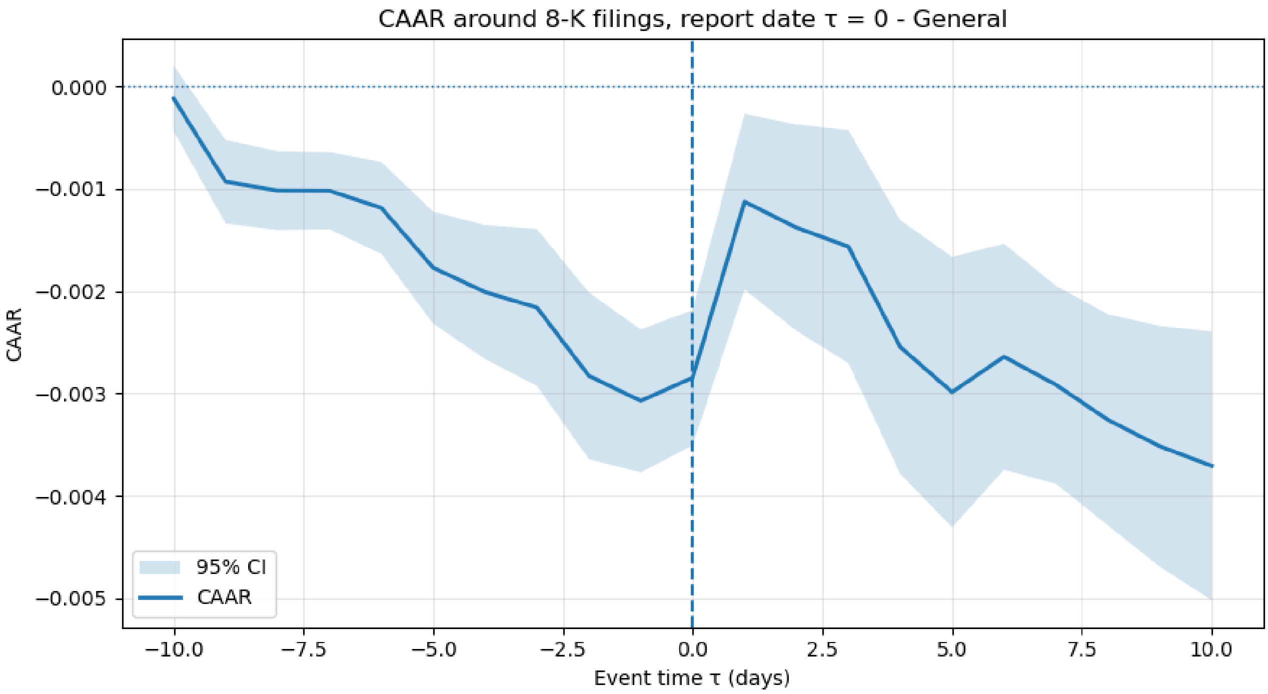


Figure 2: CAAR around report date for all 8-K filings events.

Here, one might have expected that an upward spike at the report date ($\tau = 0$) would be incorporated into prices and remain at the same level. Instead, the CAAR declines shortly afterward, indicating that the abnormal returns do not persist an outcome that does not support the presence of information trading. However, the scale of the effect is very close to zero (on the order of 10^{-3}), suggesting that this pattern may simply reflect mixed return reactions that offset one another rather than a meaningful economic signal. Nevertheless, we still observe a spike that begins to rise on the report date, supporting the hypothesis of the presence of informed trading.

To investigate this possibility, we examine spikes in average abnormal ITI (Average AI). Specifically, we perform an event study in which we replace Abnormal Returns with Abnormal ITI. Figure 3 shows average abnormal ITI around report dates, with actual 8-K events on the left and randomly selected non-8-K events on the right. Figure 4 presents the same analysis around filing dates.

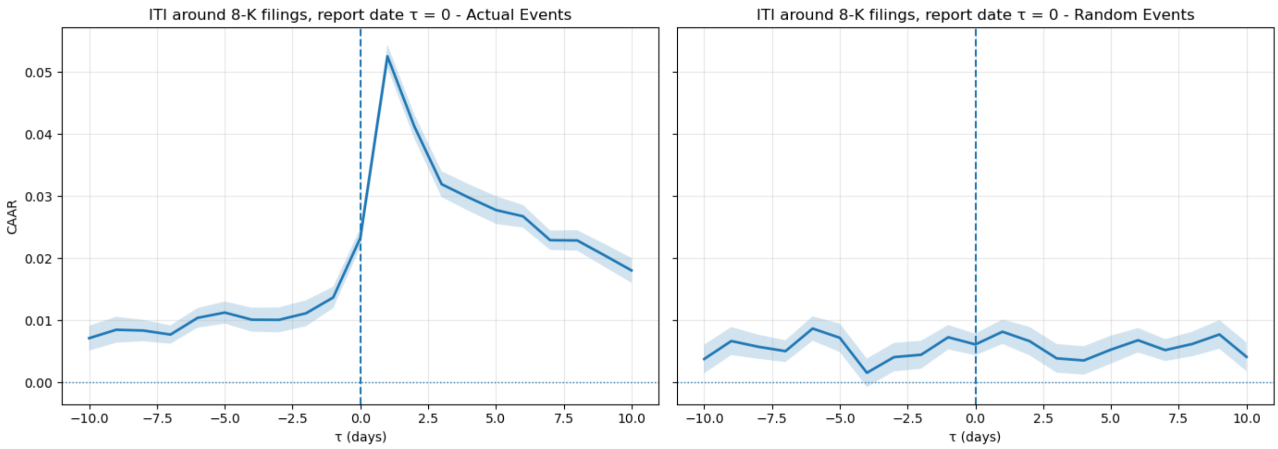


Figure 3: Average Abnormal ITI around report dates vs Average Abnormal ITI around random events with confidence interval.

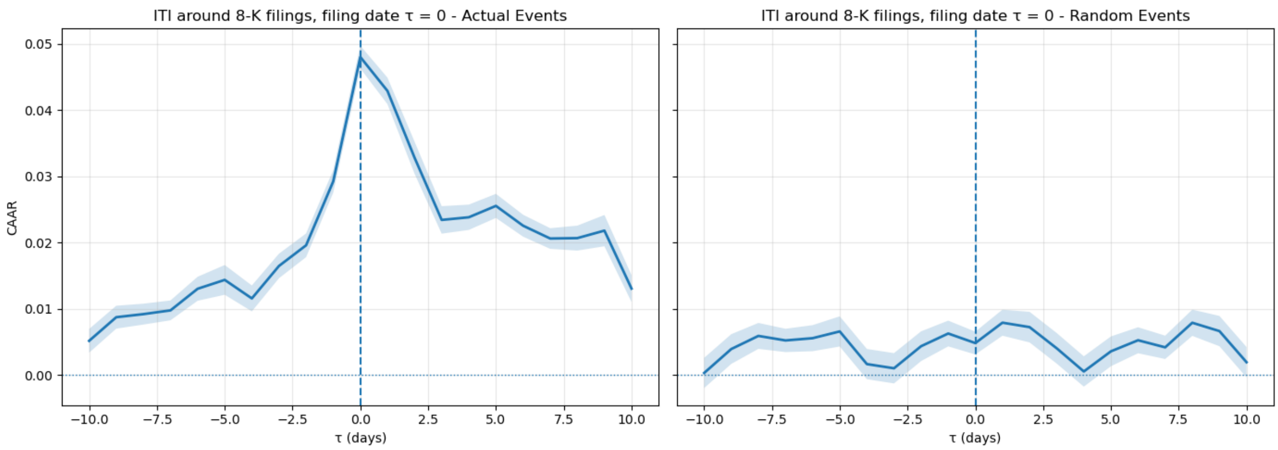


Figure 4: Average Abnormal ITI around filing dates vs Average Abnormal ITI around random events with confidence interval.

The average abnormal ITI behaves as intended: it exhibits a sharp increase immediately after the report date, followed by a rapid decline between $\tau = 0$ and $\tau \approx 2.5$, which is consistent with the fact that the filing date typically occurs within the $\tau \in [0, 4]$ window. This pattern indicates that average abnormal ITI spikes around the information release and then gradually reverts as the filing becomes publicly available. Surprisingly, the average abnormal ITI begins to rise slightly before the event date, suggesting that some information may be rumoured in advance. In contrast, placebo events show no systematic structure and fluctuate close to zero, reinforcing that the observed dynamics around actual events are not random but driven by information effects. The filing date graph further confirms our interpretation by displaying a pronounced downward spike at the filing date, consistent with the rapid decrease of Abnormal ITI once the information becomes fully public.

This event study using ITI clearly suggests that Form 8-K filings may convey CAAR-informative patterns that warrant further investigation with more refined processing.

4.2.2 Items Decomposition

Since a single filing may include multiple items, we separate filings according to their item composition. We then identify the ten most frequent item combinations, where a combination is defined as the exact set of items disclosed in a filing (e.g., a filing containing only Items 1.01 and 2.03 is classified as the combination 1.01, 2.03). Figure 5 presents the CAAR, and Figure 6 presents the average abnormal ITI around report dates for the ten most frequent item combinations.

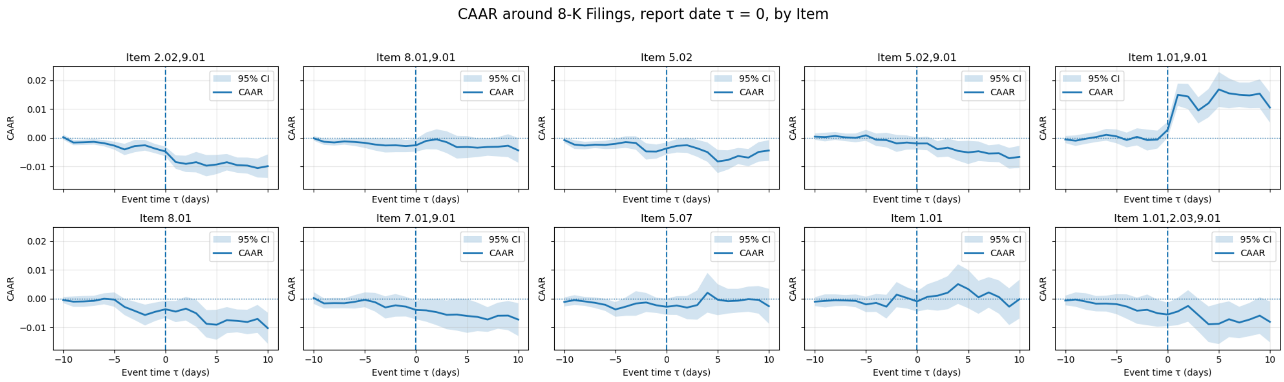


Figure 5: CAAR around report dates for the top 10 8-K filings item combinations

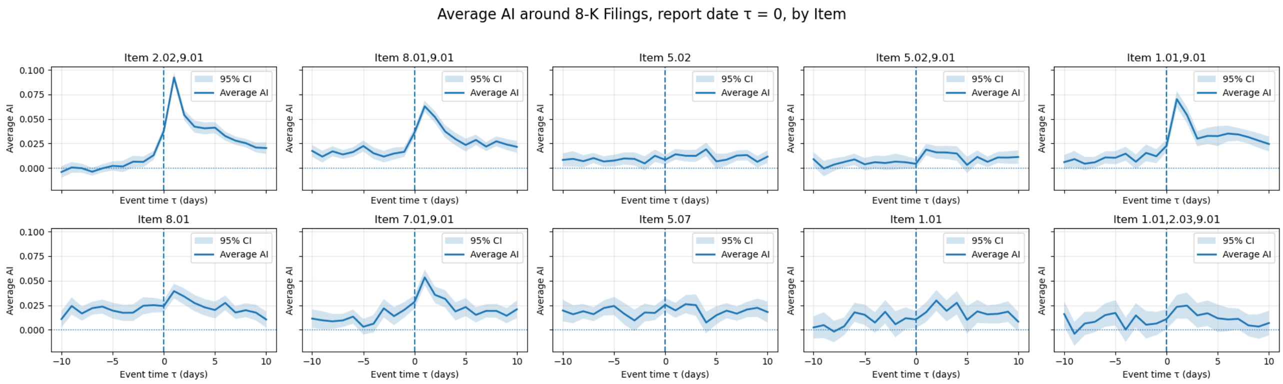


Figure 6: Average Abnormal ITI around report dates for the top 10 8-K filings item combinations

- **Item 2.02 - Results of Operations and Financial Condition.**
  This item exhibits a strong level of abnormal information intensity prior to disclosure and is associated with negative abnormal returns. Prices experience a modest decline around the report date before stabilizing, suggesting that negative information is gradually incorporated into prices ahead of the official announcement. Notably, the price decline begins prior to the event date, consistent with information leakage or insider trading before the disclosures. Overall, disclosures related to results of operations and financial condition are predominantly associated with unfavorable news.

- **Item 8.01 - Other Events.**
  Item 8.01 combined with Item 9.01 displays a moderate level of abnormal information intensity, while cumulative average abnormal return (CAAR) remains close to zero. This pattern suggests that these disclosures convey mixed informational content, combining both positive and negative news.

- **Item 5.02 - Departure and Appointment of Directors or Principal Officers.**
  This item does not exhibit elevated informative trading intensity, nor does it display systematic patterns in CAAR. This result is consistent with the ambiguous nature of these events: departures or appointments of directors and principal officers can be perceived as either positive or

negative by the market, depending on the specific context. Consequently, no clear directional price response or informed trading pattern emerges on average.

- **Item 5.07 - Submission of Matters to a Vote of Security Holders.**
  Item 5.07 similarly shows no evidence of elevated informative trading intensity and no significant CAAR around the disclosure date. This absence of abnormal trading activity and price response suggests that these filings generally convey information that is either anticipated by the market or of limited incremental informational value. Shareholder voting outcomes are often predictable or publicly discussed prior to formal disclosure, resulting in minimal information asymmetry.

- **Item 1.01 - Entry into a material definitive agreement**
  We observe that the expected pattern of informative trading intensity is clearly present for **Item 1.01** and **Item 9.01** filings. For both items, ITI increases prior to the filing date, consistent with informed trading ahead of material corporate events. Abnormal returns are positive around the report date but subsequently flatten, indicating a favorable market reaction at disclosure and rapid price adjustment thereafter. On average, these disclosures are therefore associated with good news that is largely incorporated into prices at the time of the report date. Interestingly, when adding **Item 2.03**, Creation of a Direct Financial Obligation or an Obligation under an Off-Balance Sheet Arrangement of a Registrant, the CAAR becomes negative, and the corresponding ITI response is considerably less pronounced. As these disclosures involve the contracting of debt or other financial obligations, they may be interpreted as negative signals by investors, leading to adverse price reactions. Moreover, such arrangements may often be subject to heightened confidentiality, which can limit information leakage and reduce the scope for informed trading prior to disclosure.

The largest increases in ITI are concentrated in filings that include exhibits (**Item 9.01**), suggesting that the presence of detailed contractual or supporting documentation is a key driver of informed trading activity under this metric.

We conducted the same analysis around the filing date. Figure 7 reports the CAARs, while Figure 8 reports the average abnormal ITI around the filing date for the ten most frequent item combinations.

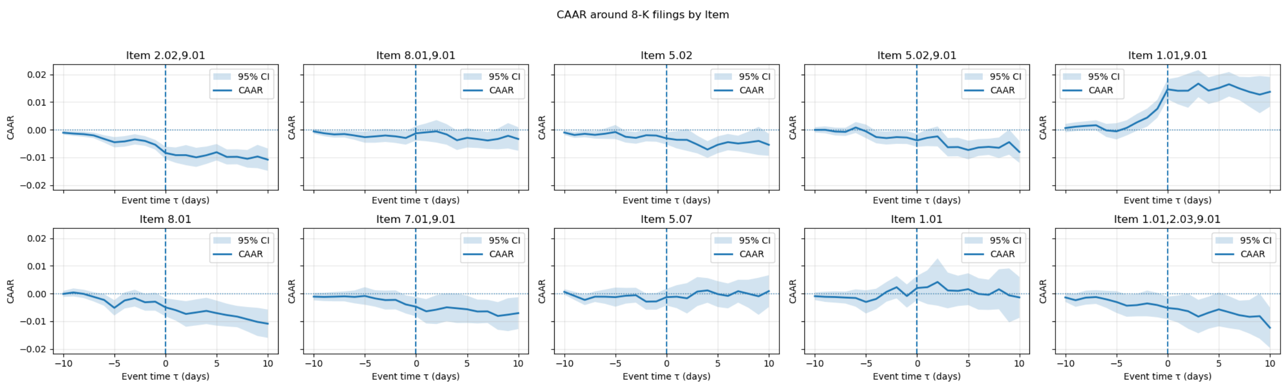


Figure 7: CAAR around filing dates for the top 10 8-K filings item combinations

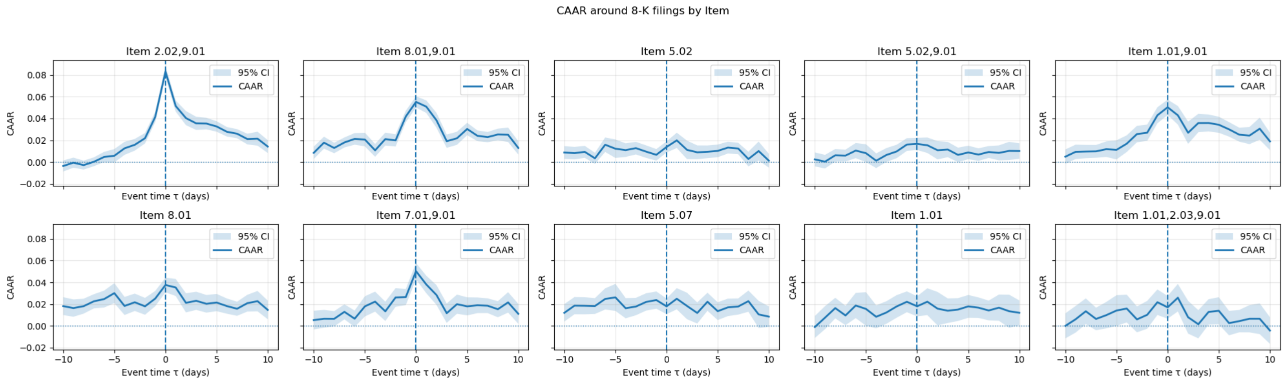


Figure 8: Average Abnormal ITI around filing dates for the top 10 8-K filings item combinations

We observe patterns that are very similar to those obtained using the report date. However, in this case, the peak is concentrated at the filing date, which is consistent with our previous observations. One could also design a trading strategy that takes a long (short) position upon the publication of a specific item combination when this combination exhibits an increasing (decreasing) CAAR following $\tau = 0$, which corresponds to the public filing date of the Form 8-K. For instance, one could take a short position when the firm publishes an Item 8.01.

This diversity in informational content across 8-K items complicates a uniform classification of disclosures based on return direction alone and motivates an alternative approach to characterizing market reactions.

4.2.3 Absolute Returns

To mitigate this issue, we attempted to classify returns based on the magnitude of the CAR at $\tau = t_i$, or alternatively using CAR quantiles, (see Appendix D), but both approaches introduced bias. To address this limitation, we instead focused on absolute returns, acting as a volatility proxy, in order to capture abnormal volatility spikes. This allowed us to investigate whether meaningful interactions arise between abnormal absolute returns and abnormal ITI.

We model absolute returns as a function of common risk-factor realizations (same setting as our previous event studies, but taking absolute returns as the dependent variable), which act as proxies for aggregate market activity and volatility. By doing so we obtained the results in Figure 9:

Figure 9: CAAR absolute returns around report date.

The graph reveals a pronounced spike in absolute returns exactly at the report date $\tau = 0$, followed by a persistent elevation over the subsequent days. This indicates a sharp and sudden increase in return volatility when the 8-K report is released, consistent with a strong market reaction to new information. In contrast, the placebo events exhibit no comparable pattern: absolute returns remain low, fluctuate without structure, and display no noticeable jump at $\tau = 0$. The difference between the two panels strongly suggests that the volatility spike observed for actual events is not random but is instead driven by the information content of the 8-K filings.

To further explore this relationship, we conducted a correlation analysis between Abnormal ITI and absolute abnormal returns over the same set of events. Figure 10 displays the resulting correlation between these two measures.

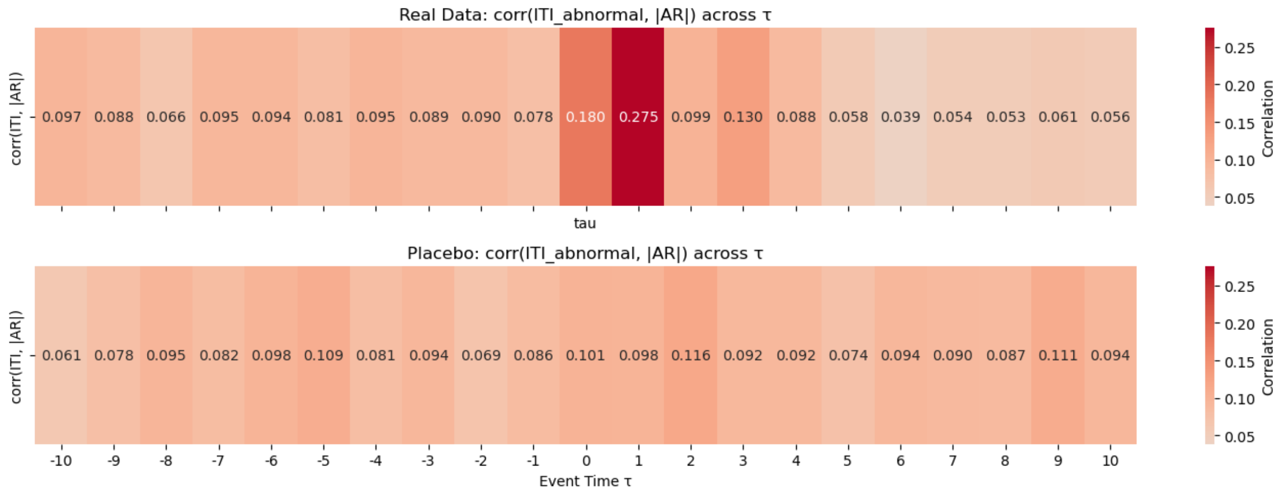


Figure 10: Correlation between Abnormal ITI and abnormal absolute returns.

The correlation analysis reveals a noticeable difference between actual events and placebo events. For the real data, the correlation between abnormal ITI and absolute returns increases sharply around $\tau = 0$, reaching approximately 0.28 at $\tau = 1$, the highest value across the entire window. This localized peak suggests that periods of unusually high information trading intensity tend to coincide with elevated return volatility shortly after the event. Outside this window, correlations are low and relatively stable.
In contrast, the placebo correlations show no such structure: they remain low (except some small spikes), fluctuate within a narrow band, and do not exhibit any pronounced peak around the event date. This flat pattern indicates that no systematic relationship exists between ITI and absolute returns when the event timestamps are randomized.
The comparison strongly suggests that the correlation spike observed in the real data is not random but likely reflects information-driven trading dynamics around the event.

The sharp increase in abnormal absolute returns around Form 8-K disclosures highlights the intensity of market reactions but also underscores a key limitation of aggregate market-based measures. Similar volatility responses can arise from disclosures conveying very different economic information, and this issue is further amplified by the heterogeneity of Form 8-K items. Some items, and particularly broad categories such as Item 8.01, bundle together disclosures with diverse and sometimes opposing implications. As a result, filings within the same item can generate mixed informational signals, making it difficult to interpret volatility or return patterns without additional structure.

4.2.4 Natural Language Processing

To complement our item-level analysis, we employ natural language processing techniques directly on the text of the disclosures. For this analysis, we restrict the dataset to filings that contain at least one Item 8.01, resulting in 13,359 observations. Our objective is to assign a sentiment score to each Item 8.01 and classify the filings into the top 10%, bottom 10%, or middle 80% based on their sentiment. We compute these sentiment scores using the three approaches introduced in Section 3.2: Simple FinBERT, FinBERT Mean, and Mistral FinBERT.

We begin by plotting in Figure 11 the average abnormal ITI around the 8.01 items report dates. (Note that this is not the same setup as in 5, since here we are considering any filings containing at least an Item 8.01.)
As in Figure 3 we see a spike immediately after the report date, followed by a rapid decline between $\tau = 0$ and $\tau \approx 2.5$, as it is here a subset of all 8-K filings, that thus follows the same pattern. In Figure 12 we show the CAAR around the 8.01 items report dates, which is also similar to Figure 2. Items 8.01 thus represent a good subset of all the 8-K filings and we will use it for our next analysis.

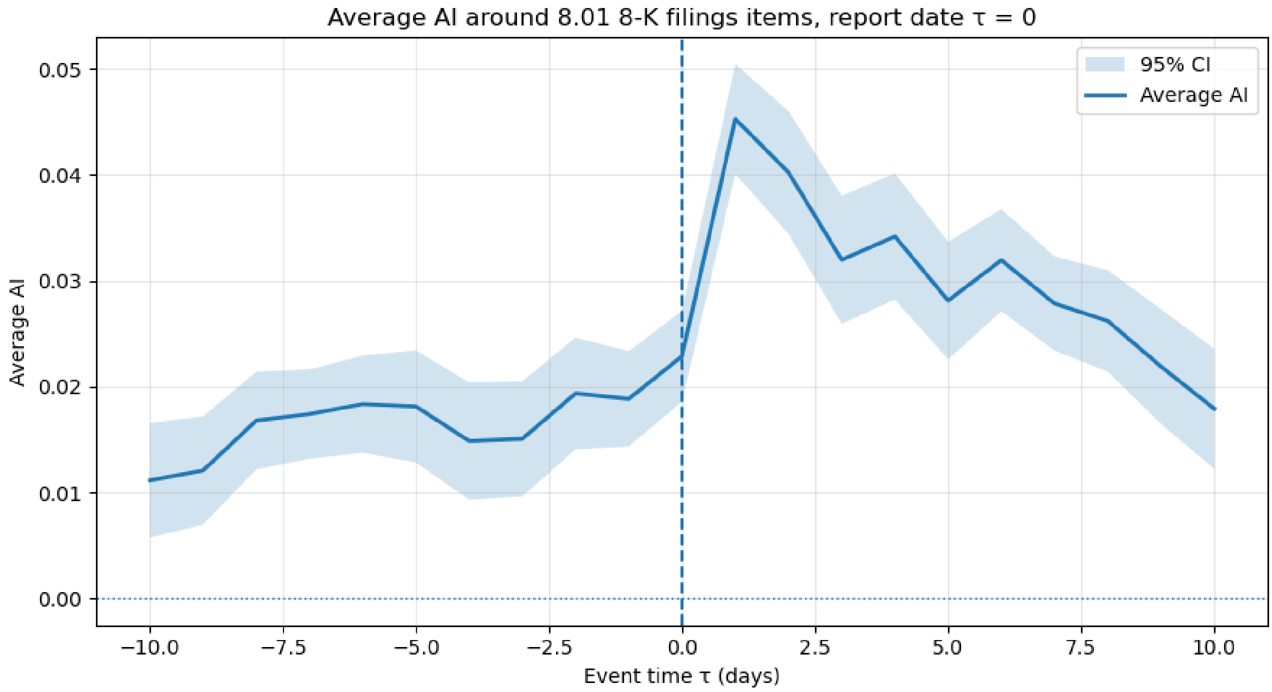


Figure 11: Average Abnormal ITI around the report dates of Item 8.01 in 8-K filings

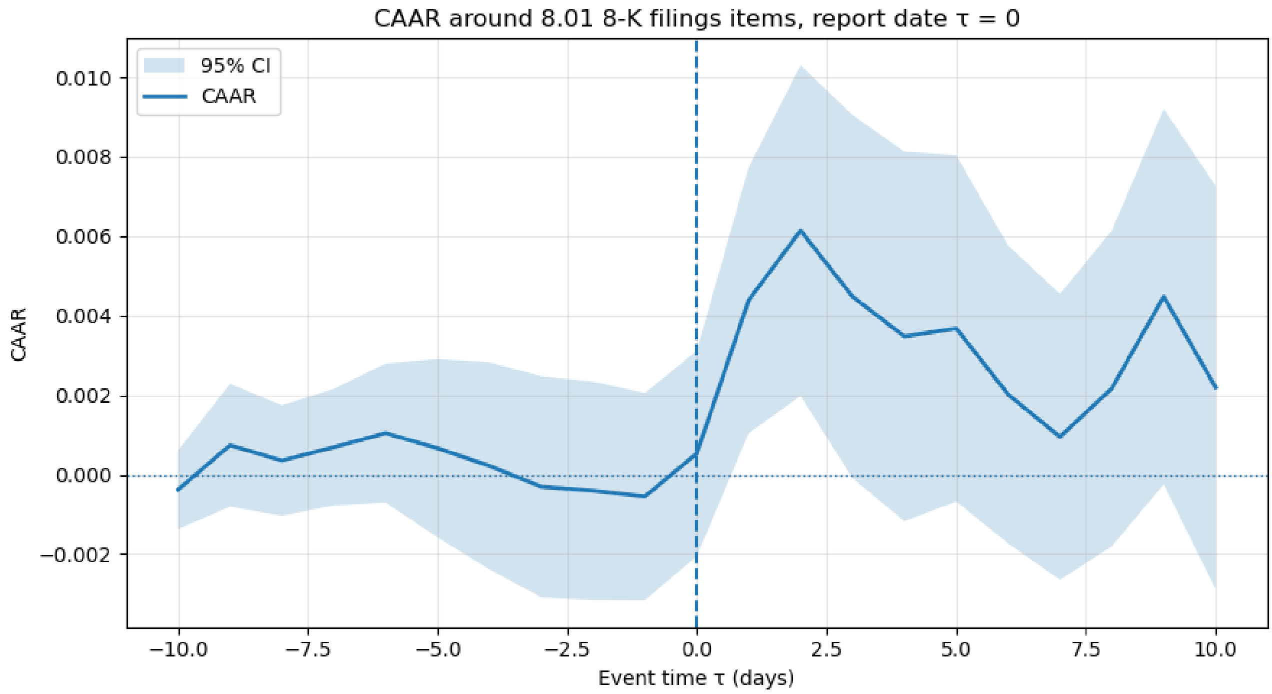
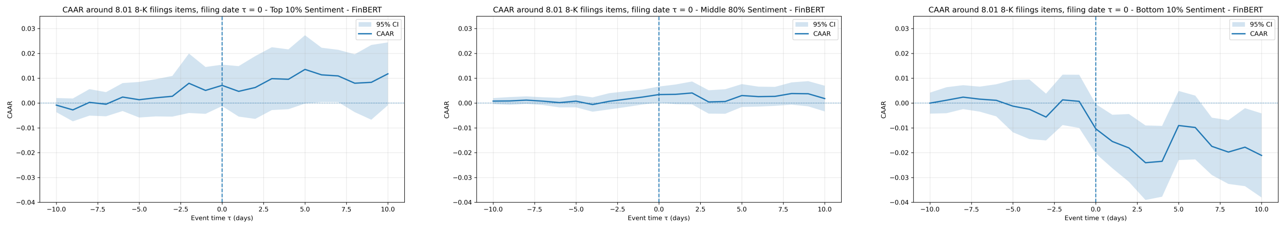


Figure 12: Cumulative Average Abnormal Return around the report dates of Item 8.01 in 8-K filings

We now apply the three sentiment extraction methods described above to the Item 8.01 disclosures, grouping them into the top 10%, middle 80%, and bottom 10% based on their sentiment scores. Figure 13 presents the CAAR around the filing dates using the simple FinBERT method, which relies solely on the first 512 tokens of each disclosure:

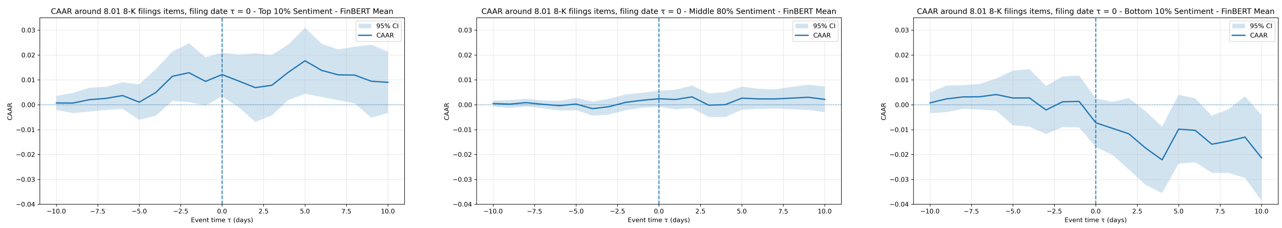


(a) Top 10% sentiment (b) Middle 80% sentiment (c) Bottom 10% sentiment

Figure 13: Cumulative Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT sentiment quantiles

For the top 10% sentiment group, the CAAR increases in the days preceding the filing date and continues to rise afterward. In contrast, the bottom 10% group exhibits a decline in CAAR following the filing date, while the middle 80% group remains relatively stable.

We next apply the FinBERT Mean method, which averages sentiment scores across all 512-token chunks of each Item 8.01 disclosure. Figure 14 shows the corresponding CAAR:

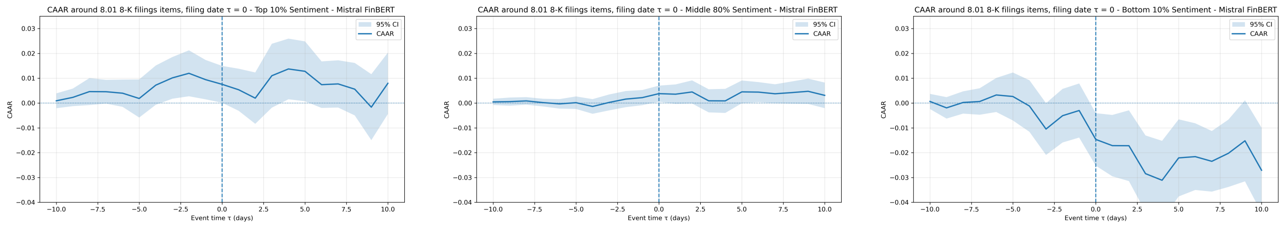


(a) Top 10% sentiment (b) Middle 80% sentiment (c) Bottom 10% sentiment

Figure 14: Cumulative Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT Mean sentiment quantiles

The CAAR dynamics are broadly similar to those obtained with the simple FinBERT method. The top 10% sentiment group exhibits an upward trend prior to $\tau = 0$, the middle 80% remains relatively flat, and the bottom 10% group shows a decline starting around $\tau = -4$. The similarity between the CAAR patterns obtained with the simple FinBERT method and the FinBERT Mean approach suggests that using only the first 512 tokens of an Item 8.01 disclosure is largely sufficient to capture its overall sentiment. In practice, the key information and tone of Item 8.01 filings are typically conveyed in the opening paragraphs, which summarize the nature of the event and its economic implications. As a result, extending sentiment analysis to the full text through chunking and averaging does not materially alter the classification of disclosures or the associated return dynamics.

Finally, we consider the Mistral FinBERT method, which first summarizes each Item 8.01 disclosure using a dedicated prompt (see Appendix C) and then applies FinBERT to the summarized text. Figure 15 reports the CAAR around filing dates:



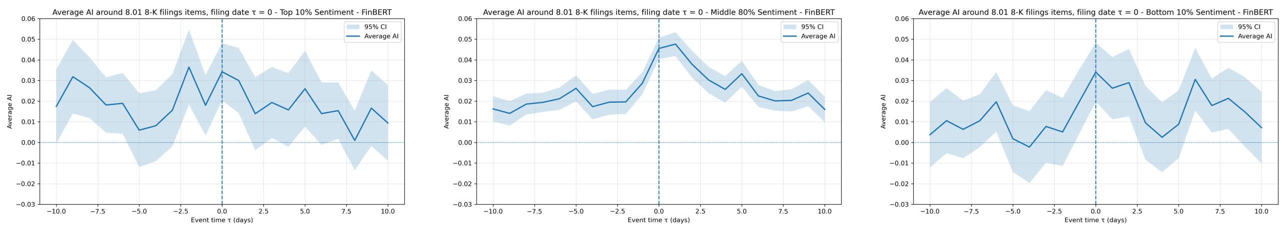
(a) Top 10% sentiment (b) Middle 80% sentiment (c) Bottom 10% sentiment

Figure 15: Cumulative Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by Mistral FinBERT sentiment quantiles

In contrast to the previous methods, the CAAR patterns are more clearly separated across sentiment groups. The top 10% group exhibits an increase, the middle 80% shows a mild upward trend, and the bottom 10% group displays a clear and persistent decline starting around $\tau = -4$ and continuing after the filing date, with the confidence interval remaining below zero for $\tau \geq 0$. This improved separation suggests that combining Mistral-based summarization with FinBERT sentiment scoring captures economically relevant information that is not fully exploited by the simple or Mean FinBERT approaches.

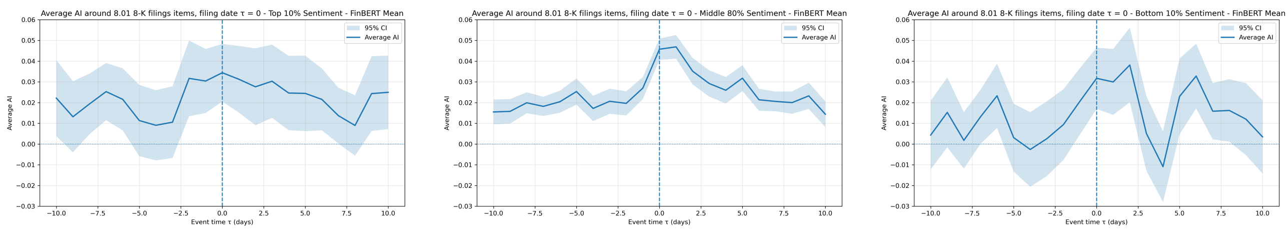
Importantly, this classification is based solely on textual content and sentiment extraction, avoiding the mechanical bias that arises when sorting events directly on CAR magnitude or CAR quantiles.

We also tried to separate the average abnormal ITI by sentiment score, using the same 3 methods. Figures 16, 17 and 18 present the average abnormal ITI around the filing dates for the three sentiment-based groups using the simple FinBERT, FinBERT Mean, Mistral FinBERT, respectively.



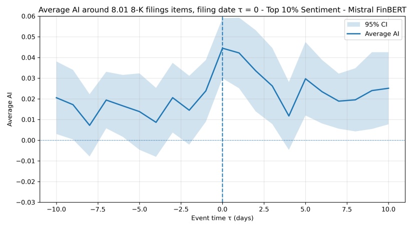
(a) Top 10% sentiment (b) Middle 80% sentiment (c) Bottom 10% sentiment

Figure 16: Average Abnormal ITI around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT sentiment quantiles

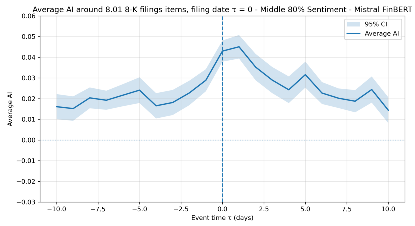


(a) Top 10% sentiment (b) Middle 80% sentiment (c) Bottom 10% sentiment

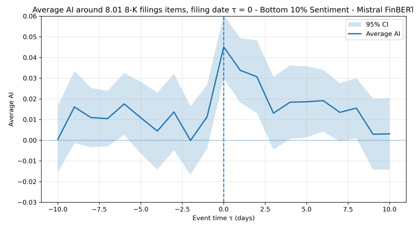
Figure 17: Average Abnormal ITI around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT Mean sentiment quantiles



(a) Top 10% sentiment



(b) Middle 80% sentiment

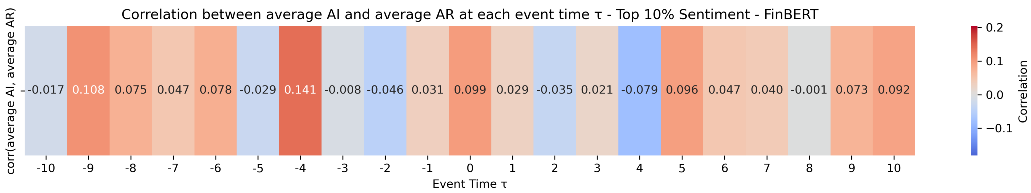


(c) Bottom 10% sentiment

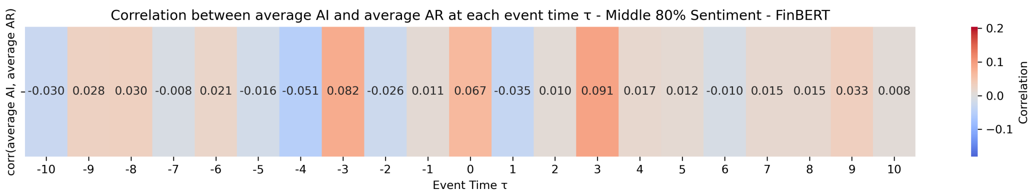
Figure 18: Average Abnormal ITI around the filing dates of Item 8.01 in 8-K filings, grouped by Mistral FinBERT sentiment quantiles

While the CAAR exhibits a clear separation across the top 10%, middle 80%, and bottom 10% sentiment groups, the pattern for average abnormal ITI does not show such differentiation. In all three sentiment groups, average abnormal ITI increases up to the filing date and subsequently declines. This pattern is consistent across all sentiment extraction methods. Although the confidence intervals appear wider for the top and bottom 10% groups, this reflects the smaller sample sizes in these groups compared with the middle 80%.

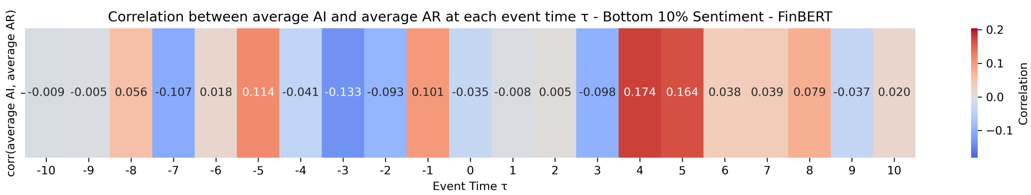
Finally, we computed the correlation between average abnormal ITI and average abnormal Return at each event time $\tau \in [-10, 10]$. Figure 19 presents the results for the top 10%, middle 80%, and bottom 10% sentiment groups using the simple FinBERT method, Figure 20 shows the same using the FinBERT Mean method, and Figure 21 corresponds to the Mistral FinBERT method:



(a) Top 10% sentiment

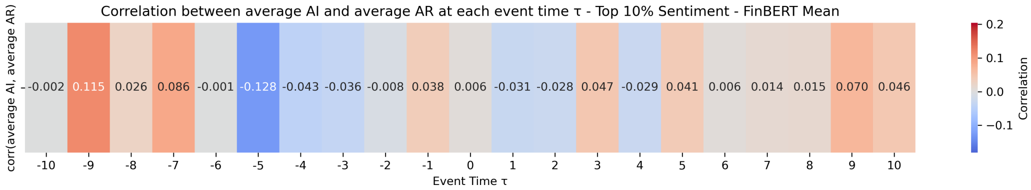


(b) Middle 80% sentiment

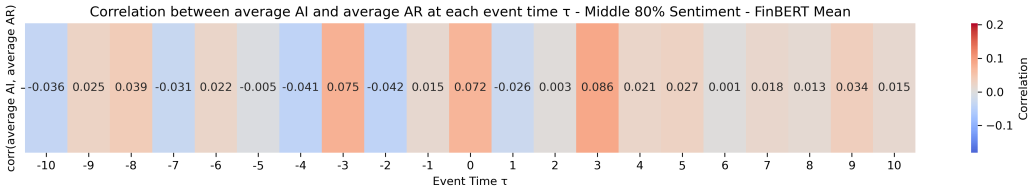


(c) Bottom 10% sentiment

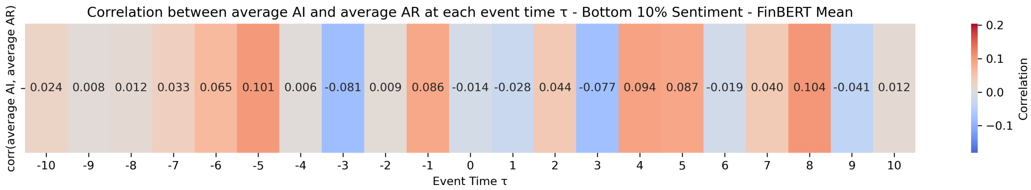
Figure 19: Correlation between Average Abnormal ITI and Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT sentiment quantiles



(a) Top 10% sentiment

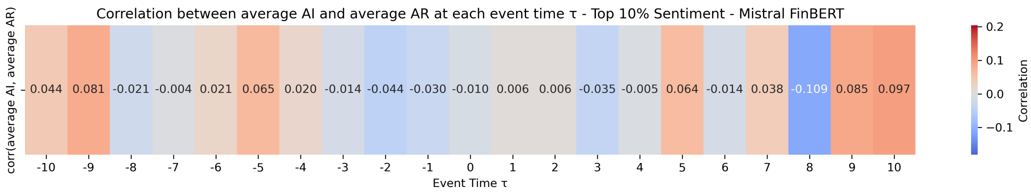


(b) Middle 80% sentiment

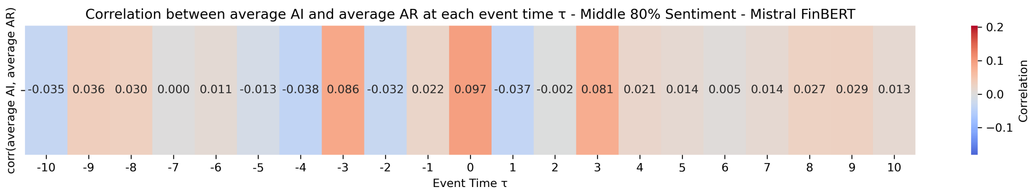


(c) Bottom 10% sentiment

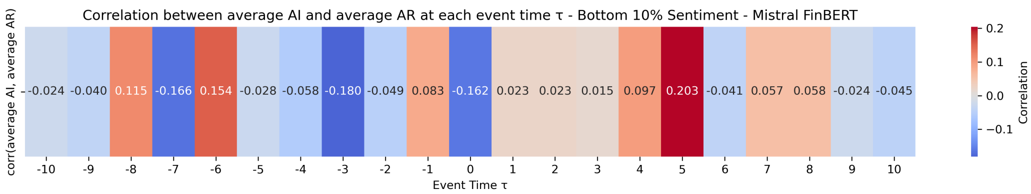
Figure 20: Correlation between Average Abnormal ITI and Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by FinBERT Mean sentiment quantiles



(a) Top 10% sentiment



(b) Middle 80% sentiment



(c) Bottom 10% sentiment

Figure 21: Correlation between Average Abnormal ITI and Average Abnormal Return around the filing dates of Item 8.01 in 8-K filings, grouped by Mistral FinBERT sentiment quantiles

Across all methods, we observe no systematic temporal pattern in the correlation between average abnormal ITI and average abnormal Return around the event date. Correlations alternate in sign across event times and remain close to zero for the top 10% and middle 80% sentiment groups,

suggesting a negligible association.
A notable difference emerges for the bottom 10% sentiment group. While the FinBERT Mean approach yields correlations comparable in magnitude to the other sentiment quantiles, both the raw FinBERT and Mistral FinBERT methods display higher dispersion, with a larger fraction of event days exhibiting correlations exceeding ± 0.1. This difference likely stems from the way negative disclosures are processed. In particular, averaging sentiment across all text chunks smooths extreme negative signals, whereas scoring only the beginning of the disclosure or a model-generated summary concentrates on the most material adverse information, leading to more extreme sentiment classifications and greater cross-sectional variation.
Despite these methodological differences, the overall magnitude of the correlations remains small, indicating limited economic significance and no robust evidence of a stable relationship between insider trading intensity and abnormal returns.

5 Conclusion and Future Work

This project investigates how ITI behaves around corporate information disclosures and how it interacts with both market activity and textual information. We tried different sources of financial news, and finally focused on Form 8-K filings, for which we combined event-study methods, item-level decomposition, and NLP applied to disclosure text to provide a clear and robust picture of how information is processed around corporate disclosure events.
Across all items, we document a pronounced spike in abnormal ITI around the filing date, even when cumulative abnormal returns remain close to zero.
A key finding of this study is the strong heterogeneity across different 8-K items. Filings associated with economically material actions such as entry into definitive agreements (Item 1.01), results of operations and financial condition (Item 2.02), and filings containing detailed exhibits (Item 9.01) exhibit pronounced increases in ITI prior to disclosure. These patterns are consistent with informed trading in settings where private information is likely to exist and where economic stakes are high. In contrast, items related to corporate governance, management changes (Item 5.02), or shareholder voting outcomes (Item 5.07) display little to no abnormal ITI, suggesting limited information asymmetry and a lower scope for informed trading.
At the return level, heterogeneous price reactions across disclosures imply that positive and negative abnormal returns often offset each other in the aggregate, resulting in average effects close to zero. This reinforces the importance of focusing on absolute abnormal returns, which better capture the intensity of market reactions to information releases.
Accordingly, we focused on abnormal absolute returns, which display a pronounced spike at the disclosure date, indicating a sharp increase in volatility and market activity following information release.
The correlation between abnormal ITI and absolute returns peaks around real event dates but not for placebo events, reinforcing the information-based nature of this relationship.
To strengthen our analysis, because signed returns are difficult to interpret in the presence of heterogeneous disclosures, we rely on textual analysis to classify news based solely on disclosure content. Focusing on Item 8.01 filings, we show that sentiment-based classification produces economically meaningful separation in return dynamics. While simple FinBERT-based approaches already capture much of this information, combining large language model–based summarization with FinBERT sentiment scoring leads to a clearer and more persistent separation of abnormal returns across sentiment groups. This suggests that summarization helps concentrate economically relevant information in complex disclosures.
Finally, while sentiment strongly affects return dynamics, ITI behaves similarly across sentiment groups, rising before disclosure and declining thereafter. Correlations between ITI and signed absolute returns remain small and unstable for Items 8.01, indicating that ITI should be interpreted as a measure of information intensity rather than a directional predictor of returns.

Several avenues for future research emerge from this analysis. First, the item-level approach could be refined by restricting the sample to single-item filings. Additional filtering, such as excluding very short disclosures, may further reduce noise in NLP-based measures.

Second, future work could extend the NLP framework beyond sentiment to incorporate topic modeling or semantic embeddings, enabling a richer characterization of the economic nature of disclosures. The analysis could also be extended to include exhibit filings. Such methods may help distinguish between operational, financial, legal, and strategic events that sentiment alone cannot fully capture.

Third, while this study focuses on contemporaneous relationships, further research could explore lead-lag dynamics between ITI, returns, and volatility. Examining whether changes in ITI systematically precede market reactions would provide deeper insight into the timing of informed trading relative to price discovery.

Finally, extending the analysis to alternative and higher-quality news sources, such as proprietary financial news datasets, could improve the precision of information-event identification.

Appendix

A Additional Datasets

Bloomberg-Reuters

For the news dataset, our first choice was to explore the Bloomberg-Reuters dataset [5]. While this dataset has been widely used in the literature for sentiment analysis and asset pricing applications, it presents several limitations for our setting. In particular, news articles often bundle multiple pieces of information, vary substantially in informational content, and do not always correspond to a clearly identifiable economic event. Moreover, the publication timestamp does not necessarily coincide with the underlying event date, which complicates event-level alignment with market reactions and information-based trading measures.
As a result, although the Bloomberg-Reuters dataset provides a rich source of textual financial information, it introduces a significant amount of noise when used for precise event-based analysis.

FNSPID Dataset

After the Bloomberg-Reuters dataset, we tried to explore the FNSPID dataset [6]. Although FNSPID is one of the most comprehensive and well-documented financial news datasets currently available, we found it to be highly heterogeneous and noisy in practice. Many of the news entries are redundant or irrelevant to the corresponding firm events, making the dataset difficult to use directly without extensive preprocessing and filtering. Moreover, we observed that several major firms such as Apple (AAPL) or other large-cap stocks have few or no recent news entries, limiting the dataset's representativeness for contemporary market analysis. As a result, despite its scale and potential, we decided not to rely on it for the core part of our analysis.

GDELT

We also looked at the GDELT project [7] to collect financial news. We implemented a fully operational Python pipeline, available on GitHub, that downloads and parses the GDELT GKG files. GDELT provides a new GKG file every 15 minutes, each containing structured information such as the timestamp, source URL, event type, locations, organizations, and additional metadata. Our pipeline allows downloading files for a given day, month, or year, extracts the CSV content from the ZIPs, filters for financial themes (e.g., `ECON_STOCKMARKET`) and US locations, parses organizations and page titles, and saves the cleaned data into daily, monthly, or yearly Parquet files for efficient processing.
The main steps of the pipeline are:

1. Retrieve the master file list from GDELT and select URLs corresponding to the desired date(s).
2. Download each ZIP file if it has not already been retrieved.
3. Extract the CSV file from the ZIP and read it, handling encoding issues.
4. Filter the data to keep only relevant financial themes and US locations.
5. Parse additional fields: extract `PAGE_TITLE` from `EXTRASXML`, parse `ORGANIZATIONS` into a list, and convert timestamps to datetime objects.
6. Concatenate all files for a day/month/year and save as Parquet for downstream use.

Despite the functionality of this pipeline, several limitations prevent its direct use for our historical ITI dataset. First, the actual news headlines are only reliably stored in the `EXTRASXML` field starting from 2020, whereas our ITI dataset only covers up to 2019. Inferring headlines from URLs is possible in some cases (as URLs often contain the title with dashes instead of spaces), but this method is highly inconsistent and introduces excessive noise.
Another limitation is the mapping of recognized organizations to financial identifiers such as tickers or PERMCO codes. Each article may contain multiple recognized entities, often including unrelated

companies or organizations. Disambiguating which entity corresponds to the relevant financial instrument is challenging and error-prone. This further complicates the direct integration of GDELT with our ITI dataset, making it unsuitable for our historical analysis.

Schedule 13D filings

At the beginning of the project, we aimed to reproduce the Informative Trading Intensity (ITI) metric [1] using more recent data. The original study relies on approximately 13,000 Schedule 13D filings between 1994 and 2018, which we attempted to replicate by developing a Python script to scrape the SEC EDGAR database and build an updated dataset of activist investor transactions. However, we ultimately did not use this dataset in our empirical analysis, since ITI also requires access to the Trade and Quote (TAQ) database to compute intraday microstructure variables such as spreads, depths, and realized volatility. Unfortunately, this dataset is not included in the EPFL's WRDS license, preventing us from fully reproducing the ITI methodology.

B Full List of 8-K Items

A Form 8-K filing can contain one or multiple of the following items:

- **1.01** Registrant's Business and Operations
- **1.02** Entry into a Material Definitive Agreement
- **1.03** Termination of a Material Definitive Agreement
- **1.04** Bankruptcy or Receivership
- **2.01** Completion of Acquisition or Disposition of Assets
- **2.02** Results of Operations and Financial Condition
- **2.03** Creation of a Direct Financial Obligation or an Obligation under an Off-Balance Sheet Arrangement of a Registrant
- **2.04** Triggering Events That Accelerate or Increase a Direct Financial Obligation or an Obligation under an Off-Balance Sheet Arrangement
- **2.05** Costs Associated with Exit or Disposal Activities
- **2.06** Material Impairments
- **3.01** Notice of Delisting or Failure to Satisfy a Continued Listing Rule or Standard; Transfer of Listing
- **3.02** Changes in Registrant's Certifying Accountant
- **3.03** Non-Reliance on Previously Issued Financial Statements or a Related Audit Report or Completed Interim
- **4.01** Unregistered Sales of Equity Securities
- **4.02** Material Modification to Rights of Security Holders
- **5.01** Corporate Governance and Management
- **5.02** Changes in Control of Registrant
- **5.03** Departure of Directors or Certain Officers; Election of Directors; Appointment of Certain Officers

- **5.04** Amendments to Articles of Incorporation or Bylaws; Change in Fiscal Year
- **5.05** Temporary Suspension of Trading Under Registrant's Employee Benefit Plan
- **5.06** Amendment to Registrant's Code of Ethics, or Waiver of a Provision of the Code of Ethics
- **5.07** Change in Shell Company Status
- **5.08** Submission of Matters to a Vote of Security Holders
- **6.01** ABS Informational and Computational Materials
- **6.02** Change of Servicer or Trustee
- **6.03** Change in Credit Enhancement or Other External Support
- **6.04** Failure to Make a Required Distribution
- **6.05** Securities Act Updating Disclosure
- **7.01** Regulation FD Disclosure
- **8.01** Other Events (to report events that are not specifically called for by Form 8-K, that the registrant considers to be of importance to security holders)
- **9.01** Financial Statements and Exhibits

C Mistral FinBERT Prompt

The prompt used for summarizing the Item 8.01 sections before passing them to FinBERT was as follows:

```
You are summarizing a section extracted from a corporate SEC Form 8-K filing.
This section corresponds to item {item}.

{text}

Produce a concise and factual summary (3 to 5 sentences) that captures the key
events, decisions, financial impacts, risks or disclosures mentioned in the text.
Focus on information that materially affects the company's business,
performance or outlook.
Do not include any sentiment, opinion, interpretation, or speculation.
This summary will later be used as input to a separate sentiment analysis model,
so ensure the summary is neutral, factual, and captures all elements
relevant for such an assessment.
```

D Supplementary Analysis

This section presents additional experiments that did not produce significant or informative results. In Figure 22, events are classified based on their short-horizon cumulative abnormal returns quantiles, with positive (negative) events defined as those in the top (bottom) decile of the CAR distribution over the [0,+1] window. Several alternative windows were tested, but all introduced bias.

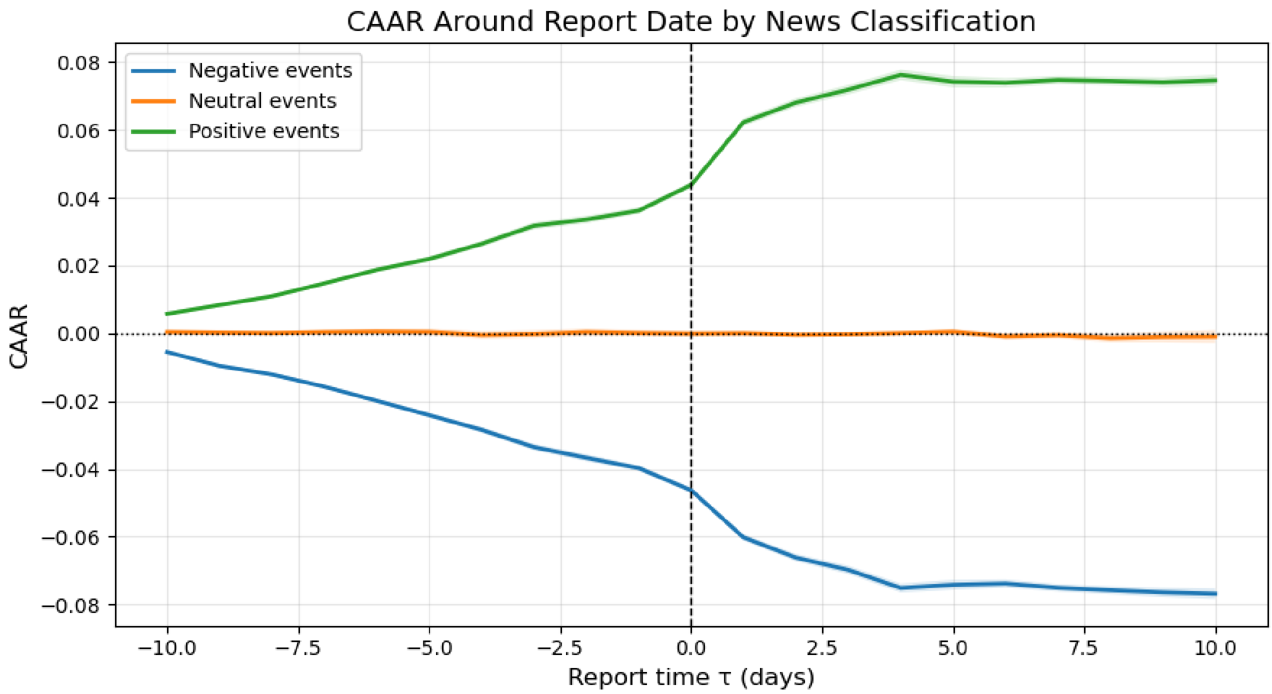


Figure 22: CAAR around report dates, classified by short-horizon CAR quantiles.

We classified the CAAR around report dates by ITI quantiles (top and bottom 10%), as shown in Figures 23 and 24. The CAAR patterns are not distinctly separated across these ITI quantiles.

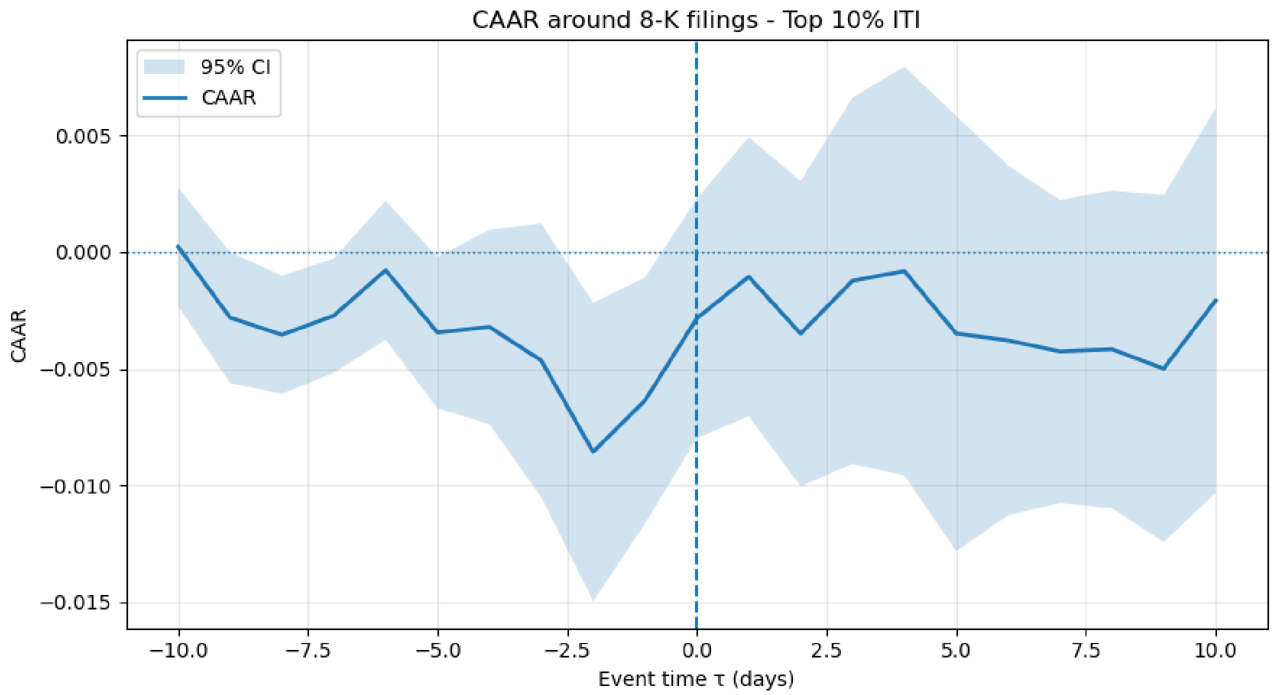


Figure 23: CAAR around report dates, for the top 10% ITI.

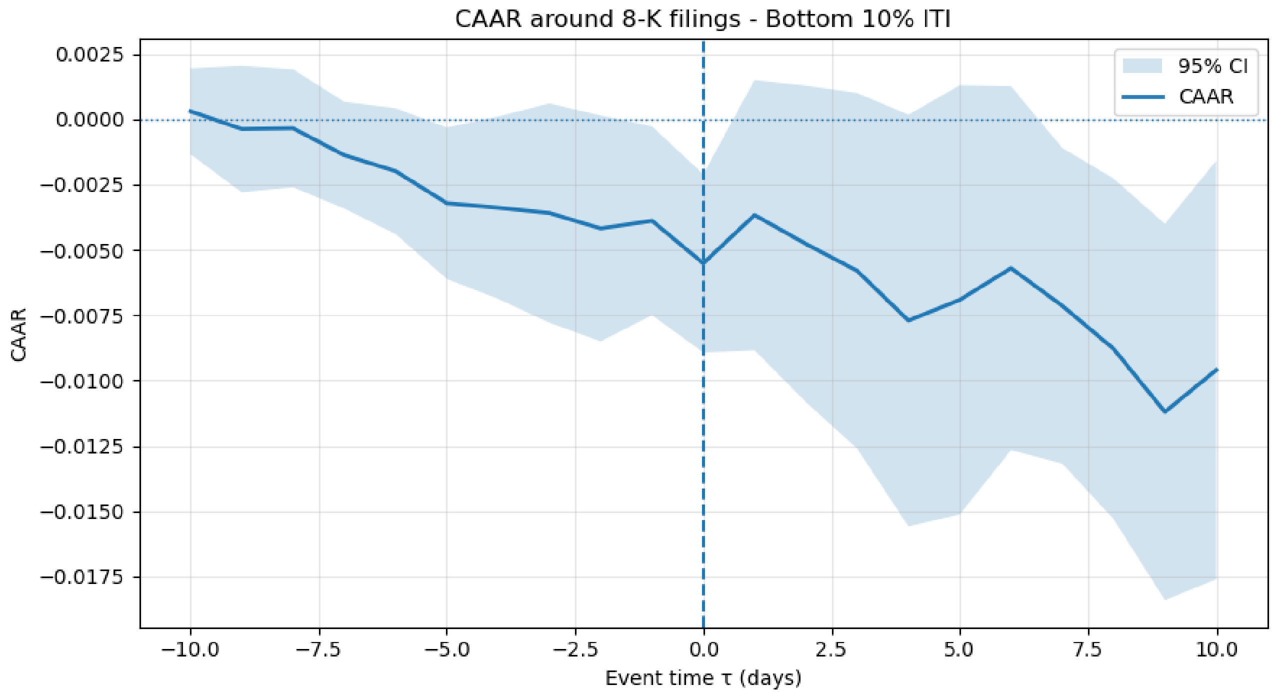


Figure 24: CAAR around report dates, for the bottom 10% ITI.

We classified events based on market capitalization, plotting only the top 10% quantile in Figure 25. This classification provides little informational value for explaining CAAR patterns.

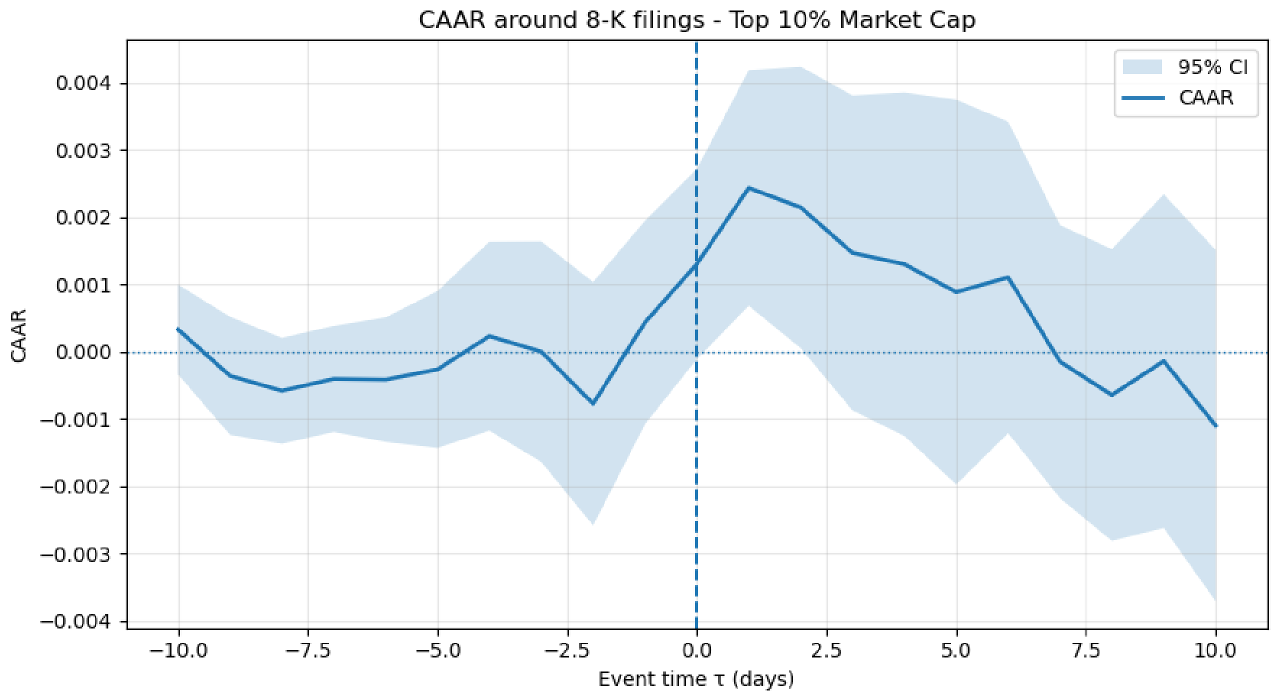


Figure 25: CAAR around report dates, for the top 10% market capitalization

References

[1] VINCENT BOGOUSSLAVSKY, VYACHESLAV FOS, and DMITRIY MURAVYEV. “Informed Trading Intensity”. In: *The Journal of Finance* 79.2 (2024), pp. 903–948. DOI: https://doi.org/10.1111/jofi.13320. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.13320. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.13320.

[2] Albert S. Kyle. “Continuous Auctions and Insider Trading”. In: *Econometrica* 53.6 (1985), pp. 1315–1335. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1913210 (visited on 12/30/2025).

[3] Lawrence R. Glosten and Paul R. Milgrom. “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders”. In: *Journal of Financial Economics* 14.1 (1985), pp. 71–100. ISSN: 0304-405X. DOI: https://doi.org/10.1016/0304-405X(85)90044-3. URL: https://www.sciencedirect.com/science/article/pii/0304405X85900443.

[4] David Easley and Maureen O’Hara. “Price, trade size, and information in securities markets”. In: *Journal of Financial Economics* 19.1 (1987), pp. 69–90. ISSN: 0304-405X. DOI: https://doi.org/10.1016/0304-405X(87)90029-8. URL: https://www.sciencedirect.com/science/article/pii/0304405X87900298.

[5] Xiao Ding Philippe Remy. *Financial News Dataset from Bloomberg and Reuters*. https://github.com/philipperemy/financial-news-dataset. 2015.

[6] Zihan Dong, Xinyu Fan, and Zhiyuan Peng. *FNSPID: A Comprehensive Financial News Dataset in Time Series*. 2024. arXiv: 2402.06698 [q-fin.ST].

[7] GDELT Project. “GDELT: Global data on events, location, and tone”. In: *ISA Annual Convention*. 2013. URL: https://www.gdeltproject.org/.

[8] Matthias Wyss and William Jallot. *iti-8k-analysis: Informative Trading Intensity and 8-K Filings with NLP-Based Sentiment*. https://github.com/Wfiles/iti-8k-analysis. Accessed: February 16, 2026. 2025.

[9] Dogu Araci. *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models*. 2019. arXiv: 1908.10063 [cs.CL]. URL: https://arxiv.org/abs/1908.10063.

[10] Mistral AI. *Mistral AI*. https://www.mistral.ai. Accessed in 2025. 2025.