Welcome to the Machine:
The Impact of News Analytics on High-Frequency Stock Market Dynamics

Danya Bouwman

A Thesis

in

The Department of Finance

John Molson School of Business

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Finance) at

Concordia University

Montreal, Quebec, Canada

February 2021

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Danya Bouwman

Entitled: Welcome to the Machine: The Impact of News Analytics on High-Frequency Stock Market Dynamics

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Business Administration – Finance Option)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Gregory Lypny *Chair's name*	Chair
Dr. Rahul Ravi *Examiner's name*	Examiner
Dr. Ravi Mateti *Examiner's name*	Examiner
Dr. Gregory Lypny *Supervisor's name*	Supervisor

Approved by Dr. Nilanjan Basu
Chair of Department or Graduate Program Director

March 9, 2021 Dr. Anne-Marie Croteau
Dean of Faculty

ABSTRACT

Welcome to the Machine:
The Impact of News Analytics on High-Frequency Stock Market Dynamics

Danya Bouwman

I investigate the impact of unscheduled, firm-specific news on high frequency stock market reactions using a five-year sample of intraday news releases and their corresponding Sentiment, Relevance, and Novelty scores generated by the Thomson Reuters News Analytics (TRNA) algorithm. My analysis features one-second interval price and volume data as well as matching trade and quote (TAQ) data for 55 Nasdaq stocks listed on the index between 2011 and 2015 inclusively. I examine cumulative abnormal returns, volumes, and trades, and further employ a quantile regression model that includes measures of news traffic to determine whether machine-readable news can effectively flag short-term trading opportunities. In line with related studies, I find significant increases in abnormal trading activity in the first few minutes surrounding a news release, with volume proving to be more sensitive to the TRNA metrics than returns. Positive news is traded more aggressively than negative news on the knee-jerk, but also experiences sharper reversals in abnormal returns in the hours following. Furthermore, results from the quantile regression analysis appear to confirm that news traffic in the run-up to and at the release time significantly impact abnormal returns. Although my results appear consistent with the hypothesis that trading activity should increase as TRNA thresholds become stricter, simulated holding period returns remain negative, highlighting the many complexities involved in algorithmically trading the news.

Keywords: Firm-specific news, News sentiment, TRNA, High-frequency data, Abnormal returns, Abnormal volume, Quantile regression, News Traffic, Nasdaq

Acknowledgments

I am deeply indebted to my supervisor, Gregory Lypny, for being a most thoughtful mentor and pillar of support throughout this entire process. Thank you for your genuine interest in and dedication to this research. Your extensive knowledge and expertise were invaluable in bringing this project to life, and your unrelenting perseverance and solidarity were a model of integrity for me. It was a sincere pleasure to work with you.

I am grateful to Arvind Ramanathan for facilitating the procurement of this unique dataset and granting me access to additional resources that were essential in conducting the analysis. Thanks to Elijah DePalma for instructive consultations on working with the TRNA dataset, and to Christopher Plenzich and Camille Ennab for lending their technical skills and expertise. Thanks to my committee members, Rahul Ravi and Ravi Mateti, whose patience and vote of support throughout the preparation and review of this manuscript were very much appreciated.

Nazim, thank you for your unwavering support throughout the many years it took to complete this thesis. Your patience cannot be overstated. To the rest of my friends and family who encouraged and endured me along the way, thank you for graciously participating in this journey.

Acknowledgement of Research Contributions

I conceived the idea for this research project and acquired the necessary data to perform the analysis. Professor Lypny led the development of algorithms and database management systems used to compress, process, and filter the data, and created key tools and tutorials that allowed me to conduct the analysis. We shared ideas and feedback that shaped the theoretical, conceptual, and analytical frameworks of the research. I analyzed the results and wrote the manuscript.

Danya Bouwman
February 2021

Table of Contents

List of Figures

List of Tables

1. Introduction

Information is the foundation of fundamental analysis and timely, unrestricted access to it allows financial markets to operate efficiently and transparently. Newswires play a quintessential role in this equation by vetting and reporting on the relevant issues affecting the global economy: news shocks act as direct catalysts for security repricing, while journalistic spin and reach proportionally affect the magnitude of the market's reactions.

For the past 20 years, an information revolution has been taking place in the world of news-based trading, bringing with it an exponential increase in the volume, breath, depth, and frequency of news data and underlying metadata available to market participants (Leinweber and Sisk, 2011). Coupled with an insatiable need for speed and the proliferation of algorithmic trading, these conditions have pushed the velocity of information to reach unfathomable levels.[1] News now breaks faster than ever before, and reaches its end users via a wider range of reputable and sometimes unconventional sources — the growing use of social media as a legitimate platform to disseminate information is a prime example of the latter. With the arrival rate of news quickly outpacing both the speed and limits of human processing, machines have become instrumental in streamlining the way we obtain, synthesize, and ultimately make decisions based on all this data.

Major newswires such as Bloomberg, Dow Jones, and Thomson Reuters have made themselves an integral part of this process by selling subscriptions to low-latency, machine-readable news feeds aimed at helping traders automate the job of making snap decisions. Broadly speaking, these feeds can be classified according to whether they supply quantitative (hard) or qualitative (soft) information. Petersen (2004) describes hard information as quantitative data that can be easily transmitted, stored, and interpreted electronically (such as quarterly GDP figures); while soft information is qualitative in nature, context-specific, and inherently subjective (such as the outcome of a G20 meeting), making it much harder to distill into a standardized metric without significant loss of information. While algorithmically trading quantitative data is a fairly straightforward process, reducing qualitative information into binary trading signals is a much more nuanced and discretionary task that has traditionally required a human touch, since the underlying connotations of a text often transcend literal words. Zhang (2012) concurs, concluding that while high-frequency traders have a clear advantage in reacting to hard information, low-frequency traders dominate at processing soft information, and tend to take their time incorporating it into security prices.

However, the two sides may be converging, as the advent of natural language processing (NLP) bridges the gap between digital data and human language, essentially blurring the lines between hard and soft information. Liddy (2001) defines NLP as "a theoretically motivated range of

---

[1] For a cursory history of how Thomson Reuters came to transit news algorithmically, see Appendix A.

computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications". As it relates to finance, NLP has been applied to a unique and robust range of textual datasets including, but not limited to, news articles, headlines, Tweets, blog posts, and chatroom discussions. NLP has also spurred the development of automated, real-time financial news analytics products that use proprietary algorithms to quantify (among other things) the tone, novelty, and relevance of a news item, metrics that help discern the news from the noise.

With NLP becoming increasingly sophisticated, it may be only a matter of time before algorithmically reacting to soft information using a machine becomes the new norm for news-based trading strategies. A 2016 MarketsandMarkets™ study estimated that the NLP industry would grow from USD 7.63 billion in 2016 to USD 16.07 billion in 2021 at a compounded annual growth rate (CAGR) of 16.1 per cent, and more recent estimates have been more bullish, with the industry valued at USD 10.93 billion in 2019 by Mordor Intelligence LLP, and forecast to reach USD 34.80 billion by 2025, which would constitute a CAGR of 21.5 per cent from 2020 to 2025. The expectation is that machines could soon replicate human interpretation and inference well enough to render seasoned traders less competitive in the marketplace, hence the focus of this analysis.

2. Literature Review

2.1. The shift towards machine-readable news

Market efficiency dictates that stock prices reflect all available information relevant to a firm's valuation, subject to the quantitative and qualitative information investors have at their disposal. Until recently, most studies had focused almost exclusively on hard data to explain market dynamics, categorically ignoring all the soft signals investors readily factor into their day-to-day trading decisions. According to Liberti and Petersen (2017), this is one of the major shortfalls of academic models — the fact that they are so heavily rooted in numerical data, while most of the information markets incorporate into stock prices is in written form. However, as enhanced processing capabilities have made analyzing increasingly large textual datasets possible, the body of financial literature devoted to modelling this wealth of unstructured information has grown (Fisher et al., 2016).

Studies on sentiment analysis have established techniques for hardening soft information using various forms of NLP, in turn demonstrating concrete links between qualitative triggers and market dynamics. One of the biggest outstanding research questions in the field is to what extent NLP can stand on its own, thereby eliminating the need for human intervention altogether. Recent studies are building an increasingly convincing case for it. Fisher et al. (2016) track the progress of NLP-related research from the emergence of text mining in the 1980's to the inclusion of machine learning and artificial intelligence in the late 1990's, noting a marked surge in the amount of research conducted from 2000 onwards. During this time, rudimentary bag-of-words or dictionary-based scoring techniques were replaced by naïve Bayes classifiers capable of analyzing news at the sentence or phrase level, producing very different and significantly improved results (Li, 2010; Boudoukh et al., 2012). Huang et al.'s (2014) scoring algorithm is a case in point, boasting a near 20 per cent improvement in accuracy over the commonly used Loughran and McDonald (2011) and Henry (2006) financial word lists.

As NLPs grew increasingly sophisticated, financial news became augmented with metadata consisting of automated, real-time descriptive metrics. This shift towards pre-scored, machine-readable news led to major changes in the nature and flow of information. For instance, the launch of the Thomson Reuters News Analytics (TRNA) suite in 2006 gave way to a fourfold increase in the monthly volume of news items released (Leinweber and Sisk, 2011), as well as a dramatic improvement in the breadth and depth of firm-specific coverage between 2003 and 2011, notably a 22 per cent increase in the number of US stocks reported on (Huynh and Smith, 2013). Furthermore, the RavenPack News Analytics (RPNA) suite led to faster price discovery following its introduction in 2009; von Beschwitz et al. (2015) observed that 35.7 per cent of the total price reaction to Dow Jones news occurred within 10 seconds of its release, compared to 28.4 per cent before RPNA went live.

The intersection of big data and speed competition has inevitably left investors with fleeting amounts of time to process increasingly vast amounts of information, conditions that dual process theory predicts will make them more reliant on intuition than deliberate reasoning to make decisions.[2] Enter our alleged hero, NLP. While the average human has the capacity to process approximately six articles an hour, news analytics software can easily sift through 10 articles a second (TRNA – White Paper, 2013). Beyond deciphering the lexical and syntactic ambiguities inherent to the written word, NLP algorithms are supposedly able to measure and score an article's tone, relevance, and novelty even before the news officially breaks. Sinha (2015) notes the TRNA sentiment engine is able to match the average tone classification of human analysts with 75 per cent accuracy, which TRNA considers impressive since humans only tend to agree with each other on any given article about 82 per cent of the time (TRNA – White Paper, 2013). Both Leinweber and Sisk (2011) and Cahan et al. (2010) observe a spike in the profitability of news sentiment portfolios shortly after this technological regime shift took place, confirming the progressively mainstreamed use of pre-scored news by the financial community.

Groß-Klußmann and Hautsch (2011), who were among the first to test the relationship between high-frequency market dynamics and unscheduled news, demonstrate that high-frequency market reactions are significantly correlated to intraday company news – crucially, when earnings announcements are excluded. Their results endorse the power of TRNA as a news filter, since the relevance score plays a critical role in identifying potentially market-moving news items, while the sentiment score helps predict intraday stock movements. Using the same event study framework outlined by Campbell et al. (1997) and employed by Groß-Klußmann and Hautsch (2011), Smales (2014b) validates the usefulness of RPNA in an international context, using a 12-year sample of unscheduled news for 33 highly liquid stocks listed on the Australian Stock Exchange (ASX). The relevance and sentiment scores are once again shown to cause definite spikes in trading activity, volatility, and spreads over 30-second intervals, with the impact of news shocks subsiding after two and a half minutes.

2.2. Sentiment analysis on newswires, analyst reports, corporate disclosures, and social media postings

In financial markets, the media acts as the primary intermediary of information disclosed by companies and shared with investors, playing a crucial role in shaping the collective beliefs of the investment community. The opinions expressed by newswires have an almost self-fulfilling-prophecy-esque quality to them, and tend to manifest themselves as price movements regardless of whether they represent new information or pure conjecture (such as rumours, sensationalism) (Shiller, 2005). With investors verifiably prone to letting heuristics and biases influence their

---

[2] Dual-process theory describes two distinct approaches for decision-making: one that is fast, impulsive, and spontaneous; and the other that is slow, deliberate, and methodical. Kahneman and Frederick (2001) explain that decisions made in the first system rely on heuristic judgement that is prone to bias. That is not to say that seasoned traders are less capable of trading news under time pressure, but rather that the processes governing their decision-making will change, leaving them more reliant on intuitive guesswork.

decision-making (Simon 1955, 1957; Tversky and Kahneman, 1974, 1979), research in sentiment analysis lends itself well to the behavioural notion of bounded rationality.[3]

One of the seminal works in sentiment analysis is Tetlock (2007), who uses a dictionary-based scoring technique to generate pessimism scores for the *Wall Street Journal*'s daily "Abreast of the Market" column. The 16-year sample, ending in 1999, shows that abnormally high and low levels of pessimism coincide with surges in trading volume, while high pessimism scores precede temporary downturns in market prices. The latter is consistent with sentiment theory that predicts short-term reversals to news that is emotive but doesn't actually contain new information. This effect is particularly strong for small stocks, reflecting the increased susceptibility of individual investors to media content, as compared to institutional investors who make up a bigger portion of the large stock audience.

Despite showing promising results, Loughran and McDonald (2011a) detect an important flaw in Tetlock's bag-of-words scoring method, in that most of the words considered negative according to the General Inquirer's Harvard-IV-4 classification dictionary don't actually have a negative connotation when used in the context of finance and accounting. This analysis, alongside similar comparisons conducted by Henry and Leone (2009) and Li (2010), confirm that generalized word categorization schemes are not appropriate for all disciplines, which led to the creation of domain-specific dictionaries that became the baseline for subsequent research in sentiment analysis.[4]

Building on Tetlock (2007), Garcia (2013) uses the Loughran and McDonald lists of financial-news-specific words to score the *New York Times*' "Financial Markets" and "Topics in Wall Street" columns from 1905 to 2005. The results obtained confirm Tetlock's (2007) findings, namely that news content helps predict daily price returns. Garcia goes on to show that investor sentiment has a more pronounced effect on the DJIA during recessions, notably on Mondays and after holidays, which is in line with results found by Smales (2014a,b) and Allen, McAleer and Singh (2013, 2015). More on this can be found in section 2.4. on 'Sentiment analysis during the financial crisis and periods of extreme positioning'.

Ferguson et al. (2015) are the first to apply a domain-specific method to a non-US sample, examining both positive and negative sentiment scores derived from FTSE 100 firm news harvested from a broad sample of UK newspapers (including the *Financial Times*, the *Times*, the *Guardian*, and the *Mirror*) between 1981 and 2010. Their major contribution comes from studying

---

[3] Simon's (1955) notion of bounded rationality addresses the shortfalls of the "economic man" by factoring in the limits imposed on his rationality by his access to information, computational capacity, and time.

[4] Creating a domain-specific dictionary essentially involves mining for terms with known polarity, which is inherently problematic when you consider how much manual knowledge engineering is often required to correctly interpret a text. NLP and machine learning are attempting to address this shortfall by developing systems for learning rules that relate to special cases.

the combined effect of tone and volume of media coverage on stock returns. While both variables have been known to significantly predict abnormal returns, they find the impact of news volume to be much more pronounced than tone. Their results also reveal increased market sensitivity to highly visible positive news, consistent with Barber and Odean's (2008) hypothesis that "individual investors are net buyers of attention-grabbing stocks".

Although the above-mentioned studies focus exclusively on news articles, we can draw important conclusions from similar analyses conducted on corporate disclosures and analyst reports. In these cases, the messenger is often as important as the message itself, and the two cannot be easily dissociated from one another. Kothari, Li and Short (2009) analyze an exhaustive sample of 326,357 firm-specific disclosures written by management, analysts, and the financial press for 889 companies over a six-year period. Their results point to news disclosed by the press as being both timelier and perceived as more credible than news released by either the company's managers or analysts, presumably because journalists face fewer agency problems. Taking this approach a step further, Dzielinski (2012) uses pre-scored TRNA news data to study how the source and tone of public news resolves information asymmetry. Using over one million firm-specific stories released between 2003 and 2011, Dzielinski concludes that only news released by both the company itself and the Thomson Reuters news agency led to smaller short-term return reversals. Neither source alone reduced information asymmetry, which confirms the important symbiotic role these agents play in financial markets. By filtering and relaying only the most relevant corporate announcements, newswires distil the news into a less noisy and more informative form. At the same time, however, the fact that newswire content alone fails to have a lasting impact on markets exposes their ineffectiveness at generating original content and their tendency to report on recent market moves in the absence of company news.

Apart from messenger-diffused content, investor-generated content has also been found to contain financially relevant information. Yu et al. (2013) confirm that social media chatter significantly relates to price movements in the stocks being referenced. Antweiler and Frank (2004) design a naïve Bayes algorithm to convert the content of one and a half million messages posted on Yahoo! Finance and Raging Bull into buy, hold, and sell ratings for 45 DJIA companies, concluding that both the volume of messages posted and the level of disagreement between them are positively related to trading volume and help predict volatility. Furthermore, the tone of these messages predicts statistically significant negative returns on the next trading day, consistent with Antweiler and Frank (2006), whose results similarly detect an overreaction to news followed by return reversals over a 10-day horizon. Das and Chen (2007) design five algorithms to classify Yahoo! Finance messages relating to 24 technology stocks on the Morgan Stanley High Tech Index between July and August 2001, noting that aggregate sentiment tracks index returns better than individual stocks. Bollen et al. (2011b) extract the tone of nine million Tweets that are then fed to a fuzzy neural network algorithm capable of predicting DJIA price movements with 87.6 per cent accuracy, while Vu et al. (2012) similarly show that Twitter sentiment can predict individual stock movements with 75-83 per cent accuracy.

2.3. Sentiment analysis and market drift

Taking a more targeted approach to content analysis, Tetlock et al. (2008) and Engelberg (2008) analyze Dow Jones News articles released ahead of quarterly earnings reports between 1980-2004 and 1999-2005 respectively, confirming that these contain hard-to-quantify information about firm fundamentals. Negative words serve as the strongest predictors of weak earnings and contain informational value in excess of quantitative indicators such as standardized unexpected earnings (SUE). Engelberg (2008) performs a cross-sectional analysis to show that "frictions" in information processing cause prices to drift, an effect that proves to be most pronounced for firms with low institutional ownership operating in informationally complex environments, such as small tech firms. This drift effect is consistent with information theory that predicts directional persistence in price movements over the longer term, though there remains some debate as to what exactly causes it. For instance, Demers and Vega (2010) and Feldman et al. (2010) find that the post earnings announcement drift (PEAD) is related to the incremental informational content of management-issued communications such as the tone of MD&As. Chan (2003) and Vega (2006) conclude that firms with more media coverage tend to experience a larger drift effect, though unlike Chan (2003), Vega's (2006) results do not confirm Daniel et al.'s (1998) behavioural theory that stipulates investors underreact to public information and overreact to private information. Instead, Chan's results support Brav and Heaton's (2002) theory of rational uncertainty that argues the distribution of information matters more than its type, such that a high arrival rate of informed traders leads to less uncertainty and smaller drifts. It also contends that the more public or private information investors have that they can agree upon, the smaller the drift effect. This naturally leads to larger, more transparent stocks experiencing less drift than small stocks.

Tetlock et al. (2008) observe that investors tend to underreact to textual cues on the knee-jerk, possibly due to skepticism or simply taking the information for granted, thereby creating opportunities to generate abnormal returns in the following trading days. Building on this result, Sinha (2015) finds that the ensuing price drift lasts much longer than the three days reported by Tetlock et al. (2008), implying that markets are even less efficient at absorbing news-based information. A weekly qualitative information (WQI) measure constructed from a sample of pre-scored Thomson Reuters news articles released between 2003 and 2010 reveals that the effect of positive articles lasts 13 weeks from initial publication, whereas that of negative articles lasts 52 weeks. This is in line with Hong et al.'s (2000) claim that "bad news travels slowly" by virtue of firm managers actively pushing out good news while trying to stall bad press. The latter hypothesis is validated using Hong and Stein's (1999) gradual-information-diffusion model and shows that as firm size and analyst coverage decline, the profitability of momentum strategies increases and the gap in reaction time between good and bad news widens.

Taking a closer look at the difference between positive and negative news, Dzielinski (2012) finds that negative and neutral news help resolve asymmetric information and reduce return reversals, while positive news does not. This is in part due to the attribution bias companies exhibit when conveying negative news — negative releases tend to be longer in length and contain fewer company-relevant words than positive releases, since companies are naturally inclined to water-down and explain away bad press. News agencies tend to strip out these irrelevant items and bring the negative aspects of the news to the forefront, thus increasing their concentration and impact. Interestingly though, newswires fail to properly correct for exaggerated positive news, suggesting a certain reluctance to interfere with corporate success stories, possibly due to their need to maintain good working relationships with corporations. This makes positive news less believable and prone to stronger return reversals than negative or neutral news. In terms of proxying for the level of informativeness, Dzielinski (2012) finds that high stock turnover on news days reduces reversals to an even greater extent, even for heterogeneous news announcements that exclude earnings releases. The impact of news on reversals appears to vary significantly with firm size, past returns, institutional ownership, and to some extent, story length and number of news items, though the bar for additional updates to impact returns is quite high.

2.4. Sentiment analysis during the financial crisis and periods of extreme positioning

Smales (2014a) examines the relationship between the tone of unscheduled news, as measured by RPNA sentiment scores, and changes in implied volatility, as measured by the CBOE Volatility Index (VIX), for stocks listed on the S&P500 Index between 2000 and 2010, finding a significant, asymmetric negative relationship between aggregated sentiment for the S&P500 and the VIX that is most pronounced during periods of negative news. The recession portion of their sample, from 2007 to 2009, proves to be an exception, in that the VIX showed more sensitivity to positive news, possibly due to the reduced marginal impact of bad news during this prolonged period of downbeat sentiment. Unlike Groß-Klußmann and Hautsch (2011) and Smales (2014b), who document near instantaneous reactions to firm-specific news, Smales (2014a) sees the relationship between firm sentiment and volatility weaken as the time intervals are shortened from one day to one hour to five minutes respectively, suggesting that aggregated sentiment measures take longer to be factored into market prices than individual stock sentiment. Smales (2015a) also makes important realizations regarding how a portfolio’s response to news might change depending on the level of investor fear; results show that markets are less driven by industry-specific news triggers during recessions, with the predictability of returns more closely tied to systemic factors instead represented by general financial news, consistent with Garcia (2013).

Allen, McAleer and Singh (2015) use an augmented Fama and French three-factor model to demonstrate that aggregated TRNA sentiment scores from 2006 to 2012 have a significant impact on the daily returns of the DJIA and its constituent firms, with the impact being most pronounced for losers, notably during the financial crisis. Sentiment scores lagged up to five days were found to be significant for some companies, suggesting that even in closely scrutinized markets, stock prices are not perfectly informationally efficient. Evidence that sentiment scores are a significant addition to traditional factor models such as the CAPM was further confirmed by Cahan et al. (2009) and Hafez and Xie (2012b) using RPNA sentiment data.

Allen, McAleer and Singh (2013) make use of TRNA data to measure the joint information content of aggregated sentiment scores for DJIA constituent stocks and their respective returns using non-linear, non-parametric entropy-based measures. Entropy metrics have been used as an alternative measure of uncertainty and dispersion that requires fewer constraining assumptions. Their results point to greater certainty in both the DJIA and TRNA sentiment measures during the financial crisis than over the broader sample period. This led to stronger herding behaviour and lower entropy values (given less uncertainty regarding how negative the news was), in turn making returns slightly more predictable during this period.

Smales (2014c, 2015b) is the first to test the application of TRNA sentiment scores to commodity-based news, focusing on the impact of news sentiment on gold futures. Smales finds a similar asymmetric reaction function to the one well-documented in equity analysis, namely that news sentiment has a significant impact on returns and volatility, with negative news having a stronger impact than positive news. Smales takes the analysis a step further by demonstrating that the net positioning of speculators and hedgers significantly impacts the news-returns relationship, such that traders react more to contrary news when positioning is at extremes, due to institutional constraints imposed on them via margin calls and position limits. This effect stems from the influence of the business cycle on net positioning, with negative news having a more pronounced effect during recessions, consistent with Garcia (2013). Borokova and Mahakena (2015) similarly focus on the impact that aggregated TRNA news has on natural gas futures prices between 2006 and 2010. They observe strong negative price trends on days with extremely negative news sentiment, and mean-reverting tendencies on days following extremely positive news sentiment. They further show that news sentiment Granger causes positive and negative price jumps. Lastly, they demonstrate that incorporating news sentiment into volatility models leads to superior forecasts.

2.5. Novelty and the information content of news

With newswires worldwide competing to cover the same breaking stories, novelty across the full scope of global news is virtually impossible to measure, and discrepancies in individual reaction times vary according to which newswires or information sets investors subscribe to. Groß-Klußmann and Hautsch (2011) find that updated TRNA news is traded more actively than "new" news, while Smales (2014b) finds evidence of significant abnormal returns up to 15 minutes before news announcements, potentially attributable to timelier channels, information leakage, or a clustering effect. Tetlock (2011) demonstrates that the magnitude of market reversals to news is positively related to its degree of staleness, proving that novelty is not merely a proxy for irrelevant news, but for information content as well. Staleness tends to be lowest on Mondays and highest on Fridays, with DellaVigna and Pollet (2009) going so far as to claim that readers pay less attention to news on Fridays. Tetlock's evidence indicates that individual investors have a harder time than institutional investors at distinguishing between new and stale information, leading them to trade more aggressively on old news.

Building on this, Huynh and Smith (2013) are the first to examine the joint impact of novelty and tone on weekly momentum returns across 21 developed countries, providing strong international evidence in favour of behavioural explanations for the momentum phenomenon. Using pre-scored TRNA data for 34,000 firms between 2003 and 2011, they demonstrate that the profitability of a weekly momentum strategy that buys winners with stale positive news and sells losers with novel negative news holds internationally, suggesting that investors worldwide are similarly biased towards underreacting to news, consistent with Hong and Stein (1999). However, contrary to Hong et al. (2000), they find that markets tend to underreact to positive news more than negative news when earnings and mergers news are excluded from the sample, leading to momentum strategies becoming more profitable as firm size increases, since large firms tend to have more positive, stale coverage.

2.6. *Market efficiency and models of informed trading*

The potential for such drawn out reactions to new and stale news raises a number of questions about market efficiency and information asymmetry. Grossman and Stiglitz (1980) argue that there is "a fundamental conflict between the efficiency with which markets spread information and the incentives to acquire information", such that if market prices were perfectly efficient and information costly, no one would be willing to pay for said information in the first place since they would have no ability to profit from it, thus leaving prices inefficient. This catch-22 supports Grossman and Stiglitz's (1980) paradoxical notion of an "equilibrium degree of disequilibrium", one that balances the heterogeneity of investor intelligence and skill with the spectrum of possible interpretations of the news. If we assume, to paraphrase Sinha (2015), that the meaning investors derive from the tone of news constitutes private information, Kyle's (1985) model of speculative trading could explain why an informed trader, having full knowledge of their informational advantage, would seek to maximize profits by slowly accumulating a position in an asset as opposed to locking in a worse average price by trading on private information too aggressively and moving the market in the process. This idea that informed traders can use news events to gain a competitive edge over uniformed traders was further confirmed by Engelberg et al. (2012), whose study of short-sellers shows that news releases actually increase information asymmetry, rather than level the playing field. Boehmer et al. (2008) also confirm that institutional short-sellers are among the most well-informed market participants since their trades lead to permanent shifts in market prices, thus making them important in maintaining market efficiency.

Foster and Viswanathan (1996) extend Kyle's model of a monopolistic trader to an oligopolistic setting, where the private information held by traders is heterogenous and not very correlated, characteristics that more closely resemble the natural flow of news. Under this model, competition is weaker and players trade less aggressively, instead engaging in a sort of "waiting game" that leads to the gradual diffusion of private information signals. However, when these same speculators have the technological ability to trade very quickly, Foucault et al.'s (2016) dynamic model of high-frequency news-based trading makes more intuitive sense in explaining why traders with a speed advantage would react more aggressively to news triggers. The model splits the speculator's return into a short-run volatility component driven by impending news and a long-run drift component that reflects the speculator's long-run price forecast. Speculators deviate from their long-run views to exploit short-run news-driven volatility — a strategy that is only profitable when the speculator can react faster than the dealer. The model predicts that as the informativeness of the news increases, the knee-jerk price reaction will be faster, as speculators trade more aggressively and represent a larger portion of trading volume. Under these assumptions, stocks with more informative news should attract more directional high-frequency traders, since being fast has more value here. At the same time, the increased informativeness of the news also puts the dealer at less risk of losing money on changes in long-run price forecasts, thus leading to increased market liquidity. Hence, when speculators are fast and the news is highly informative, Foucault et al.'s model predicts a joint increase in informed trading, trading volume and liquidity.

From this perspective, it could be argued that a positive externality of pre-scored news is a reduction in information asymmetry. By homogenizing investors' interpretation of textual data using standardized metrics such as relevance and tone, the level of disagreement between traders is significantly reduced, making the profitability of new-based strategies more a matter of speed than manual processing skills. Pre-scored news should therefore also result in asset prices taking less time to react to news once it gets released, since speculative traders would be forced to react to new information as quickly as possible in an effort to secure the best average price for their position before their competitors pile into the market. Zhang (2017) provides evidence to support this, and Foucault et al.'s (2016) findings, by showing that an index arbitrage strategy relying on machine-readable news results in improved price efficiency and quicker price discovery, with news-based trading peaking during highly volatile periods. Storkenmaier et al. (2012) also confirm a strong increase in trading activity in response to public information, but contrary to Foucault et al.'s model, they observe significantly lower liquidity on negative news days and no change in liquidity on positive news days.

In a similar vein, Groß-Klußmann and Hautsch (2011) note that while the most relevant TRNA news items for 39 stocks on the London Stock Exchange cause clear responses in firm-specific abnormal returns, volatility, trading volume, average trade sizes, bid-ask spreads, trade imbalances, and market depth, only volatility and trading volume have significant conditional relationships with the arrival of news. Similarly, Smales (2014b) concludes that trading activity and volatility are the most sensitive measures to the arrival of highly relevant RPNA news items, while Shi et al. (2016) use RPNA data to confirm that the intensity and tone of news contribute to the negative relationship between idiosyncratic volatility and expected returns. These results echo earlier work by Kalev et al. (2004), who demonstrate that volume and volatility are jointly influenced by the flow of news, consistent with Clark's (1973) and Tauchen and Pitt's (1983)

Mixture of Distribution Hypothesis (MDH).[5] They show that the number of firm-specific news announcements released over a set time interval acts as a superior proxy for the arrival rate of new information than volume, and found the news variable to have a more pronounced positive impact on the conditional variance of stock returns at shorter intervals of 30 minutes to one hour, confirming the importance of intraday news in explaining return volatility. Tetlock (2010) also found that news days coincide with elevated correlations between absolute returns and volume, with a cross-sectional analysis revealing that news resolves asymmetric information, especially in small and illiquid stocks.

While the strength of the relationship between news and market reactions bodes well for reducing information asymmetry, it also increases the risk of erroneous information becoming rapidly incorporated into market prices, potentially exacerbating market disruptions. This phenomenon was tested and confirmed by von Beschwitz et al. (2015), who show that misclassifications of news due to algorithmic errors can lead to distortionary price movements comparable to "mini" flash crashes, unnecessarily increasing market volatility and instability. By comparing the original and revised versions of the relevance scores generated by RavenPack for Dow Jones Newswire, they are the first to establish a causal relationship between news analytics and stock prices. They observe that low-relevance articles incorrectly released as having high relevance caused a knee-jerk reaction in stock prices that retraced after 30 seconds, and find a 10 per cent more pronounced five-second stock price reaction to news correctly labelled as highly relevant, as compared to highly relevant news incorrectly labelled as having low relevance. Trade volumes also spiked during these knee-jerk reactions, consistent with low-latency traders exploiting their informational advantage. These results suggest that while news analytics suites can render markets more informationally efficient, they also encourage herd behaviour and can, by the same token, increase information asymmetry due to only a small subset of the trading population having access to such sophisticated metrics.

---

[5] The Mixture of Distribution Hypothesis implies price action is driven by news trading, and uninformed traders act on large movements, leading volatility and volume to move together.

3. Data and Descriptive Statistics

3.1. Overview

The Thomson Reuters News Analytics (TRNA) database uses proprietary algorithms to generate numerical scores for firm-specific news. The indicators computed by the TRNA news engine include:

- A relevance score between 0 and 1 that measures how substantive the news is for a given company, with 1 being the most relevant.
- Three sentiment scores between 0 and 1 that measure the probability of the news being positive, negative, or neutral in tone, the sum of which must equal 1.
- A series of five novelty scores that count the number of stories with a similar lexical fingerprint released in the past 0.5, 1, 3, 5, and 7 days, that provide a gauge of how stale the news is, with 0 being the most novel.
- Parent and child topic codes, organized as a hierarchy of keywords used to describe the subject of the news (e.g., TECH is the parent code for technology, while SOFW is the child code for software).
- Sentence and word counts for the underlying text accompanying the headline or story title.

When combined, these fields of metadata are meant to act as a signaling tool for algorithmic trading systems. At its most basic level, it should help in organizing vast amounts of unstructured news data, enabling end-users to pick out only the most likely market-moving triggers from a universe of noise, thus eliminating the need for tedious and time-consuming manual filtering. In a best case scenario, it could offer a low latency solution for systematically trading firm-specific news.

The primary focus of this study is to test the effectiveness of employing such an augmented news feed for the purpose of identifying significant, unscheduled, firm-specific short-term trading opportunities. The analysis is conducted on a five-year sample of Nasdaq-100 news from the TRNA database and focuses on employing the sentiment, relevance, and novelty scores.

The TRNA dataset spans from January 1, 2011 until December 31, 2015, inclusively. Over this five-year span, Thomson Reuters released over eight million unique pieces of equity news containing nearly 12-million flagged mentions of companies for which analytical news scores were generated. Table 1 reports the yearly breakdown of news for the entire dataset as well a subset of Nasdaq news. Flagged mentions, which will herein be called ‘news items’, refer to instances where the TRNA algorithm recognizes a company as the subject of the news and computes the corresponding sentiment, relevance and novelty scores for that company as the headline or article relates to it. Since multiple companies can be mentioned in the same piece of news, the dataset contains 42 per cent more news items than unique pieces of news.

Table 1

Number of TRNA News Items and Unique Pieces of News by Year

Year	All News Items	Unique Pieces of News	NASDAQ News Items	NASDAQ Unique Pieces of News
2011	2,209,813	1,541,705	97,066	76,423
2012	2,202,063	1,485,028	97,787	71,377
2013	2,184,574	1,468,846	107,292	74,700
2014	2,443,546	1,748,665	95,949	69,650
2015	2,793,247	2,112,411	102,266	80,895
Total	11,833,243	8,356,655	500,360	373,045
	142%	100%	134%	100%

These figures summarize the volume of all equity news and Nasdaq-specific news released by Thomson Reuters from 2011 to 2015. News items refer to database entries for which the TRNA algorithm identified a company as a subject of the news and triggered the computation of sentiment, relevance, and novelty scores. A unique piece of news can contain multiple news items depending on how many companies are mentioned in it.

The TRNA news dataset also contains a unique identifier that links different news items to a specific story. A single story can come to contain multiple news items as it evolves over time; breaking news is initially reported as one or a series of headlines, followed by articles and updates as more information becomes available. Over the five-year horizon of the sample, 80 per cent of all news items belonged to stories that were released only once, while 90 per cent belonged to stories that were released no more than three times, as shown in Table 2.

A typical story contains news items classified as either Alerts or Articles. Alerts are breaking news headlines that are time-sensitive in nature and typically only tell the reader the essential facts in 100 characters or less. As their name suggests, Alerts purportedly have the most market-moving potential on the knee-jerk since they pack a punch by being among the most novel, relevant, and attention-grabbing bits of news (sometimes to the point of being sensationalized). Articles, on the other hand, are longer, less timely accounts of the story at hand, but include a lot more colour and detail than the initial Alert(s) can fit into a banner. This completeness naturally comes at the expense of timeliness, thus making Articles less pressing but more complete, and arguably more accurate, than Alerts.

According to the Thomson Reuters Journalist Handbook, a story tends to break with one or a burst of Alerts, followed by one or a sequence of Articles that add the necessary context underlying the initial Alert(s). As shown in Figure 1, Articles are by far the most prevalent story type in the dataset, comprising 68 per cent of all news items, while Alerts account for just 30 per cent. Articles also have their own set of sub-categories that describe what type of Article is being published (e.g., an update, a correction, an interview, etc.). An exhaustive list of these sub-categories is not readily available; however, I define one based on the various types discussed in

Table 2

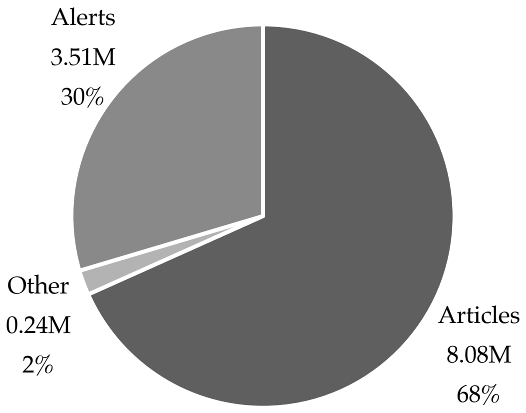
Summary of TRNA Story Frequency

Number of Times a Story is Released	Number of News Items	Frequency
1	9,473,116	80.10%
2	811,790	6.90%
3	448,726	3.80%
4	280,898	2.40%
5	190,491	1.60%
6	134,559	1.10%
7	99,190	0.80%
8	75,937	0.60%
9	58,253	0.50%
10	45,513	0.40%
11 or more (max. = 99)	213,180	1.80%
Total	11,833,243	100.00%

As a news story evolves, it may come to contain multiple news items. These figures summarize the number of news items (identified by a unique sourceID), that make up a given news story (identified by a unique story number called altID), based on the number of releases associated with the story.

Figure 1

Number of News Items by Type



Total Number of News Items: 11.83M

News items are classified as either Alerts or Articles. Alerts are breaking news headlines, while Articles are less timely, more detailed accounts that give context to the initial Alerts. A typical story breaks with one or more Alerts followed by one or a series of Articles.

the Thomson Reuters Handbook of Journalism (see Appendix B). An algorithm is subsequently used to flag 22 of the seemingly more important types based on the title of the news item. It's worth bearing in mind that the Reuters editorial tagging system is subject to discretion of each individual journalist, whose tagging conventions and consistency vary worldwide. Therefore, Article categories, much like news topic codes, are only as consistent as the teams of journalists who generate them, a limitation that's virtually impossible to quantify and outside the scope of this analysis.

3.2. Nasdaq-100 Index News

The Nasdaq-100 Index is among the most widely followed stock indices in the world; but, unlike the S&P500 and the DJIA, the Nasdaq-100 is tech-centric, excludes financials, and (similar to the S&P 500) is comprised of both US and international companies. It is also a highly liquid, large cap, market cap-weighted index, making it an appropriate complement to the FTSE 100 Index studied by Groß-Klußmann and Hautsch (2011), the springboard for this analysis.

Using data obtained from Siblis Research, 145 companies are identified as having made an appearance on the Nasdaq-100 Index between 2011 and 2015. Each constituent's add and drop dates are confirmed using press releases from Nasdaq.com (see Appendix C), revealing that 58 stocks were listed on the index continuously for all 60 months, while 40 were present for more than half of the sample horizon, and the remaining 47 appeared for 30 months or less.

For each stock in the sample, the Reuters Instrument Code (a unique identifier called RIC) is identified using known samples of firm-specific news and cross-checked against the company name associated with each RIC code in Thomson Reuters Tick History database. The persistence of each RIC code is verified over the course of the sample horizon and updated for any ticker symbol changes resulting from M&A activity or spinoffs, which typically lead to new RIC codes being generated. Given that a single company can have multiple RIC code suffixes, the sample is limited to news that uses the Nasdaq exchange suffix (.O) to ensure that news for another company with the same RIC code listed on a different exchange is not extracted accidentally.

There are a total of 500,360 news items for Nasdaq companies during the five-year sample horizon, and these correspond to 373,045 distinct pieces of news that mention one or more Nasdaq companies, a summary of which can be found back in Table 1. This equates to roughly 274 Nasdaq news items released per day over the course of five years, or one to two news items per company per day. As shown in Table 3, the average and median number of news items for each company over the five years are 3,451 and 2,200 respectively, with Apple being the most prevalent company in the news with 43,662 news items, and Joy Global being the least frequently mentioned with 283 items. The volume of news per Nasdaq company is asymmetrically distributed, such that 16 per cent of companies represent 50 per cent of all news items, while the remaining 84 per cent of companies each make up less than one per cent of the news.

The distribution of the Nasdaq news items closely resembles that of the entire dataset, with annual releases proving to be quite stable at around 100,000 per year, as shown in Figure 2a. The totals organized by month fluctuate between 35,000 and 50,000 news items, or around 600 to 800 in each of the 60 months over the five years, but dip noticeably in December, as can be seen in Figure 2c. This is a likely side effect of US markets easing their way into a holiday lull, per Leinweber and Sisk (2011). A similar dip is observed in Figure 2b, as the volume of daily news increases steadily to a peak on Thursdays, but slows down noticeably on Fridays, ushering in the weekend calm.

The intraday distribution in Figure 2d shows that the volume of news is most saturated during the US pre-market open, which isn't surprising considering this is when traders get caught up on overnight events as they gear up to start their trading day. The most pronounced spikes occur right before the market open, between 8:00am and 9:00am, and immediately following the 4:00pm close. The latter can be partially attributed to companies preferring to release sensitive information after markets close to reduce price volatility, and partly due to Thomson Reuters rerunning recent news stories on client terminals in different time zones. Notice that the US close bleeds into the Asian pre-market open, and that the gentle uptick in news around 2am precedes the European open, further reinforcing the notion that news gets rehashed multiple times throughout the day for client convenience worldwide. Apart from the surges before and after the US market open, Nasdaq news is released rather steadily throughout the trading session, with a total of 10,000 to 15,000 news items out every hour between 9:30am and 4:00pm, which is equivalent to 40 to 60 pieces hourly over the five years. The number of news releases tapers off gradually in the late after hours, before coming to a lull overnight.

Table 3

Frequency of Nasdaq News Items for Companies on the Index between 2011 and 2015

Ticker	News Items	Frequency
AAPL	43,662	8.73%
MSFT	22,413	4.48%
AMZN	17,610	3.52%
FB	17,397	3.48%
FOXA	13,741	2.75%
VOD	12,252	2.45%
INTC	11,904	2.38%
YHOO	11,504	2.30%
NFLX	9,651	1.93%
CMCSA	9,354	1.87%
EBAY	8,277	1.65%
ORCL	7,902	1.58%
CSCO	7,535	1.51%
DELL	6,764	1.35%
RIMM	6,387	1.28%
QCOM	6,211	1.24%
TSLA	5,696	1.14%
INFY	5,680	1.14%
VIAB	5,596	1.12%
AAL	5,391	1.08%
SBUX	5,209	1.04%
DISH	5,165	1.03%
TEVA	5,032	1.01%
MYL	4,951	0.99%
AMGN	4,886	0.98%
ADP	4,675	0.93%
TXN	4,461	0.89%
EQIX	4,448	0.89%
MU	4,362	0.87%
GILD	4,331	0.87%
COST	4,179	0.84%
MDLZ	4,147	0.83%
MAR	4,111	0.82%
SHLD	3,980	0.80%
BIIB	3,849	0.77%
FSLR	3,747	0.75%
BRCM	3,712	0.74%
CELG	3,489	0.70%
BIDU	3,465	0.69%
EXPE	3,446	0.69%
SPLS	3,413	0.68%
VIP	3,334	0.67%
GOLD	3,247	0.65%
DTV	3,139	0.63%
PCLN	3,134	0.63%
NVDA	3,113	0.62%
CHTR	3,095	0.62%
ADBE	3,033	0.61%
TMUS	3,014	0.60%
GMCR	2,986	0.60%
EA	2,965	0.59%
AMAT	2,910	0.58%
SNDK	2,836	0.57%
ATVI	2,750	0.55%
DLTR	2,635	0.53%
MRVL	2,575	0.51%
WYNN	2,562	0.51%
STX	2,512	0.50%
DISCA	2,485	0.50%
ESRX	2,447	0.49%
WDC	2,433	0.49%
CTXS	2,431	0.49%
SYMC	2,403	0.48%
REGN	2,402	0.48%
ILMN	2,322	0.46%
FFIV	2,318	0.46%
AVGO	2,307	0.46%
ADSK	2,298	0.46%
NTAP	2,293	0.46%
LBTYA	2,271	0.45%
ALTR	2,267	0.45%
MAT	2,262	0.45%
APOL	2,200	0.44%
INTU	2,118	0.42%
CTSH	2,079	0.42%
WFM	2,079	0.42%
VRTX	2,052	0.41%
PRGO	2,030	0.41%
GOOGL	1,994	0.40%
AKAM	1,985	0.40%
ADI	1,961	0.39%
URBN	1,931	0.39%
NXPI	1,912	0.38%
ALXN	1,822	0.36%
BBBY	1,805	0.36%
TRIP	1,751	0.35%
KHC	1,750	0.35%
SWKS	1,744	0.35%
ENDP	1,732	0.35%
ROST	1,729	0.35%
CA	1,708	0.34%
BMRN	1,663	0.33%
XLNX	1,623	0.32%
QGEN	1,597	0.32%
GRMN	1,566	0.31%
MCHP	1,481	0.30%
LMCA	1,476	0.29%
KLAC	1,464	0.29%
LRCX	1,461	0.29%
VMED	1,459	0.29%
FISV	1,449	0.29%
NUAN	1,427	0.29%
QVCA	1,415	0.28%
FOSL	1,400	0.28%
LIFE	1,397	0.28%
SIRI	1,306	0.26%
CHKP	1,269	0.25%
CERN	1,265	0.25%
ISRG	1,252	0.25%
CHRW	1,235	0.25%
MXIM	1,210	0.24%
VRSK	1,187	0.24%
WCRX	1,165	0.23%
CTAS	1,161	0.23%
MICC	1,160	0.23%
PCAR	1,160	0.23%
CTRP	1,142	0.23%
FLEX	1,115	0.22%
BMC	1,098	0.22%
SBAC	1,081	0.22%
INCY	1,055	0.21%
ORLY	1,024	0.20%
CTRX	1,010	0.20%
PAYX	1,007	0.20%
TSCO	978	0.20%
HSIC	972	0.19%
ULTA	926	0.19%
SIAL	899	0.18%
JD	893	0.18%
LLTC	884	0.18%
VRSN	802	0.16%
NIHD	797	0.16%
FAST	727	0.15%
NCLH	716	0.14%
XRAY	697	0.14%
FLIR	681	0.14%
GENZ	642	0.13%
CEPH	613	0.12%
STRZA	608	0.12%
EXPD	578	0.12%
WBA	549	0.11%
SRCL	490	0.10%
MNST	397	0.08%
PYPL	302	0.06%
JOYG	283	0.06%
Total	**500,360**	**100.00%**
Average	**3,451**	**0.69%**
Median	**2,200**	**0.44%**

All 145 stocks that were on the Nasdaq-100 Index at some point between 2011 and 2015 are organized in order of decreasing number of news items along with the frequency, which represents each Nasdaq stock's weight in the sample. The volume of news is asymmetrically distributed, such that 16% of companies represent 50% of all Nasdaq news items, while the remaining 84% of companies each make up less than 1% of the news.

Fig. 2a

Annual Distribution of Nasdaq News Items

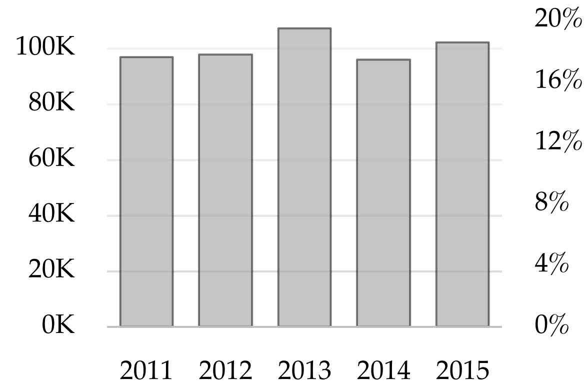


Fig. 2b

Total Count of Nasdaq News Items by Day of the Week

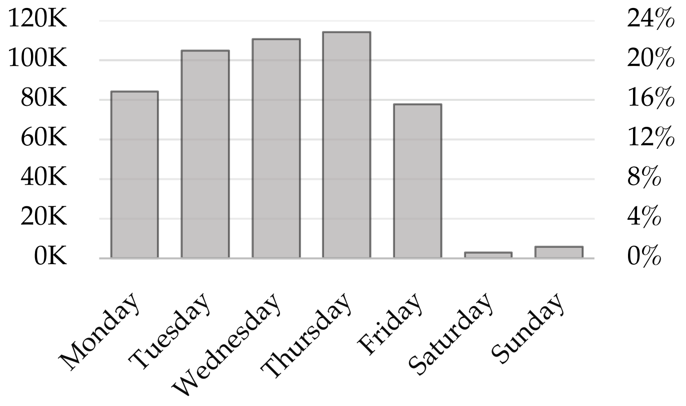


Fig. 2c

Total Count of Nasdaq News Items by Month

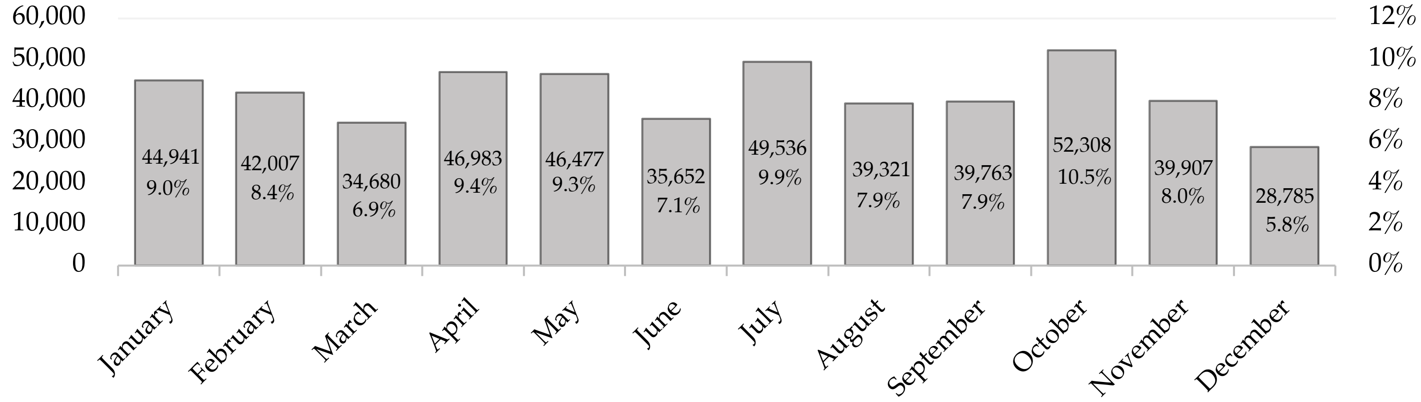
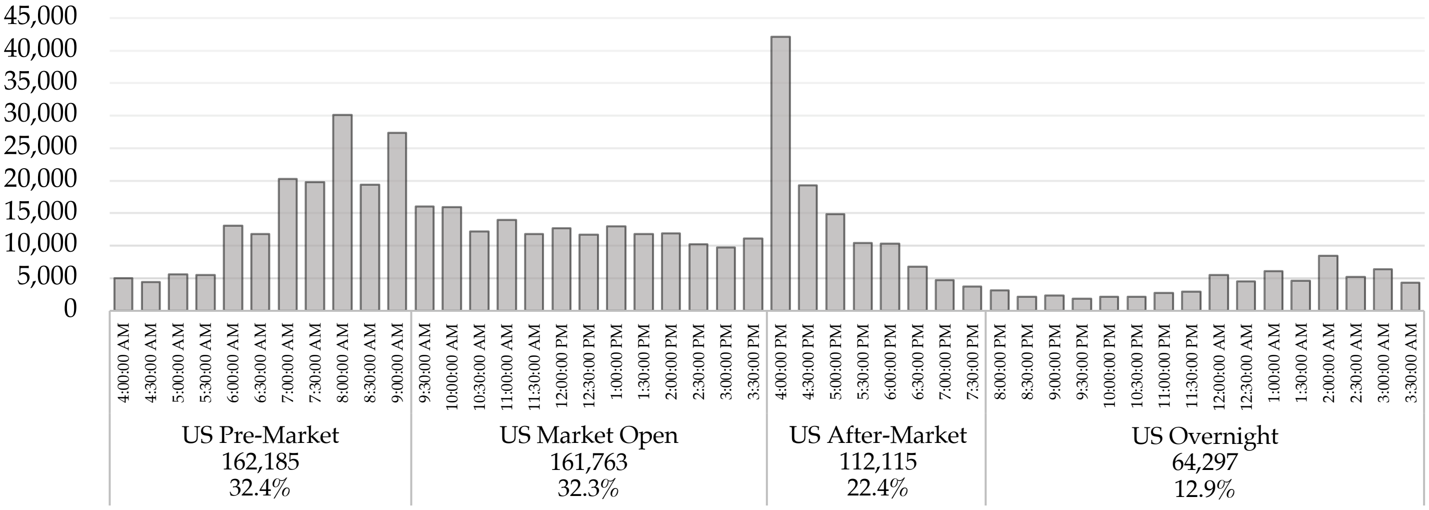


Fig. 2d

Total Count of Nasdaq News Items by Trading Period



Note that the total counts in Figures 2b, 2c, and 2d are cumulative tallies based on the entire sample of 500,360 Nasdaq news items for 145 companies identified as constituents at some point between 2011 and 2015 inclusively, the list of which is shown in Table 3.

3.3. News Traffic

In the context of high-frequency trading, it warrants looking beyond the big picture of how news is released and examine how news flows on a micro level. The idea of news traffic as a measure of momentum can help conceptualize the flow of news, making it easier to understand how the volume and pattern of releases affect a trader's, and by extension, the market's, ability to digest information.

To measure news traffic, unique news stories are grouped into one-second buckets for which a series of 38 different variables are computed for each target story in order to quantify the distribution of news released in the three-hours immediately before and after its release. The number of one-second timestamps with news are counted in both the Before and After windows, in addition to the number of Events (unique news releases) contained in each of those one-second timestamps, since it's not uncommon to find "bursts" of news, defined as instances where more than one story is released in the same one-second interval. It follows that the more saturated the news flow in each of the Before and After Windows, the higher the level of noise and distraction around the target news release. Nasdaq news is typically released in median bursts of two stories, or four and a half on average, and as Table 4 reports, the release in closest proximity to the target typically comes out 20 seconds beforehand on weekdays. Figure 3 illustrates with more granularity the distribution of seconds since the previous story, confirming the near continuous

Table 4

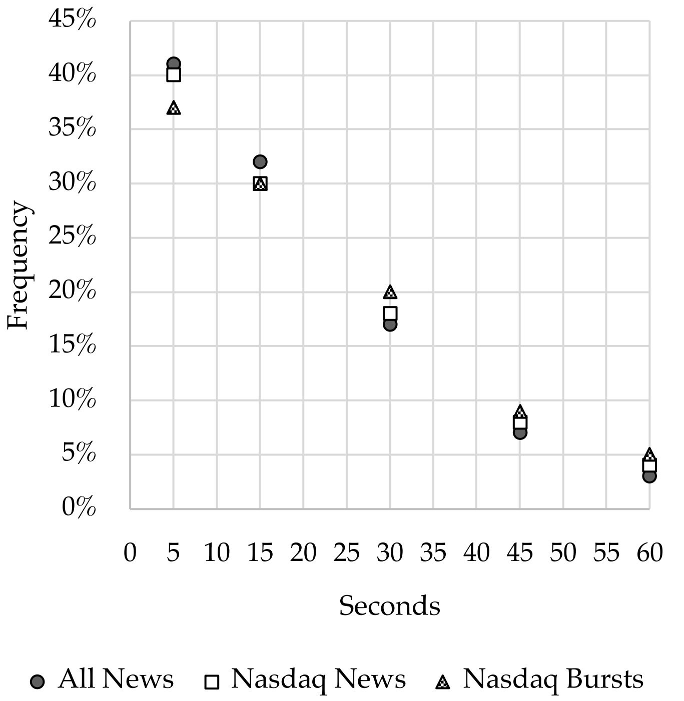
Summary of Traffic Variables for Nasdaq News Stories by Day of Week

	Target Timestamp Averages		Before Window Averages *3 hours before target timestamp*		
	Event Count	Seconds Since Previous Story	Event Count	Mean Seconds Between Events	Median Seconds Between Events
Monday	4.6	21	1,398	11	1.7
Tuesday	4.4	19	1,494	9	1.4
Wednesday	4.3	19	1,500	9	1.4
Thursday	4.3	18	1,640	9	1.4
Friday	4.8	28	1,258	12	1.9
Saturday	3.3	704	153	275	90.4
Sunday	5.0	623	112	252	39.8
All Days	4.5	30	1,456	14	2.3
Excluding Weekends	4.5	20	1,476	10	1.5

Target timestamps refer to the one-second interval containing the Nasdaq news story being analyzed, while the Before window refers to the three hours immediately preceding the target timestamp. Event Count captures the number of unique stories released in the same one-second timestamp, and are not limited to Nasdaq stories. "Bursts" of news occur when Event Count/Timestamp Count > 1. Seconds since previous story measures proximity in time of the last news story relative to the target, while the mean and median seconds between Events summarizes the time intervals between all the Event Counts in the Before window.

Figure 3

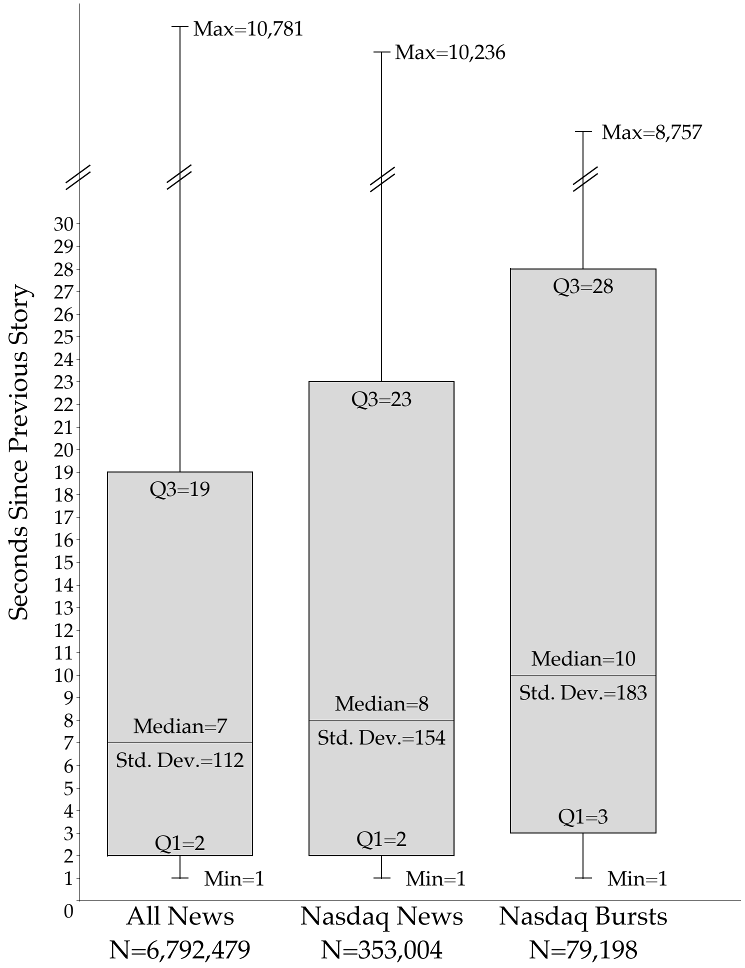
Frequency of Seconds Since Previous Story



News is released within five seconds of the target story 40% of the time, and within five to 15 seconds 30% of the time, emphasizing the near-continuous nature of the news dataset. These figures remain constant even as the traffic sample is narrowed from All News to Nasdaq News to Nasdaq Bursts.

Figure 4

Box Plot of Seconds Since Previous Story



The distribution of seconds since the last story was released (relative to the target story) shows increasing dispersion as the traffic sample is narrowed from All News to Nasdaq News to Nasdaq Bursts.

nature of the news dataset. The time since the previous story's release is within five seconds of the target story 40 per cent of the time, within five to 15 seconds roughly 30 per cent of the time, within 15 to 30 seconds 20 per cent of the time, within 30 to 45 seconds 10 per cent of the time, and within 45 to 60 seconds five per cent of the time. These frequencies remain fairly constant for All News, Nasdaq News, or Nasdaq Bursts; however, the box plot in Figure 4 shows how the standard deviation of the 'Seconds Since Previous Story' increases for more specific samples, reflecting increased dispersion.

Since most stories in the sample are released in close succession, it suffices to report the relevant statistics for the Before windows only, since the After windows overlap >90 per cent of the time. On average, 1,476 stories are released in the three hours preceding the target Nasdaq timestamp on weekdays, with minimal variation in their pace and timing from Monday to Friday (see Table 4). The positions in time of each of these Before window Events is recorded both relative to one another and the target news release, and are used to generate a series of statistics characterizing the momentum of news in the run-up and aftermath of a story being released. For Nasdaq news released on weekdays, Table 4 shows that stories in the Before window tend to come out every 10 seconds; however, the average median time between them is much shorter, at just one and a half seconds, likely indicating the presence of many bursts of stories. The lower the median time between stories, the more congestion there is around the target news release, and the wider the interquartile range (IQR) of time between stories, more dispersion there is between consecutive stories. A more erratic news flow is less predictable and therefore more likely to catch traders off-guard, potentially leading to slower reaction times than a steady, orderly pace of news, which offers more opportunities to anticipate the timing of impending releases. Figures 5 and 6 illustrate that the IQR of time elapsed between stories is less than five seconds 10 per cent of the time, five to 10 seconds 46 per cent of the time, and 10 to 15 seconds 25 per cent of the time, and overall shows very little variability for the middle 50 per cent of observations for All News, Nasdaq News, and Nasdaq Bursts.

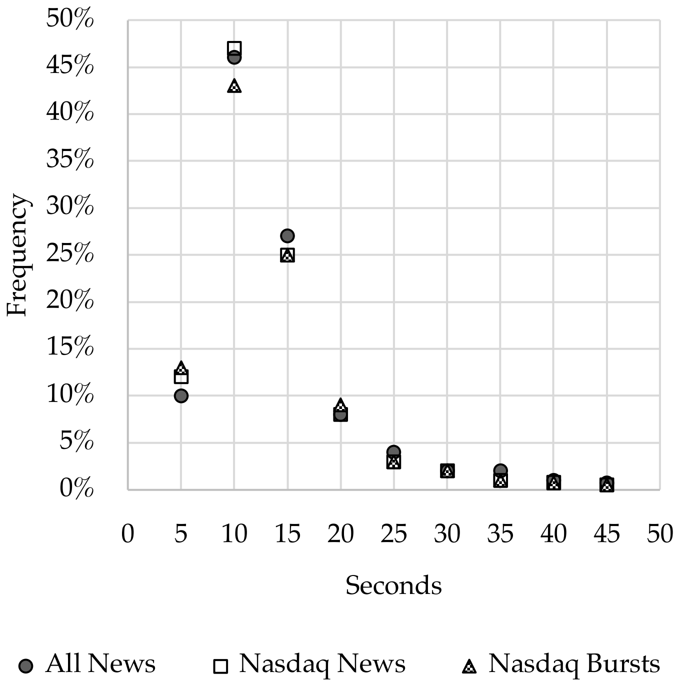
3.4. TRNA Scores

Given the proprietary nature of the TRNA algorithms, our ability to understand the inner workings of how the Sentiment, Relevance, and Novelty scores are generated is limited. However, we can still deduce some useful information from their distributions over time, which will help in structuring framework of the analysis.

According to the distribution of Relevance scores in Figure 7, 56 per cent of news items in the dataset are flagged as being highly relevant to the company identified as the subject of the news, while the remainder are skewed towards being almost completely irrelevant, suggesting the relevance scoring algorithm behaves in an almost binary manner. A similar pattern is observed in Novelty scores, with 45 per cent of news items classified as novel for at least the last seven days, while 43 per cent are stale, with similar news was already released within the last 12 hours.

Figure 5

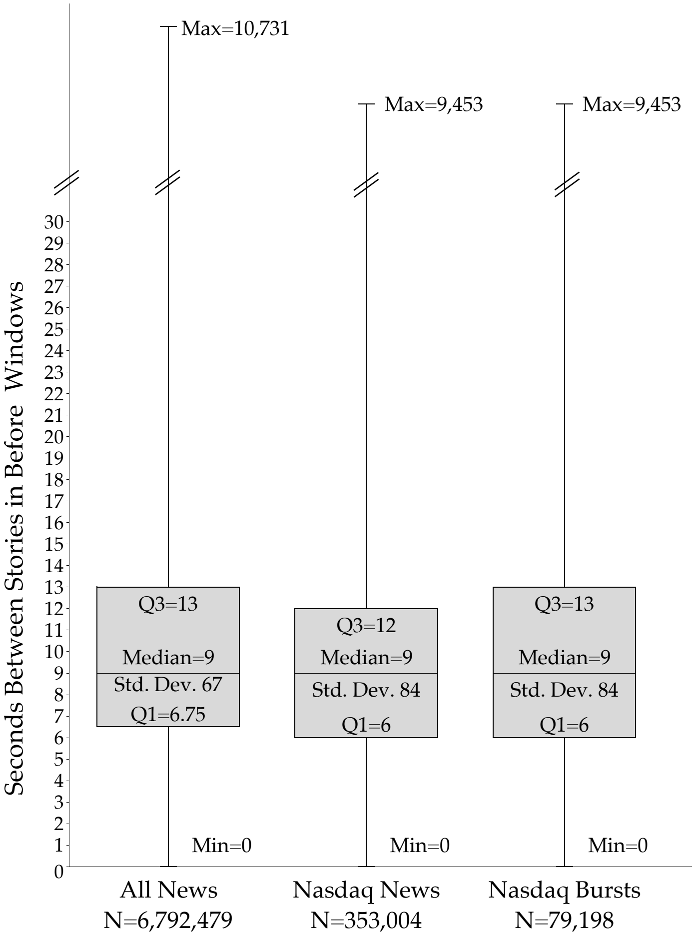
IQR of Seconds Between Story Releases



The interquartile range of seconds between stories released in the three hour window preceding the target news story is less than five seconds 10 per cent of the time, five to 10 seconds 46 per cent of the time, and 10-15 seconds 25 per cent of the time, and does not vary much as the traffic sample is narrowed from All News to Nasdaq News to Nasdaq Bursts.

Figure 6

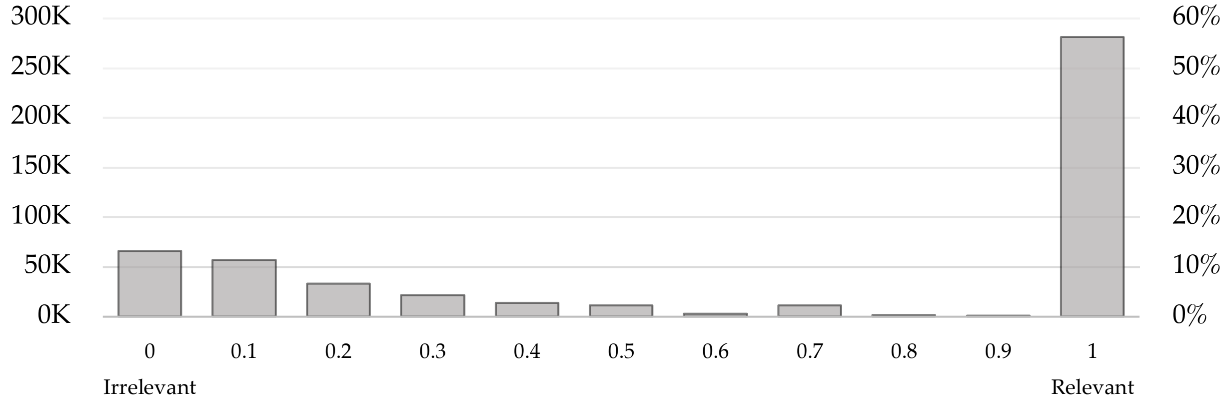
Box Plot of IQR for Seconds Between Story Releases



The distribution of the interquartile range of seconds between stories in the three-hour window preceding the target news release is quite small and fairly constant, hovering between six and 13 seconds as the sample is narrowed from All News to Nasdaq News to Nasdaq Bursts.

Figure 7

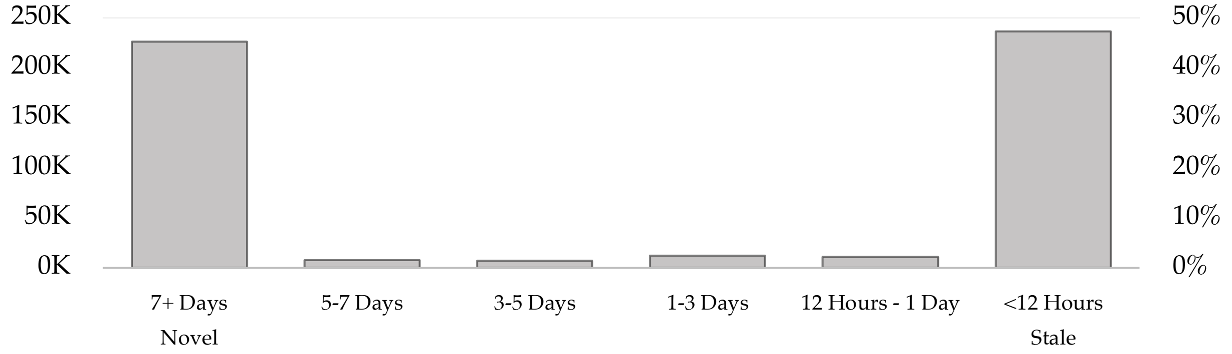
Distribution of News Items by Relevance Score



The Relevance scores for Nasdaq news show that 56 per cent of news is highly relevant (1) to the company flagged as its subject, while the remainder are skewed towards being almost completely irrelevant (0).

Figure 8

Distribution of News Items by Novelty Score



The Novelty scores for Nasdaq news reveal that 45 per cent of news items are classified as novel for at least the last seven days, while 43 per cent saw similar news stories within the last 12 hours or less. Very few news items score in the 12-hour to seven-day range.

The distribution of raw TRNA sentiment probabilities shown in Figure 9 is actually quite balanced; 34 per cent of news items have mildly positive scores, while 32 per cent have mildly negative scores. At the tails, 16 per cent of news items are very positive, while 19 per cent are very negative, however, there are actually twice as many news items with a positive net tone as net negative ones (62 percent versus 32 per cent).[6] This difference echoes the extensive body of

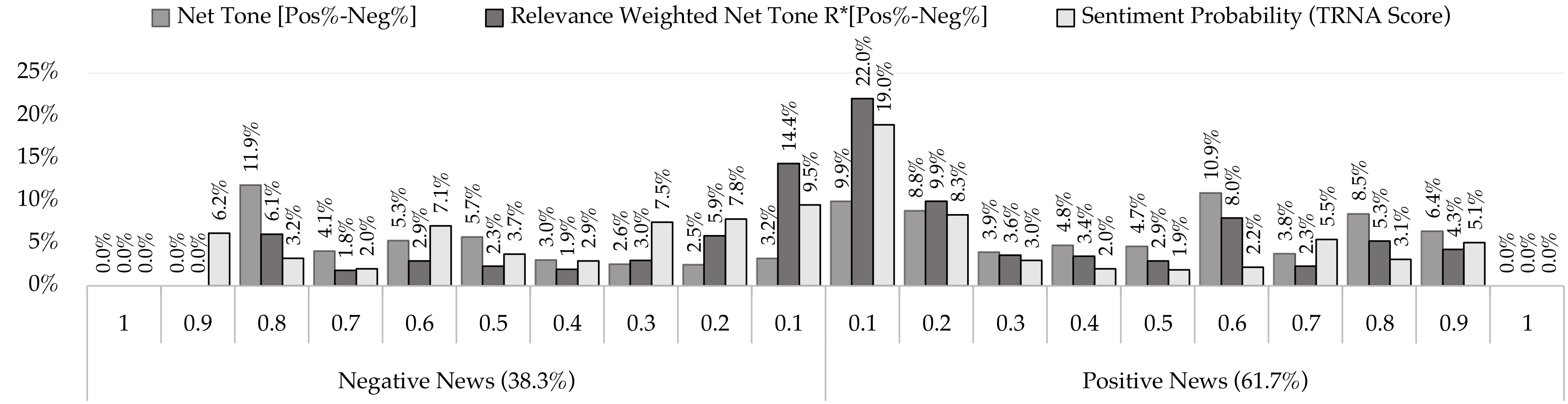
---

[6] Note that a "mild" score is defined as having a probability of being positive or negative of < 50 per cent, while "high" score has a probability of being positive or negative of > 50 per cent. Net tone is calculated as Positive % Sentiment score – Negative % Sentiment score.

research that finds corporate news is biased towards exaggerating the positives and downplaying the negatives (Hong et al., 2000; Dzielinski, 2012; Solomon, 2012). The skew holds even after the net tone of each news item is weighted by its corresponding relevance score, per Huynh, Smith (2013). Only 11 per cent of news items have highly negative relevance-weighted net tone scores, while 20 per cent have highly positive relevance-weighted net tone scores. In fact, Figure 10 shows that the average tone of Nasdaq news over the sample horizon is stubbornly positive at roughly eight per cent, reaching a monthly peak of 13 per cent in June of 2015, and dipping symbolically into negative territory in August off 2011 — which, incidentally, coincides with Black Monday. The difference between the number of positive and negative news items also shows a persistent net surplus of roughly 2000 positive news items per month (see Figure 10).

Figure 9

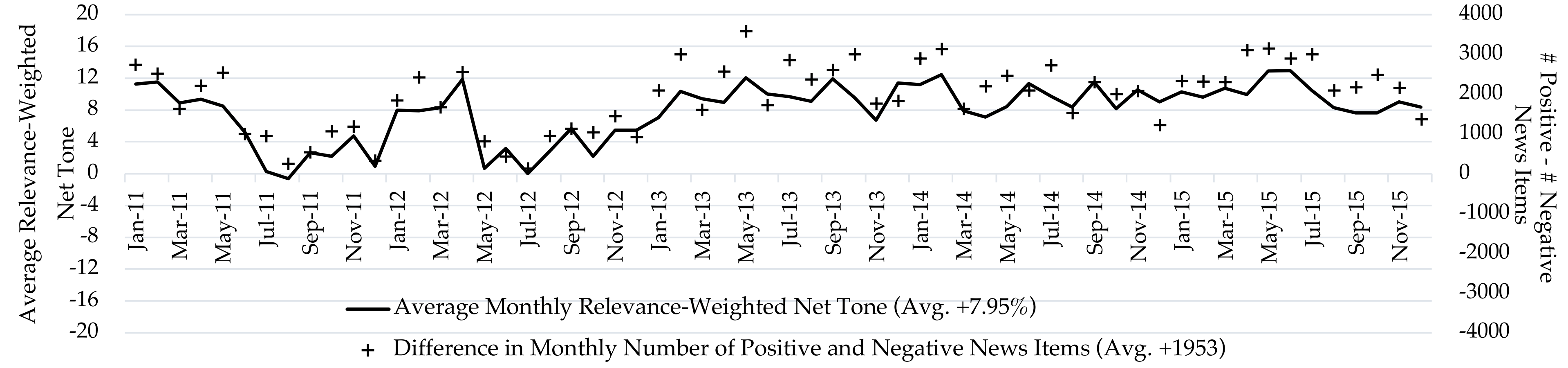
Distribution of Nasdaq News Items by Sentiment Scores, Net Tone, and Relevance Weighted Net Tone



This distribution shows the per cent of Nasdaq news items according to three metrics: Raw TRNA Sentiment Score, Net Tone (Positive % Sentiment Score – Negative % Sentiment Score), and Relevance-Weighted Net Tone (calculated by multiplying the Net Tone by its corresponding Relevance Score).

Figure 10

Monthly Average Relevance-Weighted Tone for Nasdaq News Items



Average monthly tone is measured using the Relevance-Weighted Net Tone for all Nasdaq news items (calculated by multiplying the Relevance Score by Net Tone). It is plotted alongside the monthly difference between the number of net positive and negative news items, both of which show a mildly positive bias to the news.

4. Methodology

4.1. Sample Selection for Nasdaq-100 Index News

Given the heavily skewed distribution of news items towards larger companies in the index, coupled with non-trivial processing constraints, a subset of Nasdaq-100 companies are chosen to act as a proxy for the entire index. All constituents not continuously present on the index for all five years between 2011 and 2015 are dropped, consistent with Tetlock et al. (2008), and any stocks with RIC inconsistencies arising from M&A activity or incomplete price data are removed. The remaining companies are sorted according to their weight in the Nasdaq-100 Index following the May 2, 2011 special rebalancing, and the top 55 stocks are selected to represent 73.5 per cent of the Nasdaq's total market capitalization and 52.2 per cent of all Nasdaq news items over the five-year horizon, as shown in Table 5. These proportions are substantial enough to make this subset a fair proxy for the index as a whole, consistent with the method used by Groß-Klußmann and Hautsch (2011).

Since earnings releases have a well-documented impact on stock prices (Demers and Vega, 2010; Feldman et al., 2010), isolating the effects of unscheduled news makes it easier to evaluate the usefulness of the TRNA database in helping traders navigate the near-continuous stream of spontaneous and heterogeneous news, per Groß-Klußmann and Hautsch (2011). In order to focus exclusively on unscheduled news, all earnings release days are removed from the sample. Quarterly earnings release dates are harvested for each company using a combination of Bloomberg's Earnings History (ERN) function and historical press releases. The resulting list of dates are systematically flagged in each company's respective news items and excluded from the final sample (see Quarterly Earnings Release Dates in Appendix D).

I focus on Nasdaq news released during the market open trading session, since this period presents the best liquidity conditions with which to conduct a high-frequency analysis. Expanding the scope of the analysis to include the pre-market, after-market, and overnight sessions would significantly complicate various computational aspects of the analysis. As an added precaution, the first and last 15 minutes of the trading session are also excluded to avoid volatility associated with the open and close. Dummy variables are generated to identify which of the four trading periods a news item was released in (see Nasdaq Trading Periods table in Appendix D). Attention is paid to daylight savings offsets and transitions when translating each period's start and end times to their corresponding UTC timestamp to match the TRNA dataset. Weekends, holidays and early-closes for the Nasdaq Index are also identified using NasdaqTrader.com and flagged as dummy variables (see UTC Equivalent US Holiday Start and End Dates in Appendix D).

Table 5

Number of News Items for Top-55 Nasdaq-100 Constituents Post-2011 Reweighting

Count	Ticker	Company Name	2011 Weight Post-Rebalance	News Items	Frequency	2011 Cumulative Weight Post-Rebalance	Cumulative News Items	Cumulative News Frequency
1	AAPL	Apple Inc.	12.33%	43,662	8.73%	12.33%	43,662	8.73%
2	MSFT	Microsoft Corp.	8.32%	22,413	4.48%	20.65%	66,075	13.21%
3	GOOGL	Alphabet Inc.	5.77%	1,994	0.40%	26.42%	68,069	13.60%
4	INTC	Intel Corp.	4.20%	11,904	2.38%	30.62%	79,973	15.98%
5	CSCO	Cisco Systems, Inc.	3.66%	7,535	1.51%	34.28%	87,508	17.49%
6	QCOM	Qualcomm, Inc.	3.48%	6,211	1.24%	37.76%	93,719	18.73%
7	AMZN	Amazon.com, Inc.	3.16%	17,610	3.52%	40.92%	111,329	22.25%
8	CMCSA	Comcast Corp.	2.03%	9,354	1.87%	42.95%	120,683	24.12%
9	AMGN	Amgen Inc.	1.92%	4,886	0.98%	44.87%	125,569	25.10%
10	EBAY	eBay Inc.	1.58%	8,277	1.65%	46.45%	133,846	26.75%
11	BIDU	Baidu.com, Inc.	1.46%	3,465	0.69%	47.91%	137,311	27.44%
12	GILD	Gilead Sciences, Inc.	1.32%	4,331	0.87%	49.23%	141,642	28.31%
13	COST	Costco Wholesale Corp.	1.26%	4,179	0.84%	50.49%	145,821	29.14%
14	ESRX	Express Scripts, Inc.	1.15%	2,447	0.49%	51.64%	148,268	29.63%
15	SBUX	Starbucks Corp.	1.08%	5,209	1.04%	52.72%	153,477	30.67%
16	CELG	Celgene Corp.	1.05%	3,489	0.70%	53.77%	156,966	31.37%
17	ADP	ADP, Inc.	1.00%	4,675	0.93%	54.77%	161,641	32.30%
18	PCLN	The Priceline Group	0.98%	3,134	0.63%	55.75%	164,775	32.93%
19	CTSH	Cognizant Technology Corp.	0.97%	2,079	0.42%	56.72%	166,854	33.35%
20	VOD	Vodafone Group plc	0.95%	12,252	2.45%	57.67%	179,106	35.80%
21	YHOO	Yahoo! Inc.	0.86%	11,504	2.30%	58.53%	190,610	38.09%
22	AMAT	Applied Materials, Inc.	0.80%	2,910	0.58%	59.33%	193,520	38.68%
23	PCAR	PACCAR Inc.	0.76%	1,160	0.23%	60.09%	194,680	38.91%
24	BIIB	Biogen, Inc.	0.68%	3,849	0.77%	60.77%	198,529	39.68%
25	NTAP	NetApp, Inc.	0.68%	2,293	0.46%	61.45%	200,822	40.14%
26	ADBE	Adobe Systems Inc.	0.67%	3,033	0.61%	62.12%	203,855	40.74%
27	INTU	Intuit, Inc.	0.65%	2,118	0.42%	62.77%	205,973	41.16%
28	CTXS	Citrix Systems, Inc.	0.55%	2,431	0.49%	63.32%	208,404	41.65%
29	SYMC	Symantec Corp.	0.55%	2,403	0.48%	63.87%	210,807	42.13%
30	ISRG	Intuitive Surgical Inc.	0.52%	1,252	0.25%	64.39%	212,059	42.38%
31	ATVI	Activision Blizzard	0.51%	2,750	0.55%	64.90%	214,809	42.93%
32	BBBY	Bed Bath & Beyond Inc.	0.48%	1,805	0.36%	65.38%	216,614	43.29%
33	CA	CA, Inc.	0.48%	1,708	0.34%	65.86%	218,322	43.63%
34	PAYX	Paychex, Inc.	0.45%	1,007	0.20%	66.31%	219,329	43.83%
35	MU	Micron Technology, Inc.	0.44%	4,362	0.87%	66.75%	223,691	44.71%
36	CHKP	Check Point Ltd.	0.42%	1,269	0.25%	67.17%	224,960	44.96%
37	SNDK	SanDisk Corp.	0.42%	2,836	0.57%	67.59%	227,796	45.53%
38	NVDA	NVIDIA Corp.	0.41%	3,113	0.62%	68.00%	230,909	46.15%
39	ADSK	Autodesk, Inc.	0.39%	2,298	0.46%	68.39%	233,207	46.61%
40	MYL	Mylan, Inc.	0.39%	4,951	0.99%	68.78%	238,158	47.60%
41	FAST	Fastenal Co.	0.38%	727	0.15%	69.16%	238,885	47.74%
42	VRTX	Vertex Pharmaceuticals	0.38%	2,052	0.41%	69.54%	240,937	48.15%
43	CERN	Cerner Corp.	0.37%	1,265	0.25%	69.91%	242,202	48.41%
44	FISV	Fiserv, Inc.	0.36%	1,449	0.29%	70.27%	243,651	48.70%
45	MAT	Mattel, Inc.	0.34%	2,262	0.45%	70.61%	245,913	49.15%
46	ROST	Ross Stores Inc.	0.33%	1,729	0.35%	70.94%	247,642	49.49%
47	XLNX	Xilinx, Inc.	0.33%	1,623	0.32%	71.27%	249,265	49.82%
48	ORLY	O'Reilly Automotive, Inc.	0.32%	1,024	0.20%	71.59%	250,289	50.02%
49	KLAC	KLA-Tencor Corp.	0.30%	1,464	0.29%	71.89%	251,753	50.31%
50	SRCL	Stericycle, Inc.	0.30%	490	0.10%	72.19%	252,243	50.41%
51	LLTC	Linear Technology Corp.	0.29%	884	0.18%	72.48%	253,127	50.59%
52	DLTR	Dollar Tree, Inc.	0.28%	2,635	0.53%	72.76%	255,762	51.12%
53	AKAM	Akamai Technologies, Inc.	0.27%	1,985	0.40%	73.03%	257,747	51.51%
54	STX	Seagate Technology Holdings	0.26%	2,512	0.50%	73.29%	260,259	52.01%
55	HSIC	Henry Schein, Inc.	0.25%	972	0.19%	73.54%	261,231	52.21%

This table shows the individual and cumulative weights, number of news items, and news frequencies for the top 55 stocks selected to represent the Nasdaq-100 Index from 2011 and 2015 (post the May 2, 2011 rebalancing), as compared to the initial sample of 145 stocks. Note that companies not present on the index continuously for all five years were dropped, as well as any stocks with RIC inconsistencies or missing price data.

4.2. Building on the Existing Research

My analysis builds on the research conducted by Groß-Klußmann and Hautsch (2011) in a number of ways, including in scope, breadth, and design. First off, by focusing on the Nasdaq Index, I expand on the geographic scope of their study from the UK (FTSE 100 Index) to the US. In fact, in all the related literature on news analytics, no reference was found to news-based research conducted specifically on Nasdaq stocks, with many studies instead revolving around Dow Jones-related news. Second, the analysis has a wider breadth, covering 55 companies over five years, as compared to 39 companies over one and a half years. I also use more granular intraday price data, aggregated to one-second and 15-second intervals as opposed to 20-seconds. I also incorporate the underlying TAQ data, thus allowing me to look at much shorter windows with greater precision.

The immediate reaction to breaking news strikes me as the natural starting point to study the effectiveness of news analytics suites, since their very conception was geared towards triaging mass volumes of incoming news and synthesizing them into a machine-readable format. Knee-jerk reactions spanning the first few seconds to minutes after a piece of news is released have yet to be studied in much detail, and offer a chance to add to the existing body of research on the applications of news analytics databases such as TRNA and its Dow Jones counterpart, Raven Pack News Analytics (RPNA). In fact, von Beschwitz et al. (2015) and Foucault, Hombert, and Rosu (2013) observe faster price discovery following the introduction of various news analytics suites, which makes intuitive sense in this day and age of speed competition. It follows that those investors with a speed advantage would be inclined to incorporate these news signals into their algorithmic strategies and attempt to trade on them more aggressively than their competitors, consistent with Foucault, Hombert, and Rosu's (2016) predictions about short-run price movements.

Another important difference to highlight is the sample time horizon. While Groß-Klußmann and Hautsch's sample coincides with the height of the global financial crisis (GFC), a period marked by an increased proportion of negative news, heightened market volatility, and widespread market panic, mine takes place in the aftermath, a period of tenuous recovery marked by a relative sense of optimism that the worst had passed. Garcia (2013) found that investor sentiment had a more pronounced effect on the DJIA during recessions, in line with the results of Smales (2014a,b) and Allen, McAleer and Singh (2013, 2015). That being said, the relative calm in the period following the GFC may have rendered the impact of news less potent than during the height of it, but would certainly prove quite robust should my findings prove to be consistent with related studies conducted during the crisis.

4.3. Methods for Calculating Cumulative Abnormal Returns and Volume

4.3.1. Event Windows

I define three main groups of event windows for the cumulative abnormal returns and volume analysis: Information Leakage, Knee-jerk Reaction, and Secondary Reaction. These are detailed below and in Figure 11.

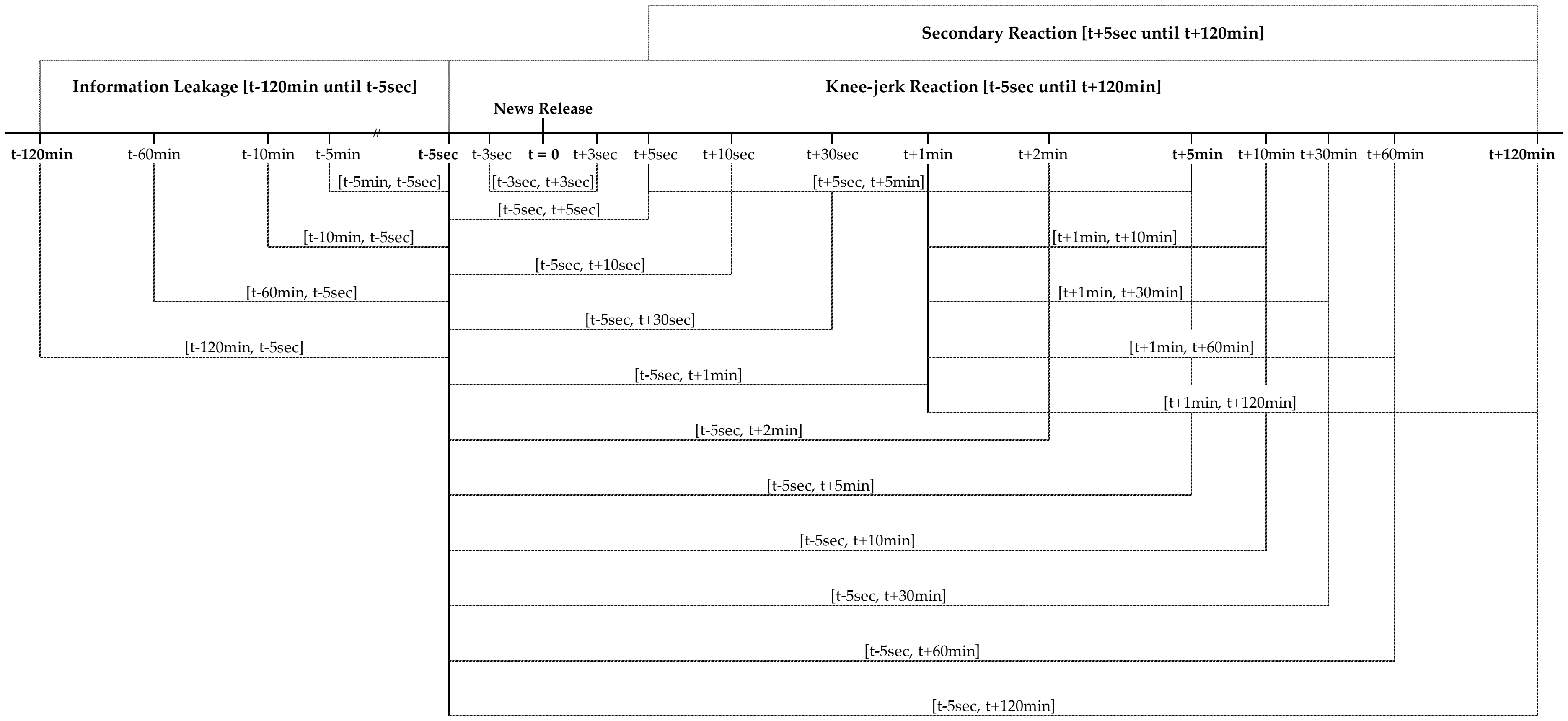
The Information Leakage block of windows refers to those periods beginning anywhere from five minutes to two hours before the news item is released, and ending five seconds before the event timestamp. Related research has observed concrete signs of price movement in the hours preceding a news release (Groß-Klußmann and Hautsch, 2011), suggesting information is being leaked via alternate news sources. Since the novelty indicator only compares the newness of Thomson Reuters news to its own database of stories, we cannot know the true timeliness of Thomson Reuters in relation to other newswires and social media outlets. The Information Leakage period accounts for this possibility by looking at four windows preceding the actual news release.

The Knee-jerk Reaction windows straddle the news release from three to five seconds before the release timestamp up to two hours afterwards. As its name suggests, the knee-jerk portion includes the shortest windows spanning from three seconds before until up to 10 seconds after the release. These are meant to capture the impact of high-frequency traders as well as latency sensitive manual traders who are likely employing strategies augmented by TRNA metadata, and executing trades based on carefully selected trigger points. Additional windows spanning beyond the first few seconds are aimed at capturing any price drift occurring in the first few minutes up to two hours following the news release. These windows account for more detailed information becoming incorporated into the stock price as reporters release stories and updates that add context to the initial burst of headlines. They also capture the impact of slower traders who may be trying to piggyback off the signals of fast traders.

Finally, the Secondary Reaction windows differ critically from the Knee-jerk Reaction in that they begin five seconds to one minute after the news is released, and end anywhere from five minutes to two hours thereafter. These windows capture the price impact observed purely after fast traders have had a chance to move the market, thus further helping isolate the impact algorithmic versus slow traders, as the expectation is to see only marginal returns captured in these post-release windows.

Figure 11

Event Windows for Cumulative Abnormal Returns and Volume Analyses



This figure illustrates the 20 event windows used to calculate Cumulative Abnormal Returns and Volumes. They are grouped into three categories: Information Leakage (four windows spanning from t-120min until t-5sec), Knee-jerk Reaction (11 windows straddling a news release from t-5sec until t+120min), and Secondary Reaction (five windows starting at t+5sec until t+120min). These three groups are designed to capture and isolate news leaked from alternate sources, trades executed by high-frequency or low latency traders, drift effects from news updates, as well as bandwagon effects stemming from late trades.

4.3.2.HPRs, CARs, and CAVs

The price, volume, and bid-ask data for individual stocks and indices are harvested from the Thomson Reuters Tick History (TRTH) database at one-second intervals. Corresponding Trade and Quote (TAQ) data are also procured from TRTH for validation purposes. Only entries for which the 0 < bid price < ask price are accepted. In order to cope with the irregularly spaced nature of the TRTH data, the "previous tick" method is used for sampling at high frequencies, per Hansen and Lunde (2006) and Wasserfallen & Zimmermann (1985).[7] It involves using the last available price with a positive closing bid and a positive closing ask to fill in the missing timeline of price data. In order to reduce the number of missing observations, but at the same time limit potential overlap between windows, stock data is accepted up to three seconds before the beginning and end of a window, while benchmark data is accepted up to five seconds before a release.

I use the method employed by Field and Hanka (2001) to calculate cumulative abnormal returns (CARs) and volume (CAVs). The stock's long, short, and midpoint holding period returns (HPRs), $r_{i,t}$, are calculated over each CAR window, and subsequently divided by the benchmark index's corresponding long, short, or midpoint (HPR), $r_{m,t}$.

Long HPRs (1) simulate buying at ask and later selling at the bid, midpoint HPRs (2) involve buying and selling at the midpoint of bid-ask, and short HPRs (3) mimic selling at the ask and later buying back at bid. HPRs are calculated according as:

$$Long\ HPR_i = long(r_{i,t}) + 1 = \frac{(b_1 - a_0)}{a_0} + 1 \tag{1}$$

$$Midpoint\ HPR_i = midpoint(r_{i,t}) + 1 = \frac{\left[\left(\frac{a_1 + b_1}{2}\right) - \left(\frac{a_0 + b_0}{2}\right)\right]}{\left(\frac{a_0 + b_0}{2}\right)} + 1 \tag{2}$$

$$Short\ HPR_i = short(r_{i,t}) + 1 = \frac{(a_1 - b_0)}{b_0} + 1 \tag{3}$$

where b is the best closing bid price and a the best closing ask in a given one-second interval.

---

[7] Sampling high frequency data that lacks time persistence is possible using previous tick extrapolation, which involves taking the last observed price prior to the sampling point. This method is preferrable to linear interpolation for simulating expected trading returns since it does not use future data to calculate a previous timestamp.

Cumulative abnormal returns are calculated using (4). Contrary to Groß-Klußmann and Hautsch (2011), I use the S&P500 as the benchmark index since I expect a strong positive correlation between individual Nasdaq-listed stocks and the broader Nasdaq index, by virtue of larger constituents acting as bellwethers for the entire tech-sector.

$$CAR_i = \prod_{t=t_0}^{t_1} \left(\frac{1 + r_{i,t}}{1 + r_{m,t}} \right) - 1 \tag{4}$$

Cumulative abnormal volumes are computed according to (5) as the percent difference between the stock's one-second average volume over the relevant window divided by its one-second average volume over the past 45 days during market open, per Tumarkin and Whitelaw (2001), and Hakim, Lypny, and Bhabra (2012). The first and last 15 minutes of the trading day are excluded from the benchmark volume calculation, similar to the newswire data.

$$CAV_i = \frac{V_{i,T}}{\frac{1}{n_i} \sum_{t-46}^{t-1} V_{i,T}} - 1 \tag{5}$$

where $V_{i,T}$ is measured as volume per second.

Note that 14 stocks underwent splits during the sample horizon, two of which split twice and one of which reverse split, the details of which can be found in Appendix E. I adjust the volume data forwards in time since most of these splits occurred in tail end of the sample, late-2013 onwards:

$$Adjusted\ V_i = \frac{Post\ Split\ V_i}{Adjustment\ Factor} \tag{6}$$

I examine the sentiment indicators on both a raw and net level, computing net tone using Dzielinski's (2012) method:

$$Net\ Sentiment = P_{+ive} - P_{-ive} \in\ [-1; 1] \tag{7}$$

where

$$P_{+ive} + P_{-ive} + P_{neut} = 1$$

and P_{+ive}, P_{-ive}, P_{neut} are the TRNA sentiment scores indicating the probability that the news item is positive, negative, and neutral in tone, respectively.

4.4. Quantile Regression Model

In order to assess whether news metrics can explain abnormal knee-jerk returns, I regress a stock's 15-second interval abnormal long and short returns against its corresponding news and traffic variables for all five years of the sample. By using abnormal returns, I implicitly account for the impact of the broader market index, allowing me to better isolate stock-specific volatility while incurring fewer multicollinearity issues in modelling.

Given the dataset's susceptibility to long periods of no news and little-to-no price action, it is useful to understand the relationships between these variables at either ends of the return distribution, as they may not be linear. To do this, I employ a quantile regression model similar to Koenker and Bassett (1978) and Koenker and Hallock (2001), an approach that provides a number of advantages in terms of versatility and robustness. For one, median-based regressions are more resistant to outliers and make no assumptions about the distribution of the residuals of the data. Since the beta coefficients of the explanatory variables behave as a function of the quantile being analyzed, the model allows us to assume that distinct quantiles of the distribution are affected differently by the explanatory variables. In other words, we can entertain the possibility that news variables impact higher return quantiles differently from lower return quantiles without having to segment the dataset and incur sample selection bias and smaller sample sizes.

The following model (8) is specified and run for. A total of five quantile regression lines are estimated for the following quantiles {0.1, 0.3, 0.5, 0.7, 0.9}.

$$\begin{aligned} AR_i = \quad & \beta_0 + \beta_1 Alert_i + \beta_2 Relevance_i + \beta_3 SetimentLogOdds_i + \beta_4 WordCount_i + \\ & \beta_5 TrafficEventCount_i + \beta_6 TrafficReferenceEventCount_i + \\ & \beta_7 TrafficSecondsSincePreviousEvent_i + \\ & \beta_8 TrafficBeforeEventCount_i + \beta_9 TrafficBeforeReferenceEventCount_i + \\ & \beta_{10} TrafficBeforeMedianSecondsBetween_i + \\ & \beta_{11} AbnormalAdjustedVolume(t-1)_i + \\ & \beta_{12} Monday_i + \beta_{13} Tuesday_i + \beta_{14} Thursday_i + \beta_{15} Friday_i + \\ & \beta_{16} OpeningHalfHour_i + \beta_{17} ClosingHalfHour_i + \beta_{18} Morning_i + \beta_{19} Midday_i \end{aligned} \tag{8}$$

Alert, Relevance, and *WordCount* refer to the TRNA scores of the stock being analyzed, with the latter used as a proxy for Article length and information completeness. Note Novelty was dropped from the model due to its high correlation with other news variables, and limited power in actually proxying for new news. The quantile regression uses a variant of the Net Sentiment score that weighs the natural log of the odds ratio of positive to negative news scores by the sum of their respective probabilities, thus placing less importance on highly neutral news:

$$Sentiment\ Log\ Odds = (P_{+ive} + P_{-ive}) \cdot \ln\left(\frac{P_{+ive}}{P_{-ive}}\right) \tag{9}$$

TrafficEventCount refers to the number of news items released in the same one second interval as the target news item, while *TrafficBeforeEventCount* refers to the number of news items released in the three hours preceding the target news item. The *Reference* versions of these variables is narrower in scope, and counts only those news items that are about the same company as the target news item. These variables act as proxies for the volume of news traffic in the run-up to and during the target release.

TrafficSecondsSincePreviousEvent and *TrafficBeforeMedianSecondsBetween* proxy for the flow of news traffic, where a larger value in the former implies less traffic immediately preceding the target release, and a larger median number of seconds between events implies more evenly spread out and less clustered news in the run-up to the release.

Note the date and time dummy variables use Wednesday and Afternoons as their baseline.

4.5. Limitations

There are important limitations to the dataset and analysis, the most significant of which relates directly to the Thomson Reuters News Analytics database. First, because Thomson Reuters uses proprietary algorithms to score the news, there is no way of knowing to what extent TRNA scores may be biased or inaccurate, and what underlying impact that may have on the results. As demonstrated by von Beschwatz et al. (2015), articles incorrectly released with a high relevance scores by RPNA did in fact cause knee-jerk reactions and volume spikes akin to "mini" flash crashes that retraced after 30 seconds, but had no fundamental basis. They also found evidence of high-frequency traders dynamically learning from and adapting to the accuracy of RPNA scores, resulting in less knee-jerk liquidity being available for stocks with more reliable news scores. Furthermore, Cao et al. (2020) are the first to detect a feedback effect that occurs in the corporate disclosures of companies subject to high machine readership. It is reasonable to assume that the TRNA database may suffer from similar drawbacks.

Another significant limitation of this analysis involves the concept of novelty. The novelty scores being used only relate to the universe of Thomson Reuters news as it compares to itself, meaning TRNA novelty scores can only tell us how many similar stories were released by Thomson Reuters in the previous 12-hours to seven days prior. This generously presumes that Thomson Reuters is always the first newswire to break a story, which borders on impossible given intense competition from the likes of Bloomberg, Dow Jones, and a host of other news outlets. This assumption is subject to even greater illusions of grandeur when we consider the growing influence of Twitter and various social media outlets as a source of breaking news, which have a speed advantage given their lack of editorial red tape. In order to focus on truly novel news, one would need to apply a novelty algorithm to all these sources collectively, which is outside the scope of this study. Although I still attempt to use the novelty indicator, my inability to account for all news venues means it would not be all that surprising to uncover signs of information leakage, similar to Groß-Klußmann and Hautsch (2011).

Furthermore, the clustering of news items presents certain challenges in determining how markets digest an influx of simultaneously released news. As this would require too many assumptions on how different market participants choose to consume the news, I opt to exclude any TRNA observations with multiple news items released for the same stock in the same one-second window, so as to reasonably attribute a single news item to a single market reaction. Once again, I cannot preclude news released by other wires.

Finally, another important limitation relates to the calculation of sentiment scores, which are generated for each piece of news in isolation of market expectations and stock-specific sentiment. That is to say, a news item may sound highly positive or negative in a linguistic vacuum, but much less so when compared to expectations, which could adversely impact the market reaction. Similarly, news may that may seem neutral in sentiment may yet prove to be positively or negatively surprising relative to what markets were bracing for. The market's psyche and pre-news pricing are missing link of sorts in trying to understand the market's reaction, or lack thereof, to a particular piece of news. I am limited in my ability to account for relative sentiment scores; that is, how positive or negative the news is compared to the market's expectation for that company or industry at the time when the news was released. It follows that new items whose sentiment differs most from market expectations should prove most surprising, and may prompt sharper short-run price movements.

5. Results

5.1. Hypotheses

Based on the intended function of the TRNA metrics, the following hypothesis should hold:

Hypothesis 1:

> Alerts with high Sentiment scores that are tagged as Relevant and Novel should have the greatest impact on a stock's returns and volume in a high-frequency setting.

To test this theory, the mean and median long, midpoint, and short cumulative abnormal returns and volume are calculated for Nasdaq news items using the following thresholds for the Relevance, Novelty, and (Net) Sentiment scores:

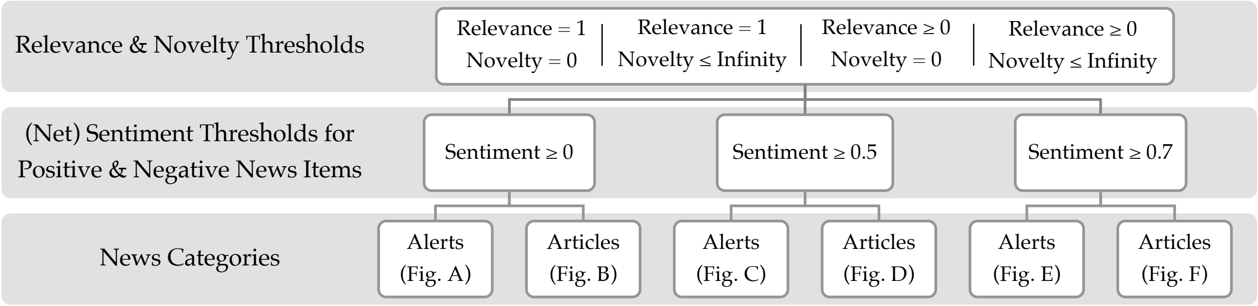
$$Relevance\ [0;1]\ \in\ \begin{array}{l} 1 = relevant\ news \\ \geq 0 = all\ relevant/irrelevant\ news \end{array}$$

$$Novelty\ [0;N]\ \in\ \begin{array}{l} 0 = novel\ news \\ \leq Infinity = all\ novel/stale\ news \end{array}$$

$$Positive/Negative\ Sentiment\ [0;1]\ \in\ \begin{array}{l} \geq 0.7 = highly\ positive/negative\ news \\ \geq 0.5 = positive/negative\ news \\ \geq 0 = all\ positive/negative\ news \end{array}$$

$$Net\ Sentiment\ [-1;1]\ \in\ \begin{array}{l} \geq 0.7\ or\ \leq -0.7 = highly\ net\ positive/negative\ news \\ \geq 0.5\ or\ \leq -0.5 = net\ positive/negative\ news \\ \geq 0\ or\ \ \leq 0 = any\ net\ positive/negative\ news \end{array}$$

All possible combinations of these variables are run, resulting in 24 separate tests for each of the CAR and CAV analyses. Results are generated for all 55 Nasdaq stocks combined, as well as for each stock individually, though in some cases, there are too few observations per stock to make the latter significant. For easier comparison, the mean and median results for positive and negative news items are reported in the same chart, with charts for Alerts and Articles shown side-by-side for increasing (Net) Sentiment thresholds. Test charts are organized into four groups of six according to their Relevance and Novelty thresholds:



The most pertinent CAR and CAV results are reported in the analysis sections that follow, and additional results, including a summary of individual stock results, are available in Appendices F, G, and H. Results from the robustness checks are available upon request.

Given the premise of the quantile regression model, the following hypotheses should also hold:

Hypothesis 2:

> Higher quantiles of abnormal returns should be increasingly positively correlated to Relevant Alerts with high absolute SentimentLogOdds scores. The coefficients of these variables should be upward sloping as the quantiles increase.

Hypothesis 3:

> Lower levels of news traffic around and preceding Relevant Alerts should lead to higher abnormal returns, as markets are less distracted from the impending news. As such, (Reference) Event Counts should be negatively correlated to abnormal returns, while SecondsSincePreviousEvent and MedianSecondsBetween should be positively correlated to abnormal returns and have more bearing on the higher quantiles.

To test these hypotheses, two series of five quantile regressions are performed using the stock's abnormal short and long returns as the independent variables, with the explanatory variables specified according model (8).

Given the significant computing constraints involved in optimizing the specification of the quantile regression model across multiple companies simultaneously, the results for Apple are used as a template for further research. The five year dataset for Apple alone involves regressing close to 1.9 million 15-second returns against more than 20 news and traffic variables, for a total of roughly 40 million datapoints. These 40 million datapoints must then be regressed a total of ten times – five quantiles for long abnormal returns, and five quantiles for short abnormal returns. Assuming the same model could be used across all 55 stocks in the sample, the results could then be compared, regrouped, and even averaged across stocks (per Groß-Klußmann and Hautsch, 2011) according to various characteristics such as industry, market capitalization, or the company's prevalence in the news.

5.2. CAR Analysis

Results of the CAR analysis show that long CARs for positive news and short CARs for negative news are most significant in the first one to two minutes following a news release, and in some cases, remain significant in all of the Knee-jerk Reaction windows spanning the first five to 10 minutes. However, even in a best case scenario (where the news is relevant, novel, and sentiment is polarized), the magnitude of CARs remains quite small at under two basis points in the first minute, as shown in Table 6 (and Figure 12e). Results for the Information Leakage and Secondary Reaction windows are less consistently significant, notably as stricter thresholds are imposed.

Table 6 (Figure 12e)

Median CARs for Nasdaq Alerts vs SPX Index when Sentiment ≥ 0.7, Relevance = 1, Novelty = 0

Window	Average N	Long CARs Positive News (bps)	Long CARs Net Positive News (bps)	Short CARs Negative News (bps)	Short CARs Net Negative News (bps)
[t-120min, t-5sec]	102	12.19*	15.70*	9.82	8.90
[t-60min, t-5sec]	145	-2.01	2.32	10.94***	11.50***
[t-10min, t-5sec]	203	4.71***	3.99***	1.93	1.29
[t-5min, t-5sec]	204	2.60***	2.60**	2.28**	2.08*
[t-3sec, t+3sec]	203	1.13***	1.13***	0.68***	0.64***
[t-5sec, t+5sec]	214	1.23***	1.37***	1.00***	0.99***
[t-5sec, t+10sec]	236	1.64***	1.64***	0.97***	0.91**
[t-5sec, t+30sec]	206	1.57***	1.52**	0.83*	0.57
[t-5sec, t+1min]	200	1.58***	1.58***	1.21*	1.21*
[t-5sec, t+2min]	203	2.57***	2.57**	0.67	0.67
[t-5sec, t+5min]	205	1.37	1.16	1.29	1.09
[t-5sec, t+10min]	205	2.37	2.37	1.42*	1.32
[t-5sec, t+30min]	193	1.16	0.46	1.15	1.42
[t-5sec, t+60min]	182	-0.58	-1.29	2.74	4.31
[t-5sec, t+120min]	154	-5.50	-15.76	2.88	2.91
[t+5sec, t+5min]	205	0.61	0.61	1.28	1.30
[t+1min, t+10min]	210	0.71	0.76	2.07**	2.12**
[t+1min, t+30min]	198	0.41	-0.34	1.98	1.54
[t+1min, t+60min]	189	-0.95	-6.01	4.30	4.30
[t+1min, t+120min]	157	-9.58	-12.84*	5.65*	6.31

Median Long (Short) Cumulative Abnormal Returns (CARs) are reported for highly Relevant, Novel Alerts with positive (negative) (Net) Sentiment scores of at least 70 per cent. Note Midpoint CARs were excluded as they were broadly insignificant at the 5 per cent level for all windows. CARs are measured in basis points (bps) of each stock's returns relative to the SPX index benchmark. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Interestingly, CARs in the Information Leakage windows more than double in the five to 10 minutes before the news breaks, consistent with Smales (2014b); however, the sharpest spike actually occurs one to two hours beforehand, which can also be seen in Table 6. Similar results were found by Groß-Klußmann and Hautsch (2011), who reported above-average trading activity more than 60 minutes prior to the news being released. This trend is most pronounced for Alerts, but is also mildly present for Articles as well (see the Information Leakage windows in Figures 12 & 13a, c, e for Alerts, and b, d, f for Articles). While this may point to competing sources releasing the news before Thomson Reuters, a closer look at the underlying news items reveals that many of the headlines with surprisingly large CAR values are actually reporting on recent stock moves as though they were news, examples of which are shown in Table 7. Notice that these news items may still be classified by the TRNA algorithm as highly Relevant and Novel, in addition to having strong Sentiment scores, even though they clearly describe past market activity.

Table 7

Sample Headlines Depicting Past Price Movements as News

Stock	Net Tone	R	N	Timestamp (ET)	Headline
YHOO	-0.76	1	0	2014-09-19 12:04:39	YAHOO SHARES HIT SESSION LOW OF $41.50 AFTER ALIBABA OPENS FOR TRADING
MYL	0.58	1	0	2014-10-02 09:58:27	MYLAN INC SHARES JUMP 4.8 PCT IN EARLY TRADING

Note R refers to Relevance (where 1 is most relevant) and N to Novelty scores (where 0 is most novel).

Notwithstanding the possible momentum effects that could arise from such headlines, it is reasonable to expect that they partially explain the sharp spike in CARs observed in the Information Leakage windows, since they essentially behave as positive feedback loops. In a rough attempt to quantify the prevalence of these types of headlines in the Information Leakage windows, I manually parse a sample of 100 positive and 100 negative Alerts with the largest CAR values in the [t0-120min, t0-5sec] window, and find that 43 per cent of negative Alerts and 25 per cent of positive Alerts fit the profile of news that describes prior price moves without offering any new information.[8] These proportions are high enough to beg the question: is the news actually being leaked, or are reporters simply writing about recent price movements? The reality may be a bit of both, but a definitive answer is outside the scope of this study. Nevertheless, further analysis (not to mention practical applications of the TRNA metadata for trading), would likely benefit from additional filtering to flag these types of news items.

---

[8] Articles were excluded from this exercise since the TRNA dataset only includes the text from the headline and not the accompanying story, making it impossible to determine whether new information was available in the underlying article. Alerts, on the other hand, are scored solely using the headline, and are not associated with any underlying text.

In the Secondary Reaction windows, negative Alerts tend to outperform negative Articles, but positive Articles outperform positive Alerts (see Figs. 13a,c,e, versus 13b,d,f). Furthermore, negative news appears to generate bigger drifts in abnormal returns than positive news, which is consistent with Hong et al. (2000) and Smales (2014c, 2015b), and not entirely surprising given the positive sentiment bias in the sample. The degree of separation is especially wide one to two hours after an Alert is released where CARs become bipolar, showing significantly positive returns for negative Alerts and significantly negative returns for positive Alerts. In cases such as Table 8 and Figure 13a, the gap between them exceeds 10bps, echoing Dzielinski (2012), and Borokova and Mahakena's (2015) conclusions about return reversals on positive news. The fact

Table 8 (Figure 13a)

Median CARs vs SPX Index for Nasdaq Alerts when Sentiment ≥ 0, Relevance ≥ 0, Novelty ≤ Infinity

Window	Average N	Long CARs Positive News (bps)	Long CARs Net Positive News (bps)	Short CARs Negative News (bps)	Short CARs Net Negative News (bps)
[t-120min, t-5sec]	1131	2.43	6.86***	0.42	15.49***
[t-60min, t-5sec]	1521	-0.31	4.79***	2.79*	11.14***
[t-10min, t-5sec]	1970	1.27**	2.76***	0.70	2.99***
[t-5min, t-5sec]	1970	0.84***	1.44***	1.16***	2.31***
[t-3sec, t+3sec]	1972	1.10***	1.09***	0.99***	0.93***
[t-5sec, t+5sec]	2016	0.98***	1.06***	0.97***	0.86***
[t-5sec, t+10sec]	2191	1.02***	1.09***	0.93***	0.94***
[t-5sec, t+30sec]	1992	0.95***	0.95***	1.01***	1.00***
[t-5sec, t+1min]	1972	0.97***	1.02***	1.06***	1.23***
[t-5sec, t+2min]	1988	1.02***	1.14***	0.99***	1.23***
[t-5sec, t+5min]	1985	0.63**	0.77**	1.36***	1.58**
[t-5sec, t+10min]	1970	0.30	0.71	1.90***	2.41***
[t-5sec, t+30min]	1906	0.26	-0.04	2.18***	1.56
[t-5sec, t+60min]	1789	-4.10***	-4.54***	5.69***	5.12***
[t-5sec, t+120min]	1499	-4.39***	-5.09***	7.14***	4.70*
[t+5sec, t+5min]	2004	0.52*	0.61	1.52***	1.53**
[t+1min, t+10min]	2149	0.12	0.54	1.90***	2.83***
[t+1min, t+30min]	2069	-0.65	-0.99	2.50***	1.59*
[t+1min, t+60min]	1931	-4.58***	-4.78***	6.33***	4.83***
[t+1min, t+120min]	1611	-4.78***	-4.96***	6.59***	5.54***

Median Long (Short) Cumulative Abnormal Returns (CARs) are reported for all Alerts without imposing any thresholds for Relevance, Novelty, and (Net) Sentiment. Note Midpoint CARs were excluded as they were broadly insignificant at the 5 per cent level for all windows. CARs are measured in basis points (bps) of each stock's returns relative to the SPX index benchmark. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

that this dichotomy is much stronger for Alerts than Articles (see Table 9 and Figure 13b for a comparison) could signal that breaking news headlines are more susceptible to positive exaggeration than the underlying stories, resulting in short-lived rallies followed by sharper reversals on positive Alerts, while negative Alerts are generally more credible and lead to stronger drifts in the stock prices.

In order to test the robustness of these results, CARs were also generated using the QQQ ETF instead of individual stock returns, yielding quantitatively similar results.

Table 9 (Figure 13b)

Median CARs vs SPX Index for Nasdaq Articles when Sentiment ≥ 0, Relevance ≥ 0, Novelty ≤ Infinity

Window	Average N	Long CARs Positive News (bps)	Long CARs Net Positive News (bps)	Short CARs Negative News (bps)	Short CARs Net Negative News (bps)
[t-120min, t-5sec]	7584	-0.07	0.04	2.36***	2.64***
[t-60min, t-5sec]	9534	0.86**	1.50***	1.64***	2.24***
[t-10min, t-5sec]	11695	0.87***	0.78***	1.39***	1.20***
[t-5min, t-5sec]	11753	0.95***	0.90***	1.14***	1.23***
[t-3sec, t+3sec]	11771	0.99***	1.01***	1.03***	1.04***
[t-5sec, t+5sec]	11825	0.96***	0.97***	1.05***	1.05***
[t-5sec, t+10sec]	12616	0.97***	0.97***	1.02***	1.01***
[t-5sec, t+30sec]	11802	1.00***	0.99***	1.03***	1.04***
[t-5sec, t+1min]	11772	0.97***	0.99***	1.05***	0.99***
[t-5sec, t+2min]	11747	0.95***	0.94***	1.15 ***	1.13***
[t-5sec, t+5min]	11758	0.97***	1.01***	1.17***	1.17***
[t-5sec, t+10min]	11705	0.81***	0.78***	1.41***	1.48***
[t-5sec, t+30min]	10775	0.27	0.31	2.12***	2.12***
[t-5sec, t+60min]	9835	-0.29	-0.30	2.71***	2.67***
[t-5sec, t+120min]	7908	-0.45	-0.57	2.79***	2.63***
[t+5sec, t+5min]	11795	0.98***	0.99***	1.09***	1.17***
[t+1min, t+10min]	12542	0.87***	0.86***	1.11***	1.15***
[t+1min, t+30min]	11543	0.42*	0.48*	1.65***	1.67***
[t+1min, t+60min]	10555	-0.51	-0.45	2.29***	2.32***
[t+1min, t+120min]	8474	-0.28	-0.55	2.68***	2.51***

Median Long (Short) Cumulative Abnormal Returns (CARs) are reported for all Articles without imposing any thresholds for Relevance, Novelty, and (Net) Sentiment. Note Midpoint CARs were excluded as they were broadly insignificant at the 5 per cent level for all windows. CARs are measured in basis points (bps) of each stock's returns relative to the SPX index benchmark. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12

Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Returns (CARs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively. Note that Long CARs are reported for (Net) Positive News and Short CARs for (Net) Negative News so as to respect the implied direction of the market reaction.

For this series of tests, Relevance is set to 1 (most relevant news), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CARs are measured in basis points (bps) relative to the SPX index benchmark. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

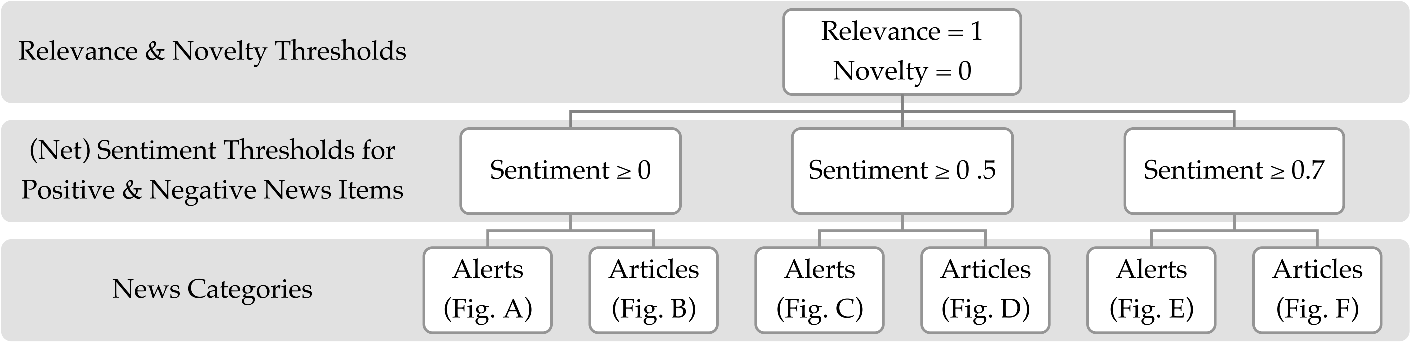
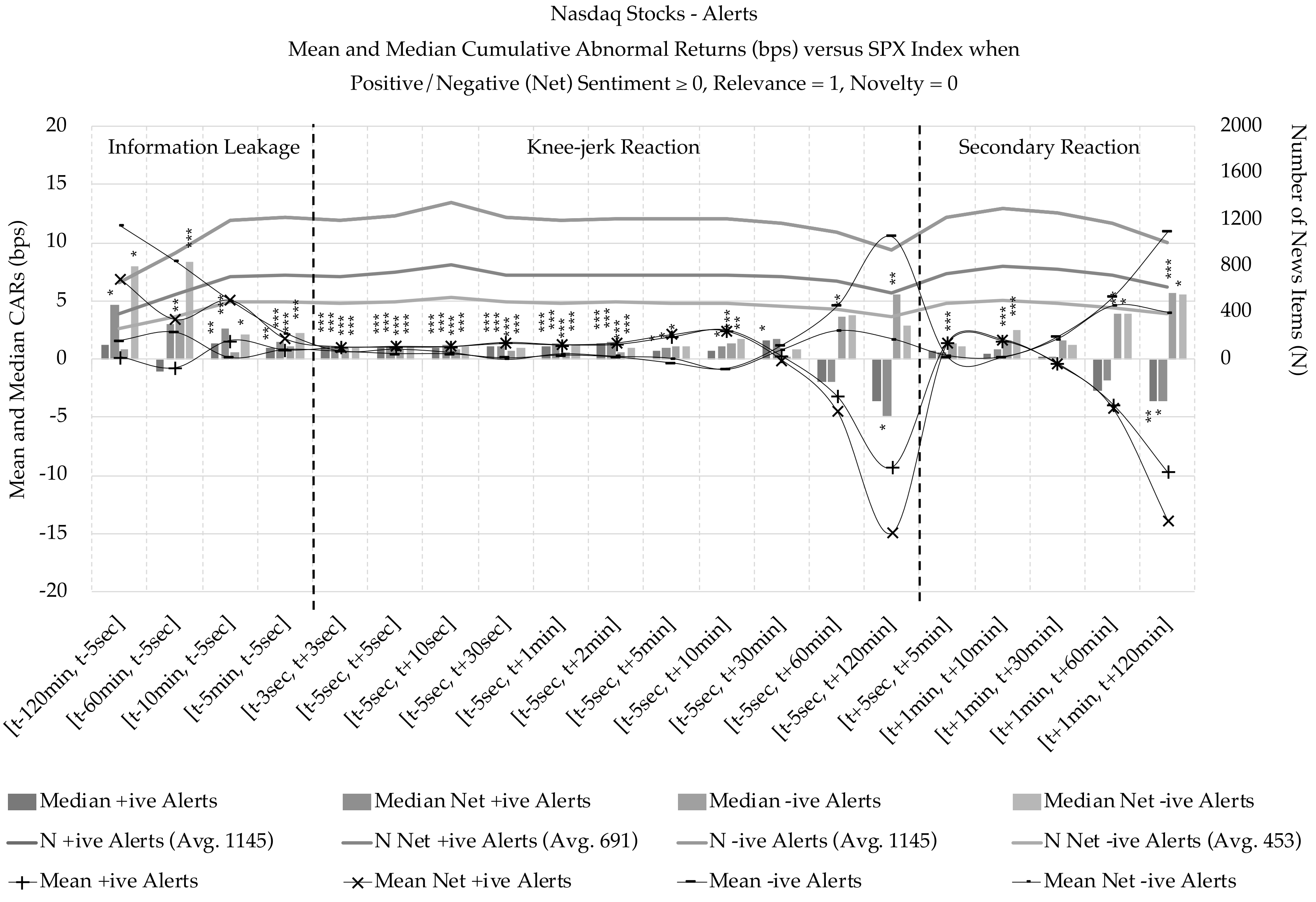


Figure 12a

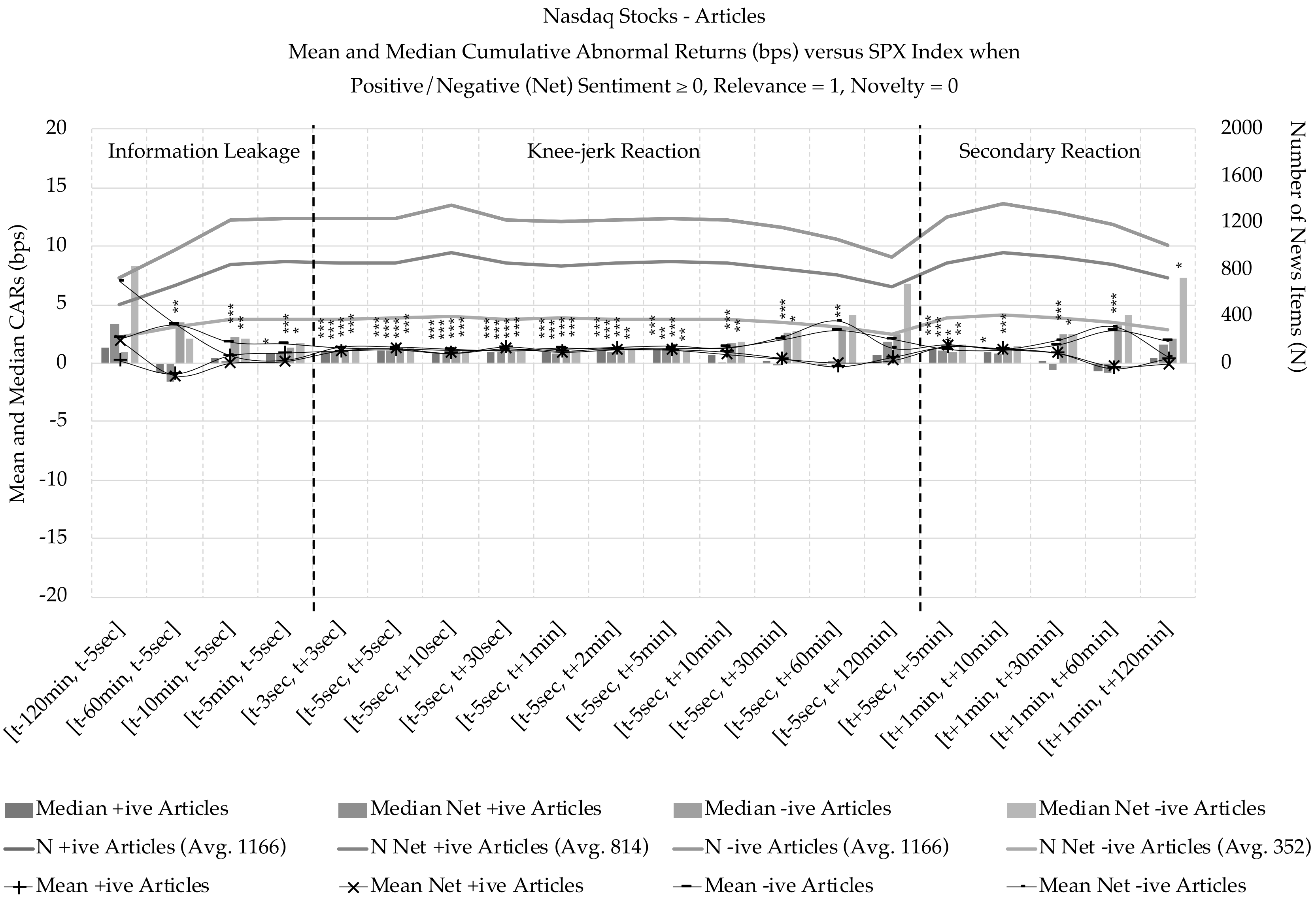
Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12b

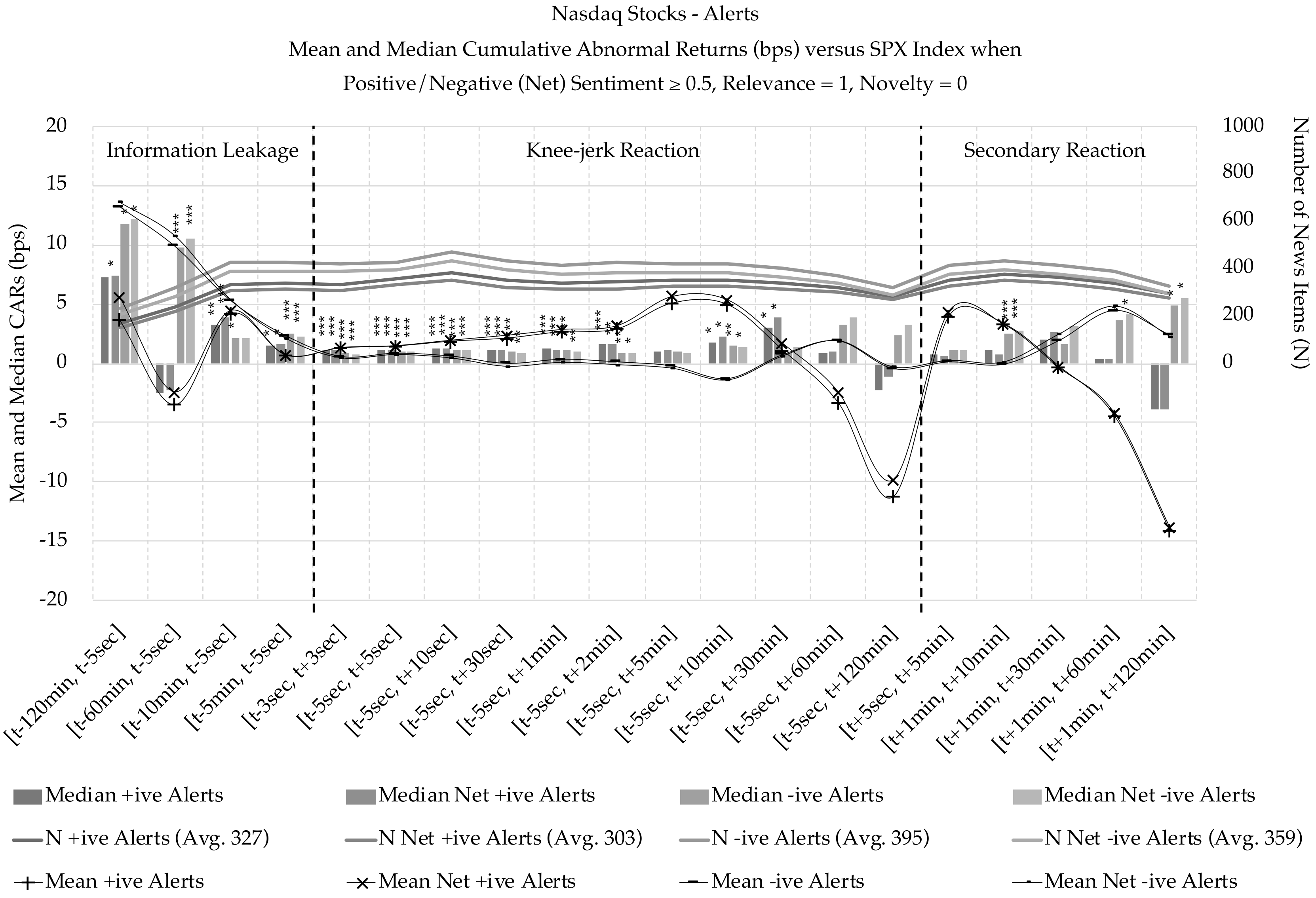
Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12c

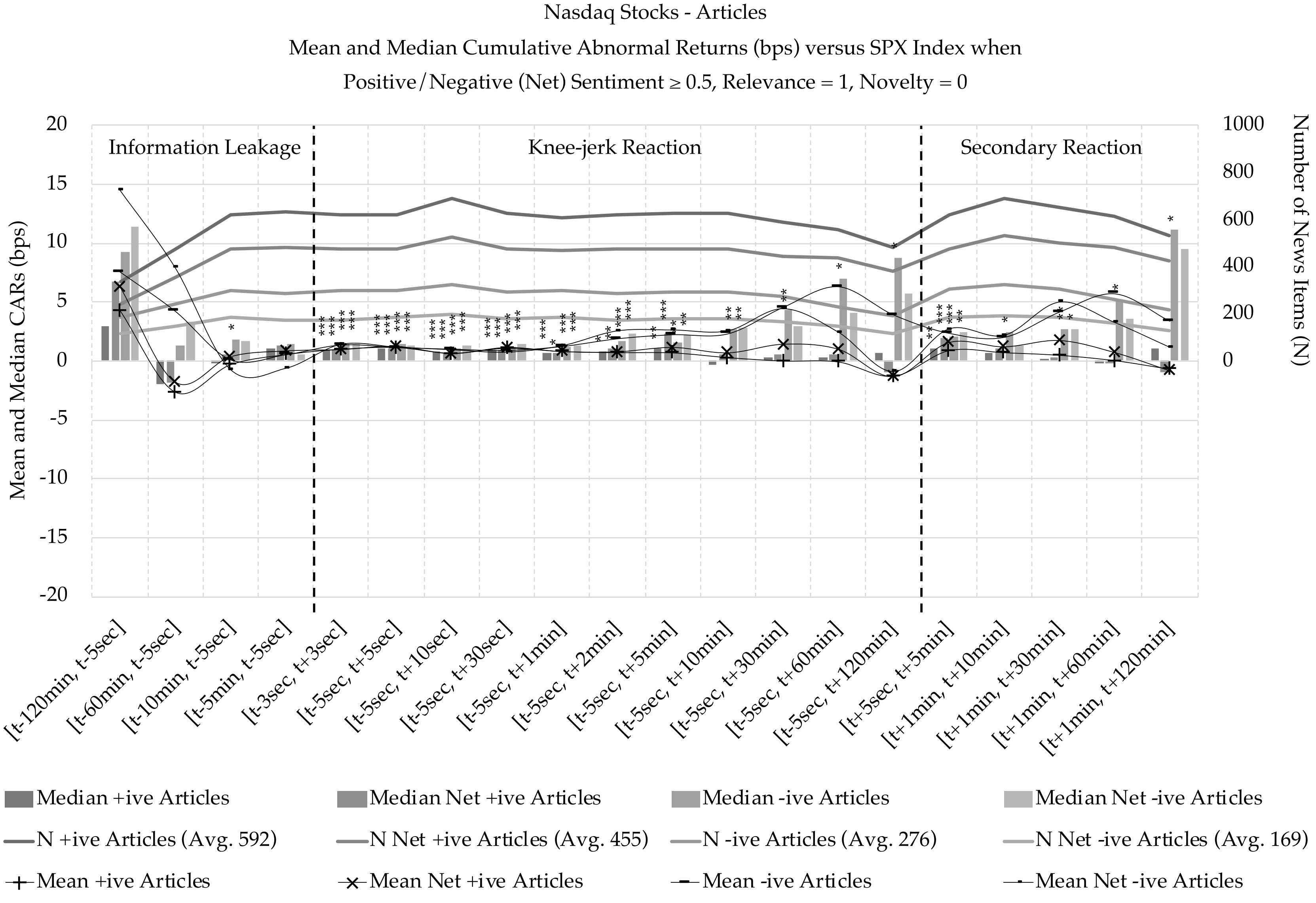
Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12d

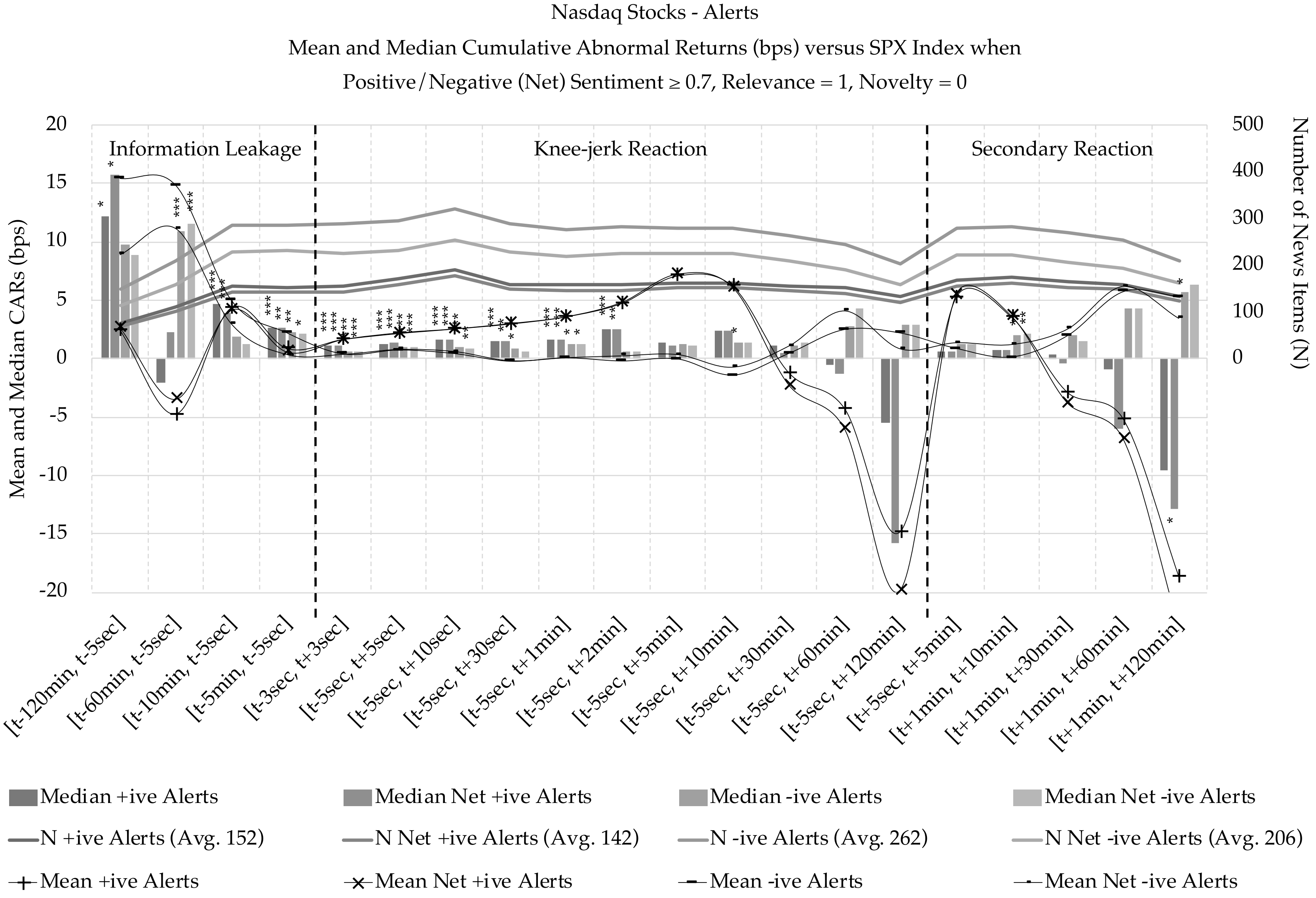
Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12e

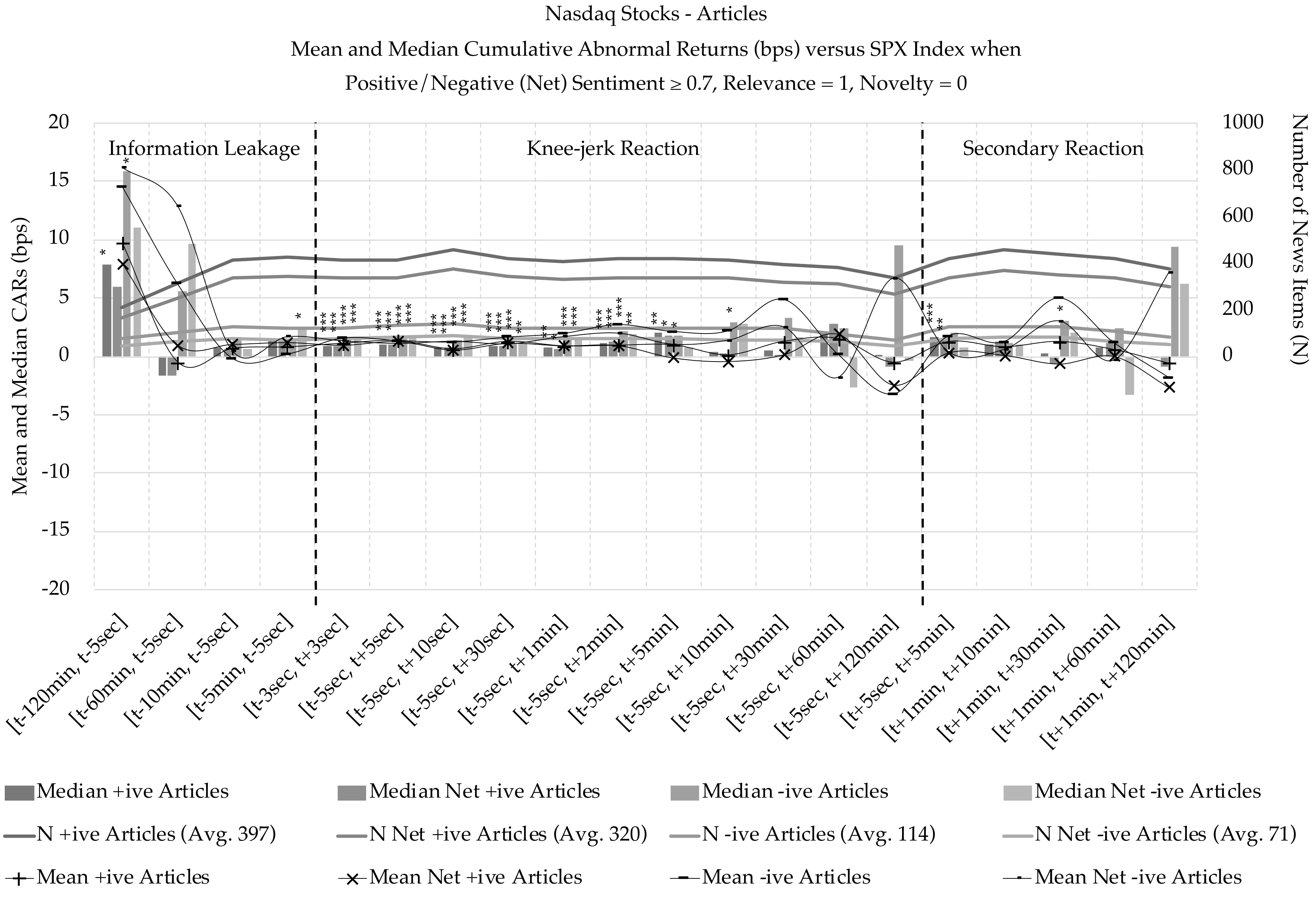
Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 12f

Mean and Median CARs vs SPX Index for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13

Mean and Median CARs vs SPX Index for Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Returns (CARs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively. Note that Long CARs are reported for (Net) Positive News and Short CARs for (Net) Negative News so as to respect the implied direction of the market reaction.

For this series of tests, no thresholds are set for Relevance and Novelty (all relevance and novelty scores are included), while absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CARs are measured in basis points (bps) relative to the SPX index benchmark. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

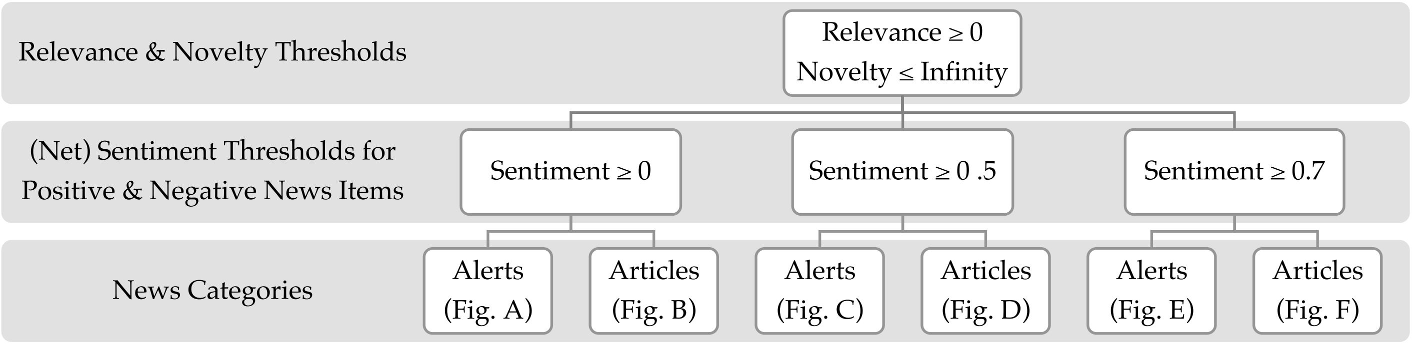
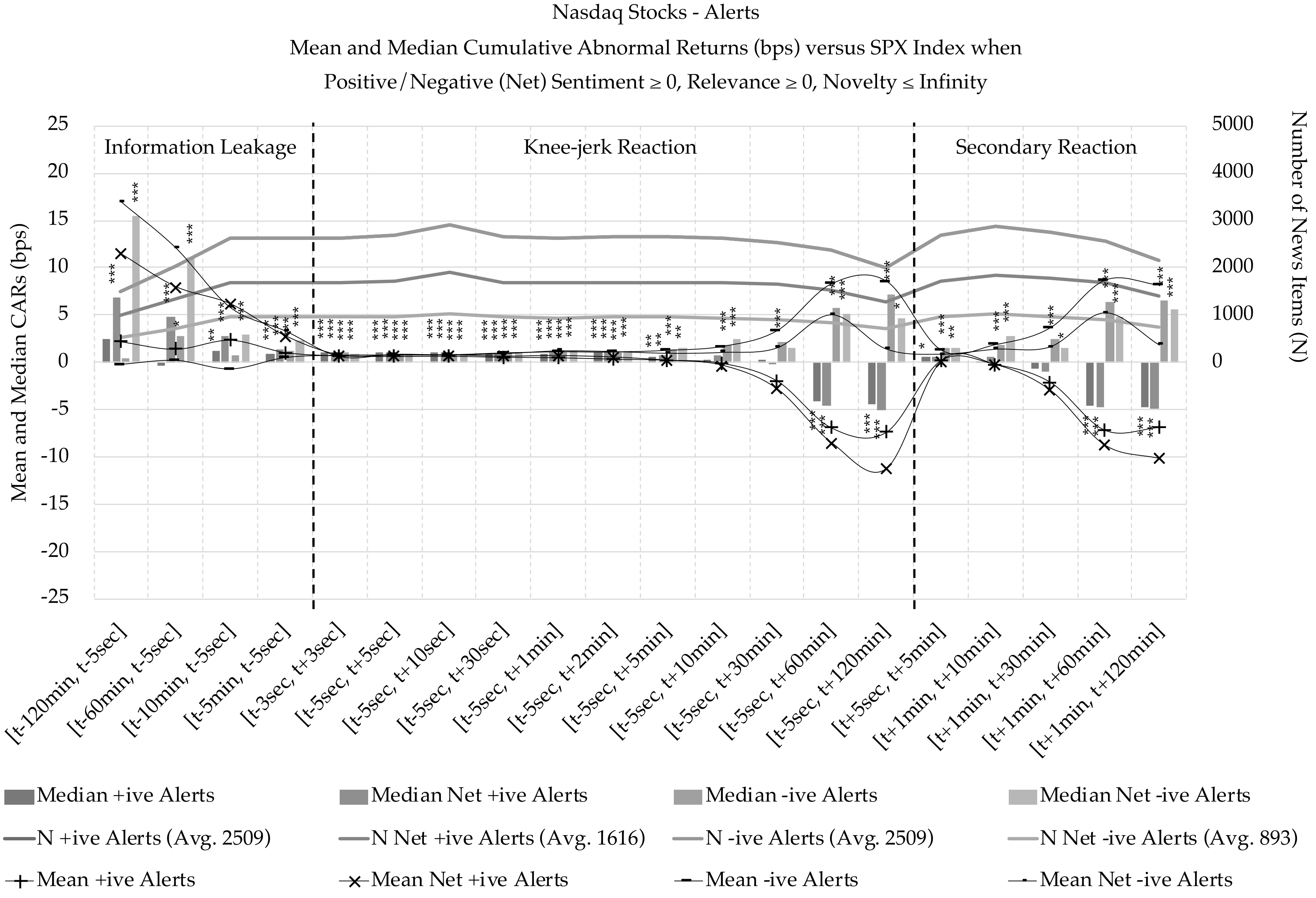


Figure 13a

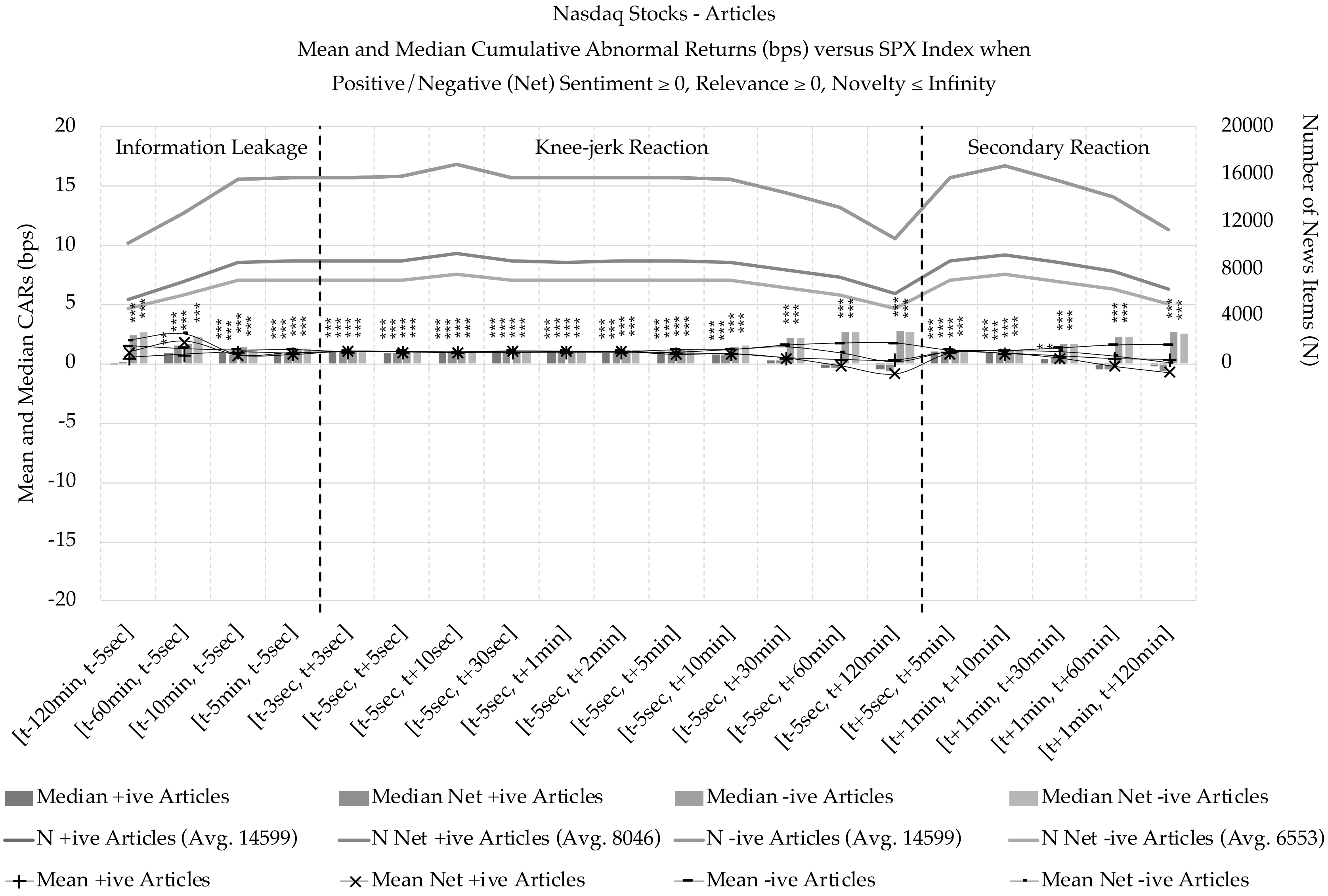
Mean and Median CARs vs SPX Index for Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13b

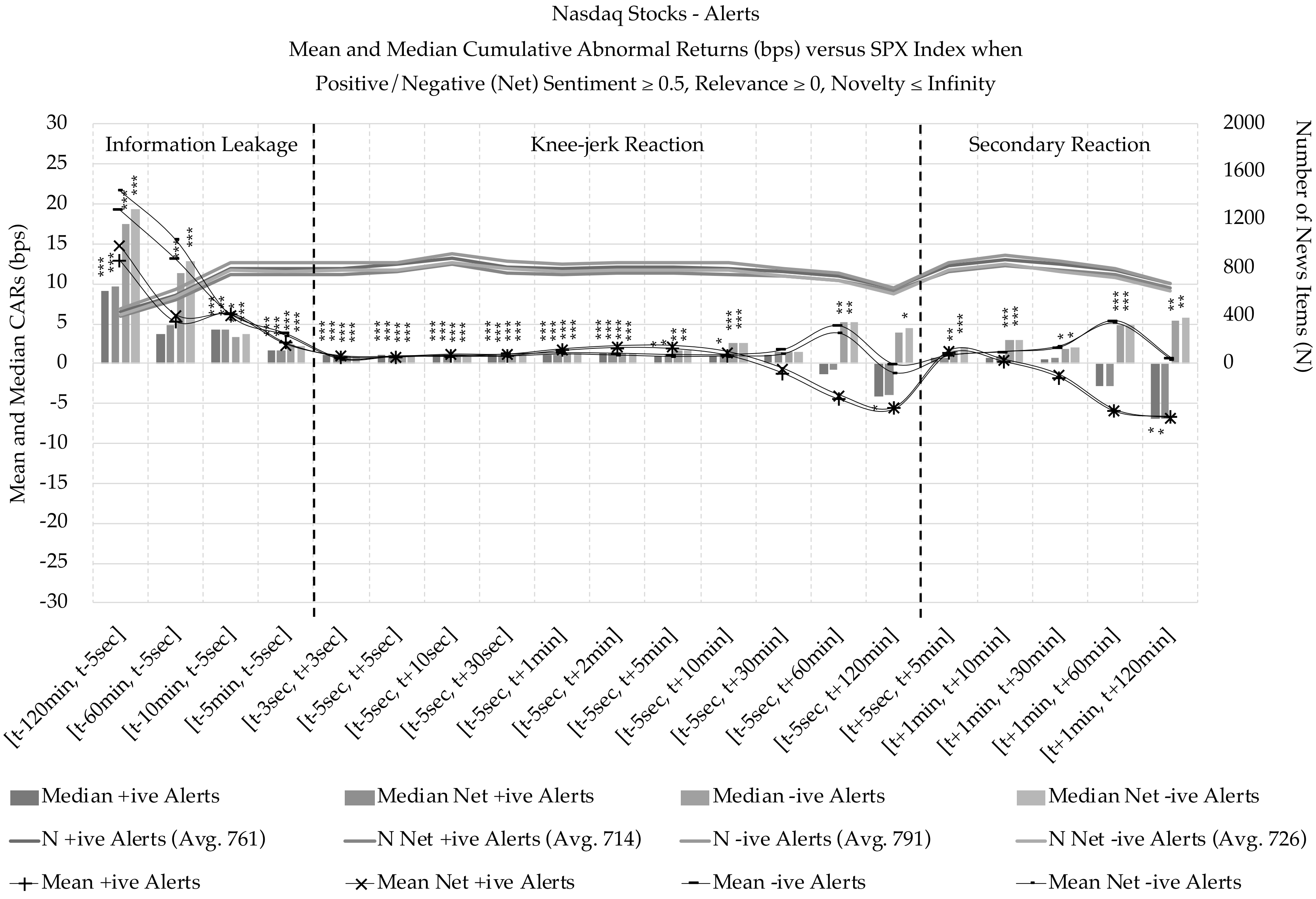
Mean and Median CARs vs SPX Index for Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13c

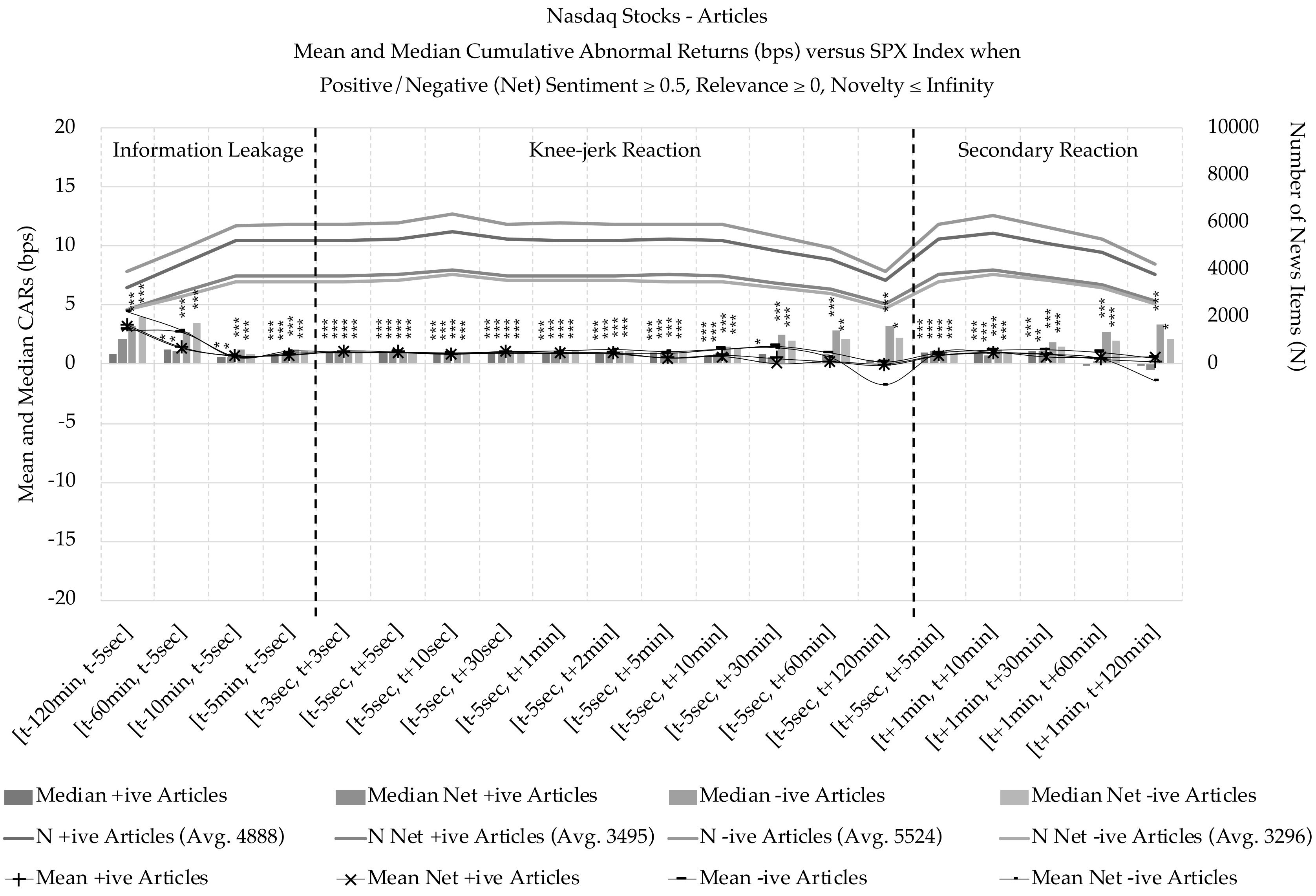
Mean and Median CARs vs SPX Index for Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13d

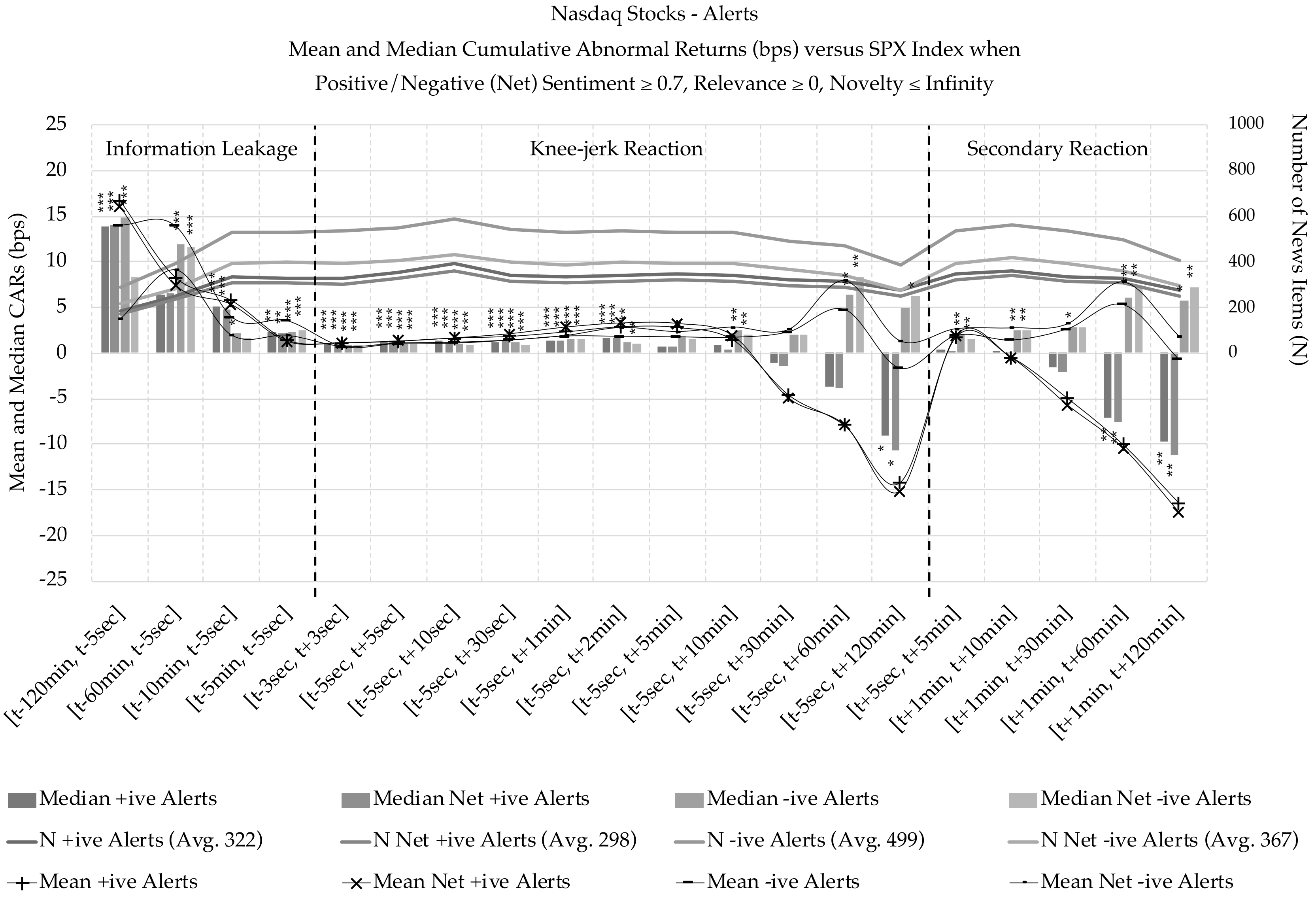
Mean and Median CARs vs SPX Index for Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13e

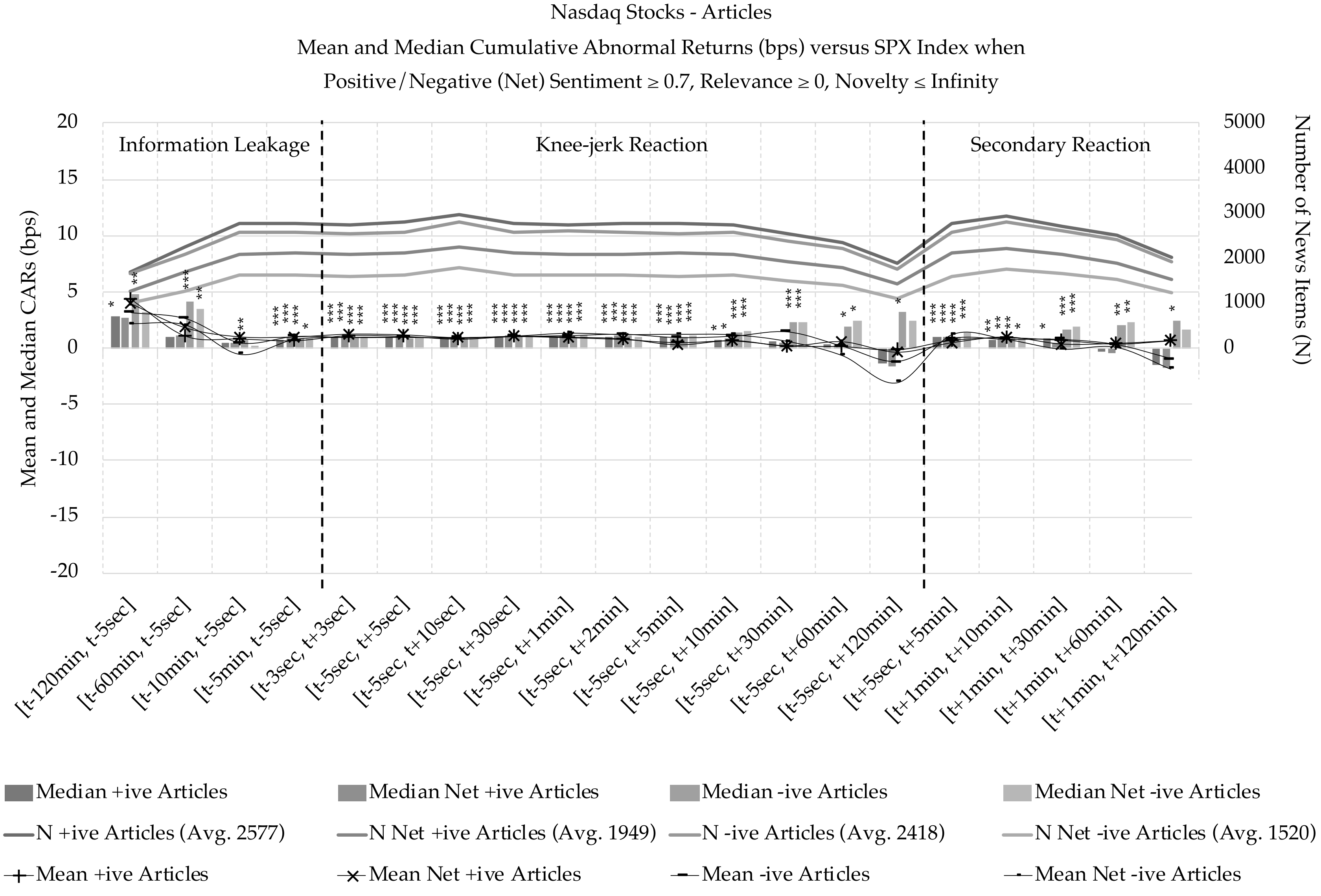
Mean and Median CARs vs SPX Index for Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 13f

Mean and Median CARs vs SPX Index for Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

In terms of which combination of variables produced the highest median CARs and ultimately performed the best across each event window, the results of Table 11 show that Sentiment was an important factor, since over 50 per cent of the maximum median CARs were derived using a Sentiment threshold of 0.7 or higher. Relevance was also key in at least 70 per cent of windows, as was Novelty for Articles and positive Alerts. While the prevalence of high Sentiment and Relevance thresholds remained fairly constant across both positive and negative Alerts and Articles, it is interesting to note that Novelty did not seem to matter much for negative Alerts, and was only a contributing factor in 10 per cent of short CARs. Once again, this may be explained by the positive sentiment bias in the sample, resulting in a higher bar for good news to move markets materially higher, while bad news was generally negative for prices regardless of how new it was.

Table 11

Properties of Maximum Median CARs

Indicator		Alerts			Articles		
		All	Positive	Negative	All	Positive	Negative
Sentiment	≥ 0.7	53%	55%	50%	60%	60%	60%
	≥ 0.5	33%	35%	30%	28%	40%	35%
	≥ 0	15%	10%	20%	13%	20%	5%
Relevance	Relevant	70%	80%	60%	75%	50%	100%
	Irrelevant	30%	20%	40%	25%	50%	0%
Novelty	Novel	48%	85%	10%	80%	75%	85%
	Stale	53%	15%	90%	20%	25%	15%

This table summarizes the properties of the maximum median CAR values for each event window across all test combinations. In the case of Sentiment and Relevance, the distributions are clearly skewed towards the stricter thresholds for both Alerts and Articles, while the Novelty score is more important to Articles.

The broader distributions of all CAR values in Figure 14 reflect many of these same properties. Increasing relevance is associated with bigger absolute CARs for Alerts (Figure 14a), with the distribution fanning out noticeably at the 0.5 and 0.7 thresholds, and fat tails present when relevance reaches its maximum score. When isolating the impact of increasing Relevance on a test-by-test basis though, the indicator actually seems to provide more marginal benefit to the median CARs of Articles over Alerts, and perhaps appropriately so when considering that the subject of an Alert is typically quite obvious from the headline itself, whereas Articles are longer in length and could mention many companies in varying capacities. Indeed, it turns out that 91

Figure 14

Distribution of CARs for [t-5sec, t+2min] Window

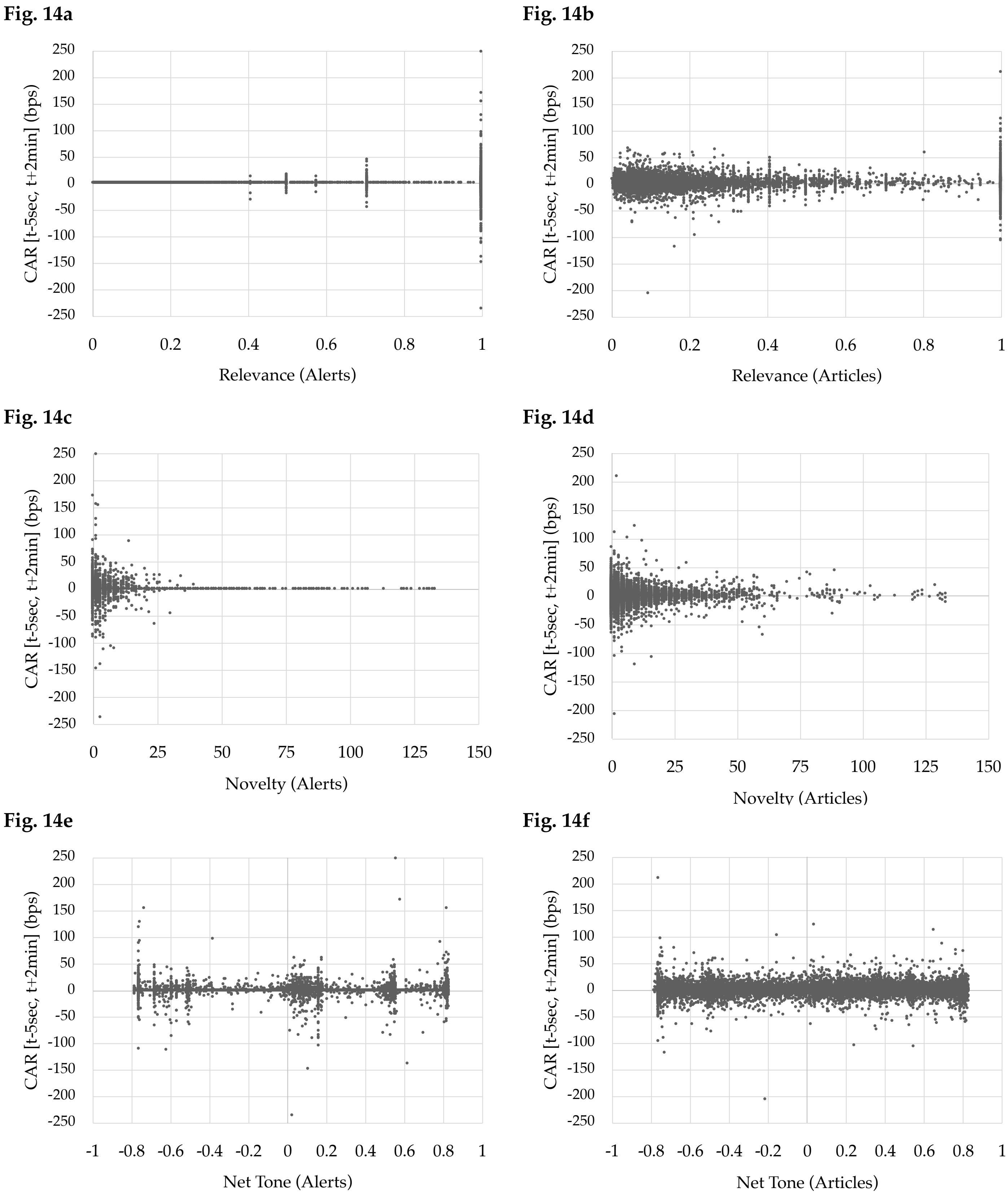


Fig. 14a

Fig. 14b

Fig. 14c

Fig. 14d

Fig. 14e

Fig. 14f

These figures show the distribution of all [t-5sec, t+2min] CARs for Alerts and Articles according to the Relevance (1 = most relevant), Novelty (0 = most novel), and Net Tone (1/-1 = most positive/negative) scores.

per cent of all Alerts in the sample are also tagged as highly relevant, compared to just 32 per cent of Articles.

The distributions of CAR values by Novelty score shown in Figures 14c and 14d similarly show the CARs fanning out as Alerts and Articles become more novel, but somewhat ironically, on a test-by-test basis, filtering the news using the Novelty indicator doesn't seem to benefit the median CARs in the Knee-jerk windows.

The distributions of CARs by Net Tone show less clearly defined trends for Articles (see Figure 14f), however the outliers for Alerts do tend cluster around absolute scores of 0.5 to 0.8, as shown in see Figure 14e. This is consistent with the results observed on a test-by-test basis, as increasing the sentiment threshold did lead to larger median CARs for Alerts, with the 0.7 threshold providing a marginal improvement of up to 0.7bps over two minutes. These gains did not translate as consistently to Articles.

Although there were few striking differences between the median CARs for Alerts and Articles, positive Alerts did appear to marginally outperform positive Articles, while negative Articles marginally outperformed negative Alerts in the first two minutes post-release, and this difference became wider once Relevance, Novelty, and Sentiment thresholds were imposed, but never exceeded 1.5bps.

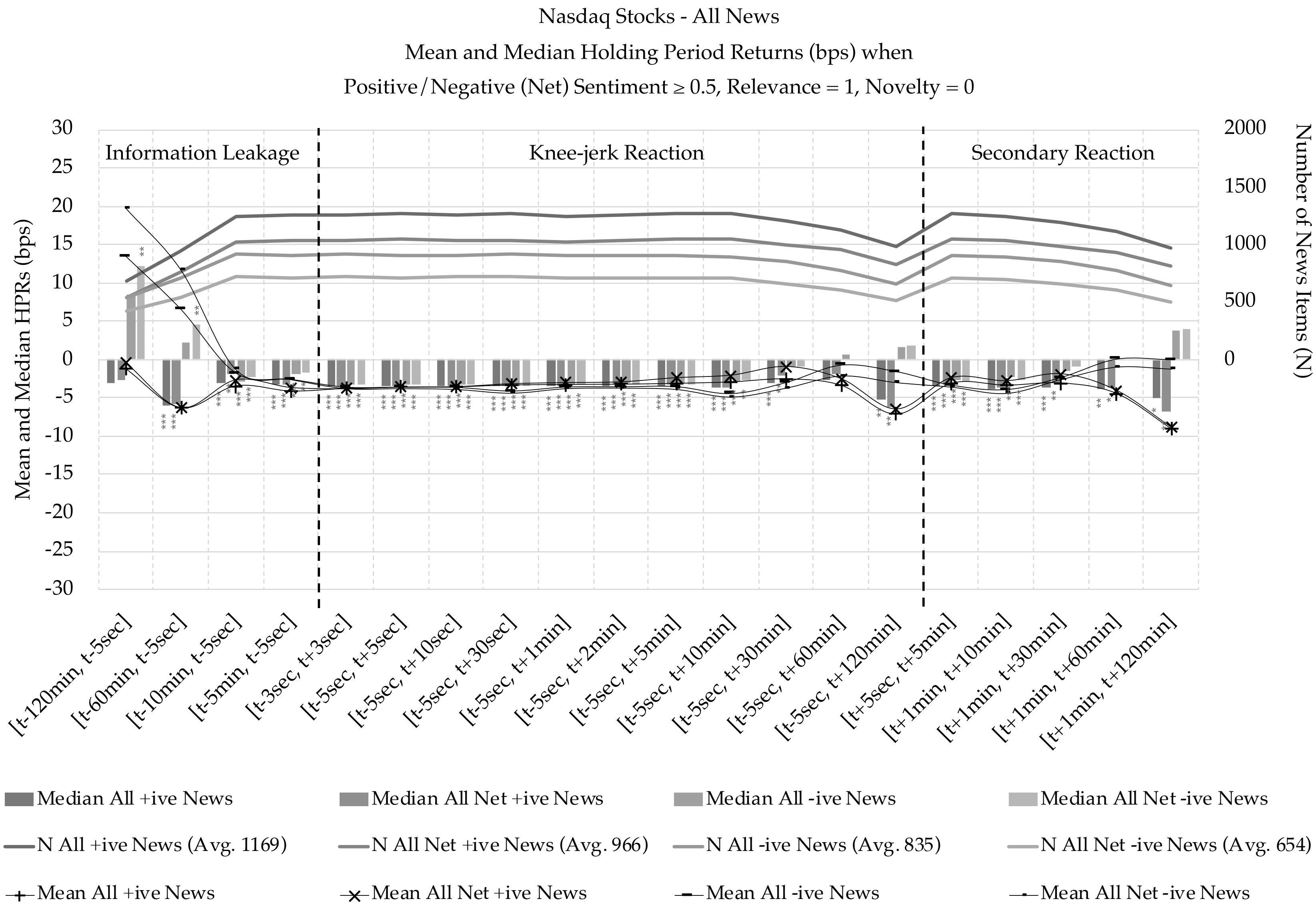
5.3. Naïve Trading Returns

For an algorithm blindly applying the Sentiment, Relevance, and Novelty metrics to a trading strategy, the HPRs underlying the CAR analysis provide a good back-test for potential short-term profits that would have been generated from buying and selling a stock based on its TRNA triggers. As shown in Figure 15, HPRs are highly significant in almost all windows for relevant, novel Nasdaq news with an absolute tone of at least 0.5, but almost never positive. Table 12 further suggests that trading a stock based on the TRNA signals alone generates a median loss of roughly two to four basis points over time for all possible test combinations, pointing to a likely inability to beat the bid-ask spread over high-frequency trading horizons. In fact, the only windows with any significant HPR gains were for (net) negative Alerts traded one to two hours ahead of the TRNA release timestamp (see Figure 15), where returns varied between four and 20bps as the sentiment threshold was increased.

Note HPRs generated by trading the QQQ ETF instead of the individual stocks yielded quantitatively similar results.

Figure 15

Mean and Median HPRs vs SPX Index for Relevant, Novel Nasdaq News when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Table 12

Summary of HPR Performance across different (Net) Sentiment, Relevance, and Novelty Thresholds for Alerts and Articles

	Test Thresholds			Positive News HPRs (bps)			Net Positive News HPRs (bps)			Negative News HPRs (bps)			Net Negative News HPRs (bps)		
				Min	Max	Median	Min	Max	Median	Min	Max	Median	Min	Max	Median
	Sentiment	Relevance	Novelty	*for all 20 Windows*		*[t-5sec, t+5min]*	*for all 20 Windows*		*[t-5sec, t+5min]*	*for all 20 Windows*		*[t-5sec, t+5min]*	*for all 20 Windows*		*[t-5sec, t+5min]*
Alerts	≥ 0	= 1	= 0	-13.25	-2.14	-3.55	-13.95	-1.72	-3.55	-3.66	7.39	-3.32	-3.66	6.51	-3.61
	≥ 0	= 1	≤ ∞	-11.24	-3.19	-4.05	-11.65	-1.82	-4.04	-3.40	5.22	-2.27	-3.32	17.17	-2.47
	≥ 0	≥ 0	= 0	-12.97	-2.56	-3.59	-13.62	-1.86	-3.58	-3.69	7.05	-3.27	-3.67	6.41	-3.67
	≥ 0	≥ 0	≤ ∞	-11.17	-3.23	-4.01	-12.01	-2.00	-3.87	-3.48	5.16	-2.39	-3.41	14.87	-2.74
	≥ 0.5	= 1	= 0	-18.77	0.00	-3.18	-18.64	0.00	-3.05	-3.63	8.29	-3.63	-3.68	10.93	-3.68
	≥ 0.5	= 1	≤ ∞	-13.85	0.00	-3.66	-13.84	0.00	-3.57	-3.34	18.59	-2.58	-3.36	23.21	-2.60
	≥ 0.5	≥ 0	= 0	-17.02	-1.09	-3.13	-17.02	0.00	-2.96	-3.60	6.68	-3.57	-3.62	8.89	-3.60
	≥ 0.5	≥ 0	≤ ∞	-13.66	0.00	-3.58	-13.51	0.00	-3.53	-3.41	15.47	-2.58	-3.42	19.17	-2.60
	≥ 0.7	= 1	= 0	-21.66	0.00	-2.84	-26.85	0.00	-2.84	-3.68	10.92	-3.51	-3.79	10.10	-3.45
	≥ 0.7	= 1	≤ ∞	-21.31	0.00	-4.05	-21.66	0.00	-4.10	-3.25	18.78	-1.66	-3.29	12.56	-1.84
	≥ 0.7	≥ 0	= 0	-21.31	0.00	-2.27	-24.29	0.00	-2.19	-3.60	8.59	-3.45	-3.74	6.94	-3.38
	≥ 0.7	≥ 0	≤ ∞	-20.49	0.00	-3.72	-21.31	0.00	-3.66	-3.31	15.40	-1.83	-3.34	10.40	-2.21
Articles	≥ 0	= 1	= 0	-5.37	0.00	-3.57	-4.68	0.00	-3.42	-4.23	0.00	-3.19	-3.36	7.79	-2.76
	≥ 0	= 1	≤ ∞	-5.57	-2.97	-3.13	-6.08	-2.25	-3.34	-3.12	0.00	-3.07	-3.35	0.83	-3.17
	≥ 0	≥ 0	= 0	-4.69	-3.08	-3.26	-4.07	-2.82	-3.19	-3.35	-0.79	-3.12	-3.68	0.00	-3.04
	≥ 0	≥ 0	≤ ∞	-4.67	-3.02	-3.21	-4.24	-2.74	-3.21	-3.19	-0.51	-3.00	-3.16	0.00	-2.97
	≥ 0.5	= 1	= 0	-6.52	0.00	-3.57	-7.11	0.00	-3.19	-3.96	9.76	-1.47	-3.49	13.17	-1.96
	≥ 0.5	= 1	≤ ∞	-4.81	-3.01	-3.29	-4.67	-1.15	-3.35	-3.54	2.08	-3.21	-4.03	4.77	-2.72
	≥ 0.5	≥ 0	= 0	-4.68	-1.85	-3.30	-3.73	-1.23	-3.25	-3.74	0.00	-3.00	-3.87	0.00	-2.88
	≥ 0.5	≥ 0	≤ ∞	-4.17	-2.92	-3.25	-3.94	-2.13	-3.34	-3.21	0.99	-3.02	-3.36	1.58	-2.99
	≥ 0.7	= 1	= 0	-4.71	0.00	-3.19	-4.09	0.00	-3.33	-3.25	15.48	-2.33	-3.29	15.48	-1.72
	≥ 0.7	= 1	≤ ∞	-6.13	-1.13	-3.42	-7.54	0.00	-3.19	-4.11	5.93	-3.43	-7.06	7.79	-2.76
	≥ 0.7	≥ 0	= 0	-3.88	-0.12	-3.14	-3.87	-0.34	-3.26	-3.96	1.38	-2.82	-4.16	3.20	-3.03
	≥ 0.7	≥ 0	≤ ∞	-5.26	-1.65	-3.32	-5.98	-1.44	-3.25	-3.47	1.75	-3.02	-3.58	3.35	-3.00

This table summarizes the minimum, maximum, and median holding period returns (HPRs) that would have transpired from trading (net) positive and (net) negative Nasdaq stock Alerts and Articles using the TRNA triggers under all possible test combinations using Relevance, Novelty, and Sentiment thresholds.

The related literature has already established a strong relationship between news and volume, with recent studies directly linking elevated trading volume to news analytics (Groß-Klußmann and Hautsch, 2011; Smales, 2014b; von Beschwitz et al., 2015). The CAV analysis is a useful complement to the CAR results, offering a different perspective on how automated news feeds affect high frequency market dynamics. Figure 16 confirms that the distributions of CAVs closely resemble the trends observed in the CAR analysis, with CAVs increasing as thresholds become stricter. The clearest relationships occur when Relevance is one (Fig. 16a) and Novelty converges to zero (Fig. 16b), while CAV outliers tend to fan out as absolute Sentiment increases (Fig.16c).

Figure 16

Distribution of CAVs for [t-3sec, t+3sec] Window

Fig. 16a **Fig. 16b**

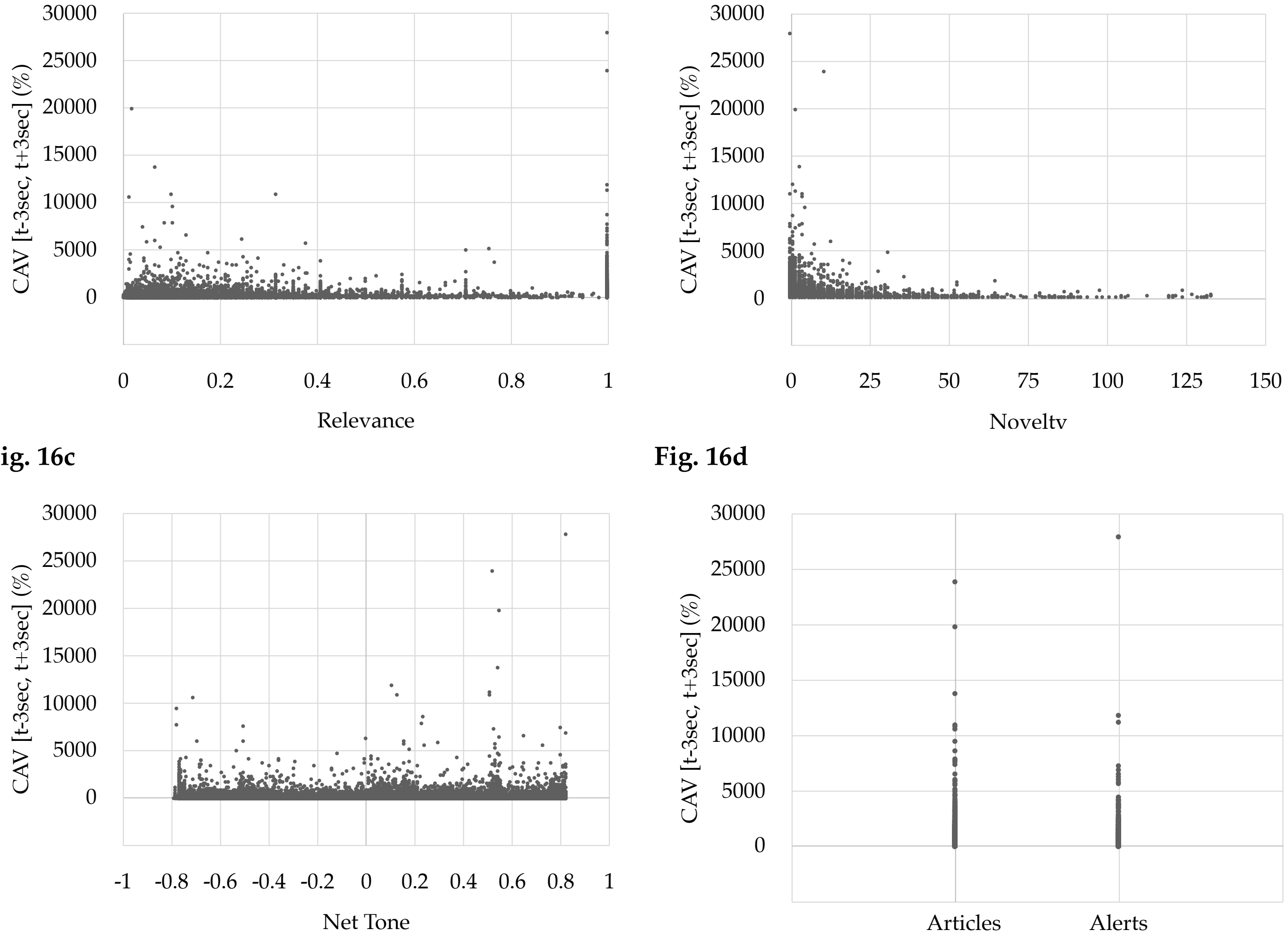


Fig. 16c **Fig. 16d**

These figures show the distribution of all [t-3sec, t+3sec] CAVs for Nasdaq Alerts and Articles combined, according to their Relevance (1 = most relevant), Novelty (0 = most novel), and Net Tone (1/-1 = most positive/negative) scores.

The median CAVs in Table 13 and Figure 18c show significant spikes in abnormal volume of approximately 20 per cent in the [t-10min, t-5sec] and [t-5min, t-5sec] windows (before a relevant Alert is released), followed by a sharp reversal immediately upon release, in the [t-3sec, t+3sec] window. This sudden inflection is possibly a side-effect of news-based trading, where high frequency players who trade aggressively on unscheduled news likely catch market makers off-guard, prompting them to pull their quotes from the books until they, and the market, have had a chance to adjust to the new information. This state of illiquidity and widened bid-ask spreads could potentially explain the temporary sharp dip in CAVs, as volume does begin returning to the market after 30 seconds, as evidenced by the pickup in CAV values which hover between 15 and 30 per cent above normal levels in the 30 minutes following the news release. This second

Table 13 (Fig.18c)

Median CAVs for Nasdaq Alerts when Sentiment ≥ 0.5, Relevance = 1, Novelty $\leq$ Infinity

Window	Average N	Positive News CAVs (%)	Net Positive News CAVs (%)	Negative News CAVs (%)	Net Negative News CAVs (%)
[t-120min, t-5sec]	605	-4.49	-4.55	-2.13	-1.08
[t-60min, t-5sec]	785	-2.06	-2.17	4.28	5.21
[t-10min, t-5sec]	1080	21.68***	21.81***	21.85***	22.30***
[t-5min, t-5sec]	1080	16.14***	15.74***	23.07***	24.57***
[t-3sec, t+3sec]	1080	-41.96***	-42.60***	-28.85***	-25.58***
[t-5sec, t+5sec]	1080	-20.77***	-20.85**	-6.55	-3.33
[t-5sec, t+10sec]	1080	-6.30	-6.27	2.16	4.55
[t-5sec, t+30sec]	1080	8.03	8.05	18.11**	20.63***
[t-5sec, t+1min]	1080	14.60**	13.76**	16.20***	17.32***
[t-5sec, t+2min]	1080	15.21***	14.55***	17.55***	18.45***
[t-5sec, t+5min]	1080	23.88***	23.75***	27.69***	29.25***
[t-5sec, t+10min]	1080	20.76***	20.92***	25.45***	27.60***
[t-5sec, t+30min]	1030	15.08***	15.06***	20.24***	23.27***
[t-5sec, t+60min]	964	8.25**	8.25**	13.22***	18.45***
[t-5sec, t+120min]	818	7.40	7.55	4.23	6.00*
[t+5sec, t+5min]	1080	23.40***	22.77***	28.33***	29.94***
[t+1min, t+10min]	1080	20.06***	19.71***	25.84***	27.88***
[t+1min, t+30min]	1030	13.79***	13.63***	18.32***	21.86***
[t+1min, t+60min]	964	7.70*	7.70*	12.06**	18.76***
[t+1min, t+120min]	818	6.87	7.76	4.03	5.49*

Median Cumulative Abnormal Volumes (CAVs) are reported for highly Relevant Alerts with positive and negative (Net) Sentiment scores of at least 50 per cent. Note Novelty is not controlled in this test. CAVs are measured in per cent above (below) the stock’s 45-day moving average volume traded during market open. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

wind of elevated trading activity could be due to a combination of late traders entering new positions and fast traders exiting their positions to secure short-term gains. This is why there is little difference between the Secondary Reaction and their equivalent Knee-jerk Reaction CAVs.

Compared to Alerts, median CAVs for Articles are practically unrecognizable in terms of volume metrics, with no signs of significantly heightened trading activity in any of the event windows. This is evident in Table 14 and Figure 18d, which show the corresponding results for Articles as Table 13 and Figure 18c did for Alerts, though the trend is consistent for all tests combinations. Immediately following the TRNA release, we observe the same abnormally low levels of median CAVs as we saw in the Knee-jerk Reaction windows for Alerts, in some cases over 50 per cent

Table 14 (Fig.18d)

Median CAVs for Nasdaq Articles when Sentiment ≥ 0.5, Relevance = 1, Novelty ≤ Infinity

Window	Average N	Positive News CAVs (%)	Net Positive News CAVs (%)	Negative News CAVs (%)	Net Negative News CAVs (%)
[t-120min, t-5sec]	1277	-11.12***	-12.23***	-6.05***	-3.571
[t-60min, t-5sec]	1662	-7.36***	-7.78***	-3.15	-3.148
[t-10min, t-5sec]	2090	0.58	0.80	4.11	6.42*
[t-5min, t-5sec]	2090	-4.86*	-4.95*	2.61	2.47
[t-3sec, t+3sec]	2090	-53.61***	-52.99***	-43.16***	-45.15***
[t-5sec, t+5sec]	2090	-37.87***	-35.82***	-31.17***	-31.15***
[t-5sec, t+10sec]	2090	-28.91***	-26.83***	-23.53***	-23.34***
[t-5sec, t+30sec]	2090	-15.72***	-14.75***	-12.29***	-10.97**
[t-5sec, t+1min]	2090	-10.57***	-9.87***	-5.87*	-5.30
[t-5sec, t+2min]	2090	-9.21***	-8.96***	-3.66	-3.00
[t-5sec, t+5min]	2090	-4.49*	-4.08	0.80	1.88
[t-5sec, t+10min]	2090	-2.62	-2.21	3.92	5.68
[t-5sec, t+30min]	1943	-2.42	-2.19	3.30	2.98
[t-5sec, t+60min]	1782	-3.64**	-3.95**	2.27	2.86
[t-5sec, t+120min]	1472	-7.34***	-7.70***	-1.92	-1.03
[t+5sec, t+5min]	2090	-5.79**	-5.45	1.02	1.62
[t+1min, t+10min]	2090	-2.08	-1.62	3.08	4.17
[t+1min, t+30min]	1943	-2.73	-2.30	3.28	2.66
[t+1min, t+60min]	1782	-4.05**	-4.12**	1.58	2.55
[t+1min, t+120min]	1472	-7.61***	-7.73***	-2.00	-1.03

Median Cumulative Abnormal Volumes (CAVs) are reported for highly Relevant Articles with positive and negative (Net) Sentiment scores of at least 50 per cent. Note Novelty is not controlled in this test. CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

below average. While this may just be the result of no one trading TRNA Articles, it does hint at the possibility that market makers could use news analytics signals to systematically dodge otherwise unpredictable volatility events, much like they would ahead of scheduled news and economic data releases.

Consistent with the CAR results, Alerts appear to have a larger and more lasting impact on market dynamics than Articles, with the Relevance and Sentiment indicators playing a key role in determining the degree of abnormal trading activity. However, a point of peculiarity is once the Novelty threshold is imposed as part of the test variables, median CAVs are no longer positive in the Information Leakage and Knee-jerk Reaction windows, as seen in Table 15 and Figure 17e.

Table 15 (Fig.17e)

Mean and Median CAVs for Nasdaq Alerts when Sentiment ≥ 0.7, Relevance = 1, Novelty = 0

		Positive News CAVs (%)		Net Positive News CAVs (%)		Negative News CAVs (%)		Net Negative News CAVs (%)	
Window	Average N	Mean	Median	Mean	Median	Mean	Median	Mean	Median
[t-120min, t-5sec]	168	25.10	-6.00	27.96	-5.32	11.43	-16.99 ***	10.43	-20.37 ***
[t-60min, t-5sec]	223	34.39	-8.87*	30.21	-8.17*	25.56	-9.72**	23.61	-12.83 ***
[t-10min, t-5sec]	312	83.03	-2.91	85.49	-2.17	69.62	2.25	72.90	1.74
[t-5min, t-5sec]	312	95.03	-12.50	96.26	-11.26	76.11	2.56	77.78	0.30
[t-3sec, t+3sec]	312	218.78	-60.31 ***	234.56	-60.31 ***	91.02	-42.99 ***	81.53	-45.43 ***
[t-5sec, t+5sec]	312	175.39	-51.29 ***	183.01	-51.48 ***	95.16	-18.29**	89.44	-27.11*
[t-5sec, t+10sec]	312	159.39	-28.30 ***	164.06	-33.13 ***	94.45	-21.58**	85.00	-22.31**
[t-5sec, t+30sec]	312	135.16	-18.58*	138.12	-19.56*	89.08	-9.50	74.45	-17.95*
[t-5sec, t+1min]	312	123.95	-18.79	128.01	-18.79	78.60	-9.46*	63.48	-14.42**
[t-5sec, t+2min]	312	109.36	-13.29	113.07	-13.79	75.82	-7.41	68.29	-11.61*
[t-5sec, t+5min]	312	96.54	-2.41	100.65	-1.73	70.01	3.55	64.41	-3.11
[t-5sec, t+10min]	312	98.13	-0.56	101.35	-0.33	65.56	4.39	61.12	2.96
[t-5sec, t+30min]	295	73.66	-0.59	76.64	-0.18	52.01	1.33	47.55	-1.22
[t-5sec, t+60min]	279	54.16	-8.76**	56.56	-8.76*	40.88	-5.37	37.80	-5.75
[t-5sec, t+120min]	233	36.08	-12.32**	38.27	-11.89*	22.92	-15.15**	21.92	-13.77*
[t+5sec, t+5min]	312	93.87	-3.89	97.86	-3.89	69.16	-0.72	63.56	-4.40
[t+1min, t+10min]	312	95.03	0.56	98.14	0.56	63.99	1.70	60.84	0.81
[t+1min, t+30min]	295	71.78	-0.95	74.72	-0.47	50.86	0.96	46.82	-1.60
[t+1min, t+60min]	279	52.98	-9.70 ***	55.35	-9.70 ***	40.00	-5.84*	37.17	-6.24*
[t+1min, t+120min]	233	35.28	-12.05**	37.46	-11.89**	22.22	-16.05**	21.37	-13.66*

Mean and median Cumulative Abnormal Volumes (CAVs) are reported for highly Relevant, Novel Alerts with positive and negative (Net) Sentiment scores of at least 70 per cent. CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

It is interesting that this runs completely contrary to the pattern observed in mean CAVs, which are heavily right-skewed and appear to respond to all the TRNA metrics. The gains clearly seem proportional to the level of positive sentiment, as observed by comparing Figures 17a, 17c, and 17e, but unlike the CAR analysis, mean CAV values are actually higher on average for positive news than negative news.

Table 15 shows that the mean CAVs soar as much as 235 per cent above their normal levels in the [t-3sec, t+3sec] window for highly positive, relevant, novel Alerts, which points to some stocks potentially being subject to higher levels of algorithmic trading than others. Table 16 highlights the top 11 stocks traded in the [t-3sec, t+3sec] window, which were identified as having positive median CAVs for Alerts. It is worth noting though that these results are highly volatile depending on the event window chosen, and subject to the type of news reported for each company over the sample horizon, and should thus be interpreted accordingly.[9] Note that all individual stocks had negative median CAVs for Articles, consistent with the overall results.

Note the same analysis is conducted on Trades, yielding near identical results. Additional results are reported in Appendix G and Appendix H.

---

[9] There is no way to guarantee there will be market reaction-worthy news for every individual stock over the sample horizon, nor that these will be reported by Thomson Reuters and tagged correctly by the TRNA algorithm. That does not necessarily mean that these stocks would not otherwise be traded algorithmically should the right opportunity present itself.

Table 16

Summary of CAV Results for the Top Traded Individual Stock Alerts

			Alerts												
								CAVs (%)							
Stock	Total N	% Sig. Windows	N Alerts	Avg. R	Avg. N	Avg. Net Tone +ive News	Avg. Net Tone -ive News	Information Leakage		Knee-jerk Reaction				Secondary Reaction	
								[t-120min, t-5sec]		[t-3sec, t+3sec]		[t-5sec, t+1min]		[t+1min, t+120min]	
								Mean	Median	Mean	Median	Mean	Median	Mean	Median
AAPL	6498	66.04%	948	0.99	3.40	0.31	-0.60	47.57	20.21	126.33	27.71	122.04	38.10	45.66	9.68
CELG	507	43.19%	133	1.00	1.83	0.31	-0.50	53.11	16.12	358.84	23.51	339.96	168.47	119.44	81.63
GILD	537	40.17%	109	0.98	1.28	0.35	-0.60	26.12	9.30	128.90	13.95	155.44	44.09	40.27	43.17
YHOO	1278	39.83%	175	0.97	1.21	0.33	-0.62	91.47	-6.55	427.72	27.52	382.62	58.15	285.75	23.58
BBBY	242	28.23%	26	0.99	6.92	0.51	-0.70	166.75	82.81	133.96	41.93	292.75	144.07	190.03	113.28
PCAR	128	26.81%	40	0.97	1.08	0.52	-0.72	-8.50	-20.95	1235.42	18.28	435.55	56.43	31.76	-8.95
DLTR	240	19.86%	47	0.97	2.36	0.46	-0.57	-13.17	-22.97	92.49	25.72	76.21	9.13	38.75	2.33
MAT	205	19.10%	34	1.00	0.91	0.41	-0.58	-7.60	187.32	74.69	117.59	149.44	180.68	48.28	173.90
PAYX	158	16.15%	29	0.96	0.59	0.41	-0.48	-0.04	2.57	273.25	3.99	112.84	-14.98	2.75	-3.56
INTU	149	13.51%	20	1.00	2.70	0.53	-0.66	100.64	29.85	152.08	103.99	200.35	75.02	127.66	43.74
XLNX	60	5.28%	9	1.00	0.67	0.49	-0.68	4.99	8.51	1040.19	93.30	128.90	-37.24	12.26	-17.79

This table summarizes the results for individual stock Alerts with positive median Cumulative Abnormal Volumes (CAVs) in the [t-3sec, t+3sec] window. Median and mean CAVs are reported for four event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance (R), Novelty (N), and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests). Note results for all individual stocks can be found in Appendix G.

Figure 17

Mean and Median CAVs for Relevant, Novel Nasdaq News across Different (Net) Sentiment Thresholds

---

The six charts that follow report the median and mean Cumulative Abnormal Volumes (CAVs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, Relevance is set to 1 (most relevant news), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

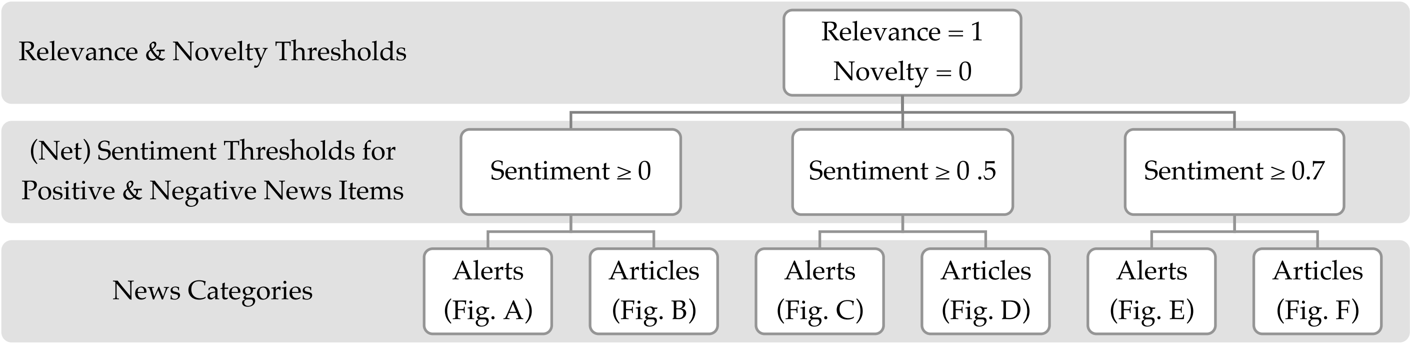
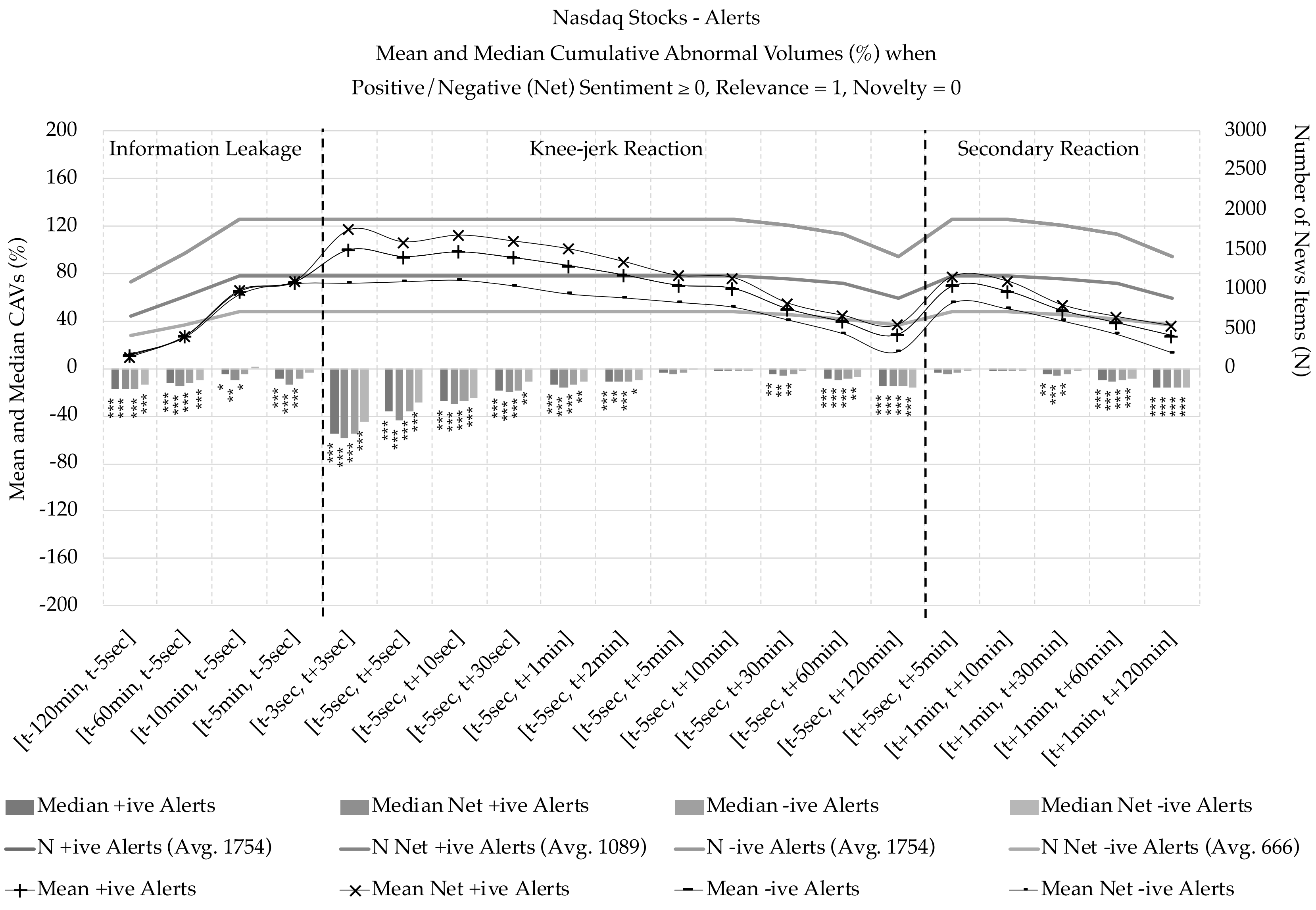


Figure 17a

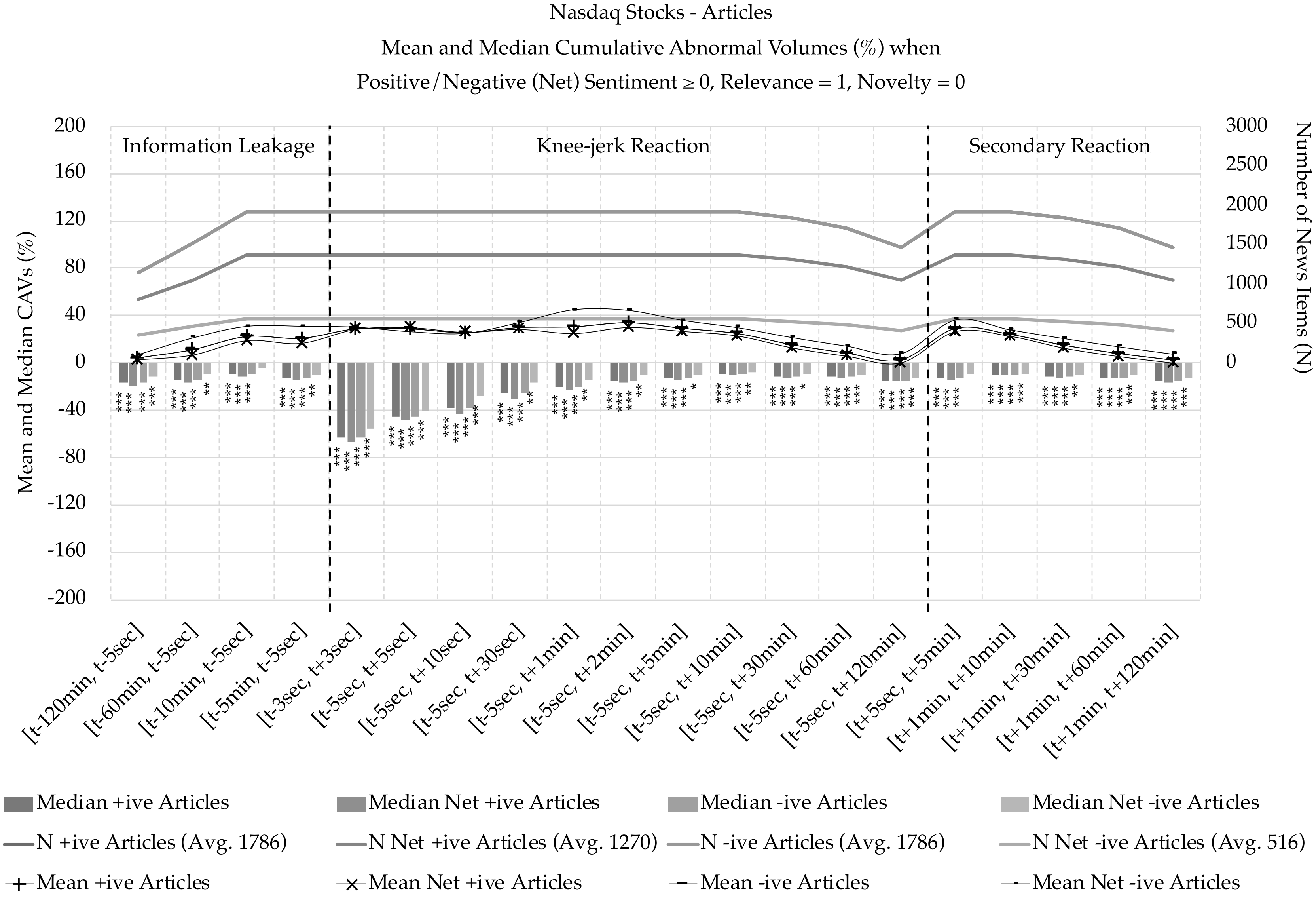
Mean and Median CAVs for Relevant, Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 17b

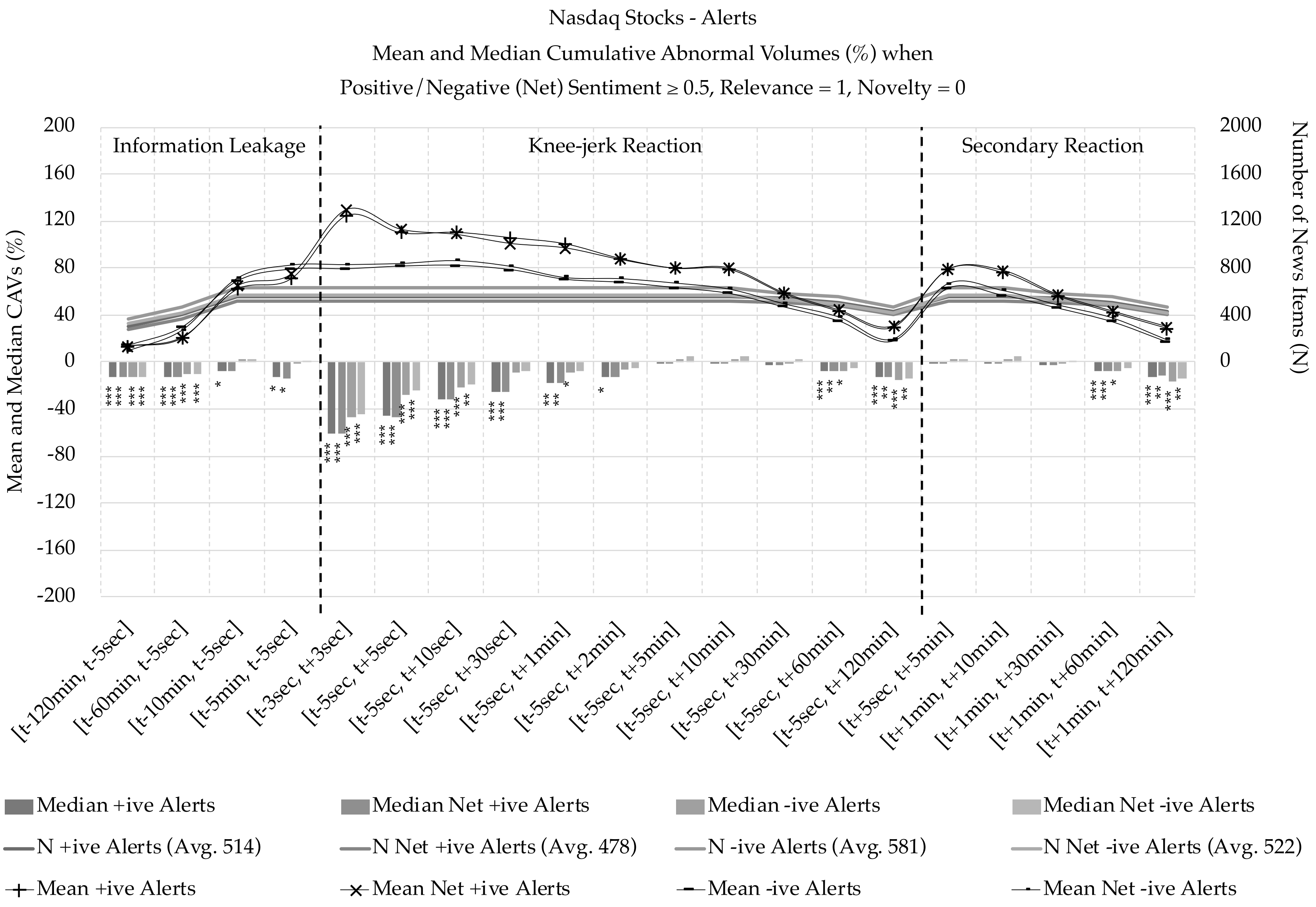
Mean and Median CAVs for Relevant, Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 17c

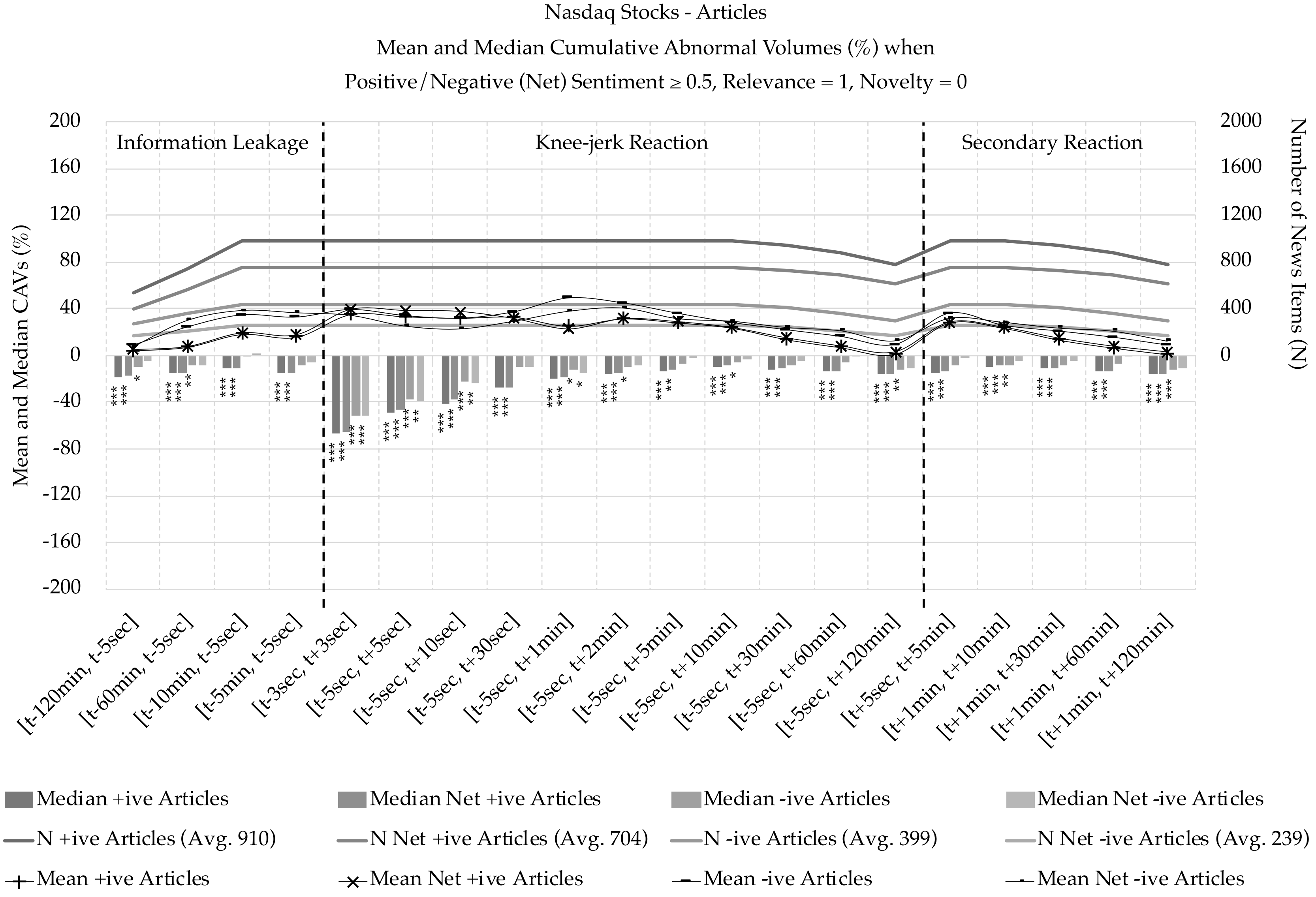
Mean and Median CAVs for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 17d

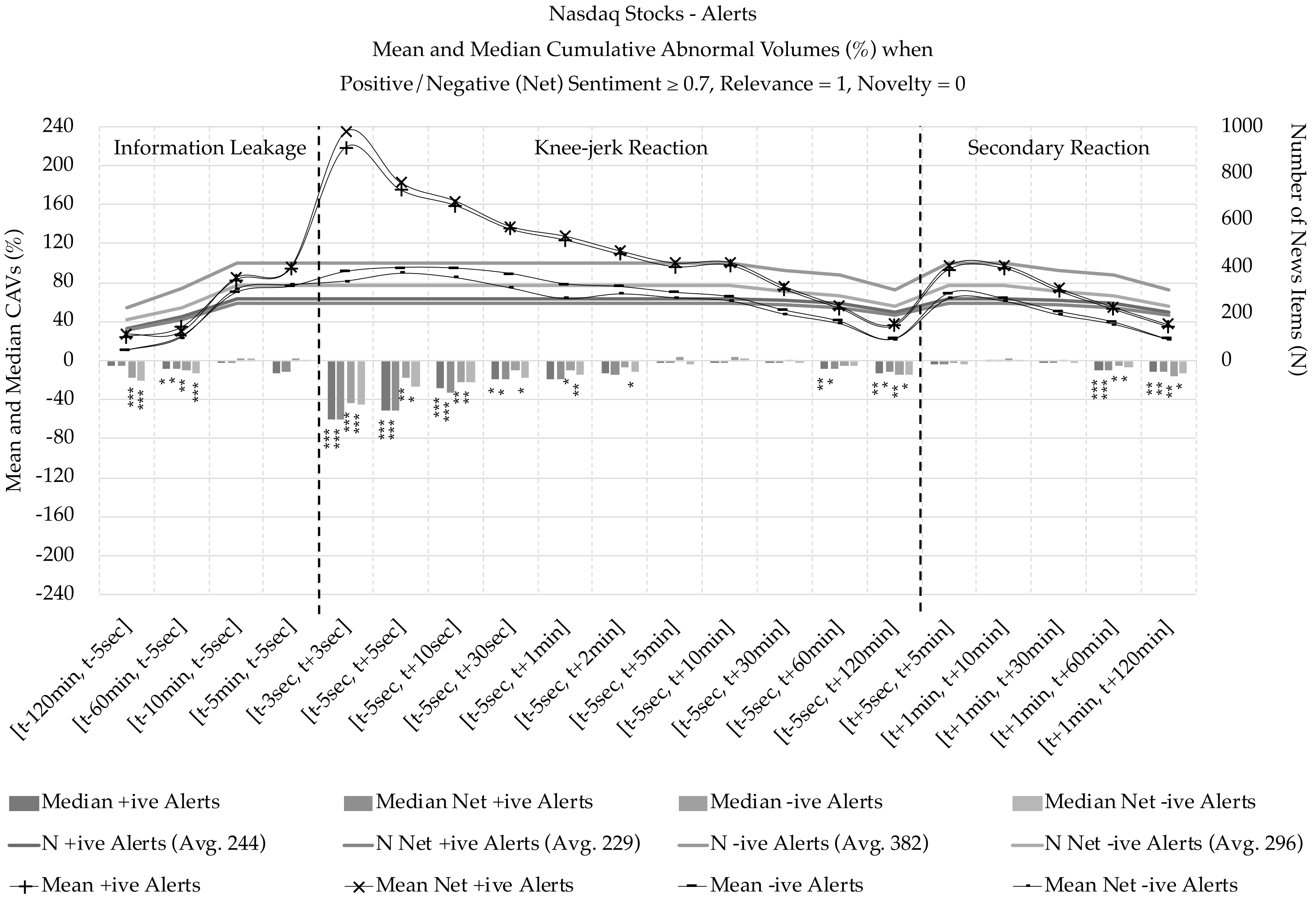
Mean and Median CAVs for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 17e

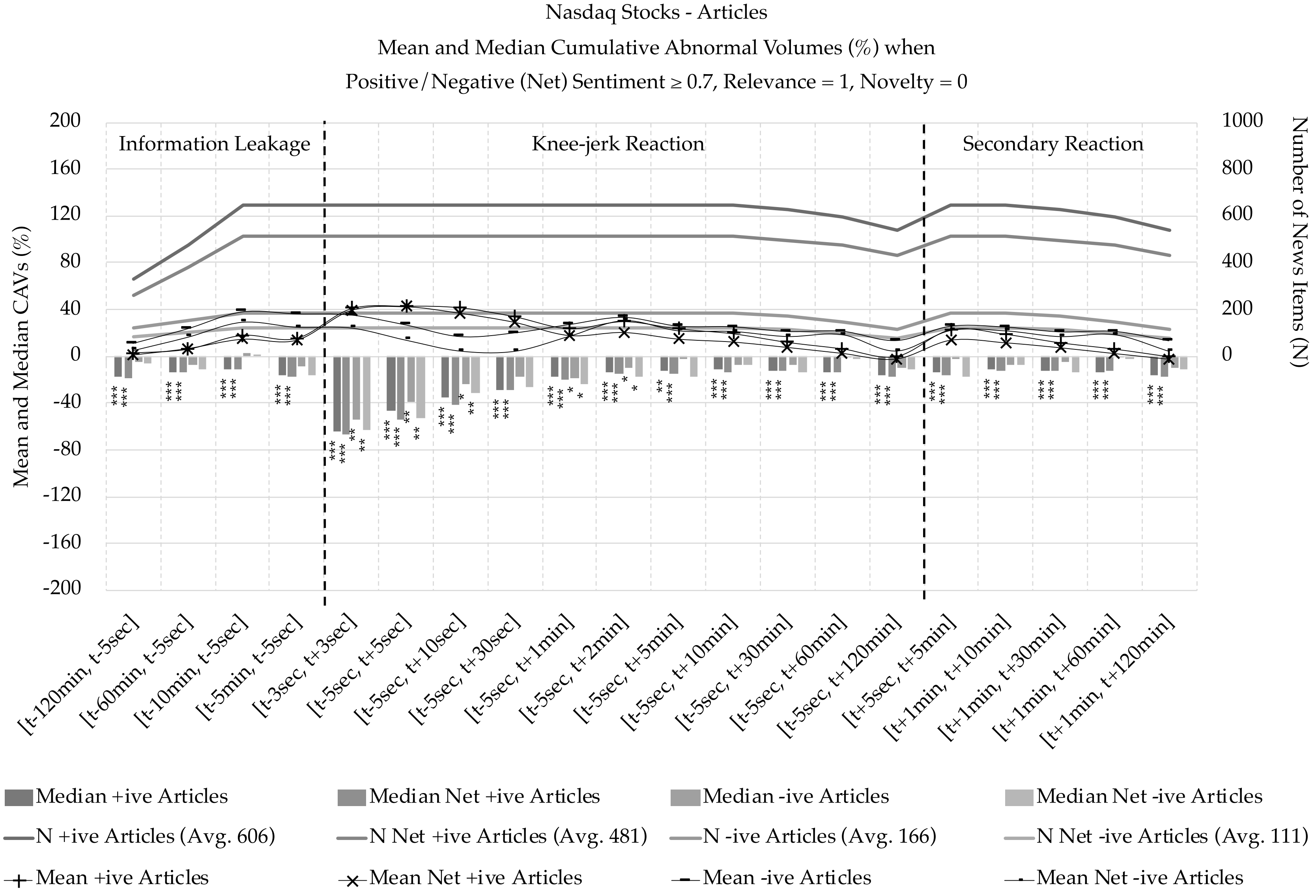
Mean and Median CAVs for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 17f

Mean and Median CAVs for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18

Mean and Median CAVs for Relevant Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Volumes (CAVs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, Relevance is set to 1 (most relevant news), no threshold is set for Novelty (all novelty scores are included), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

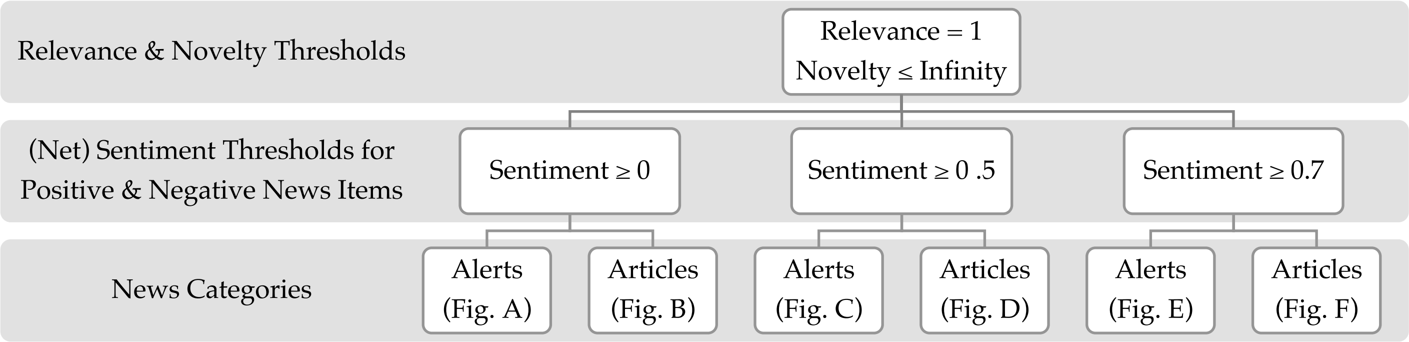
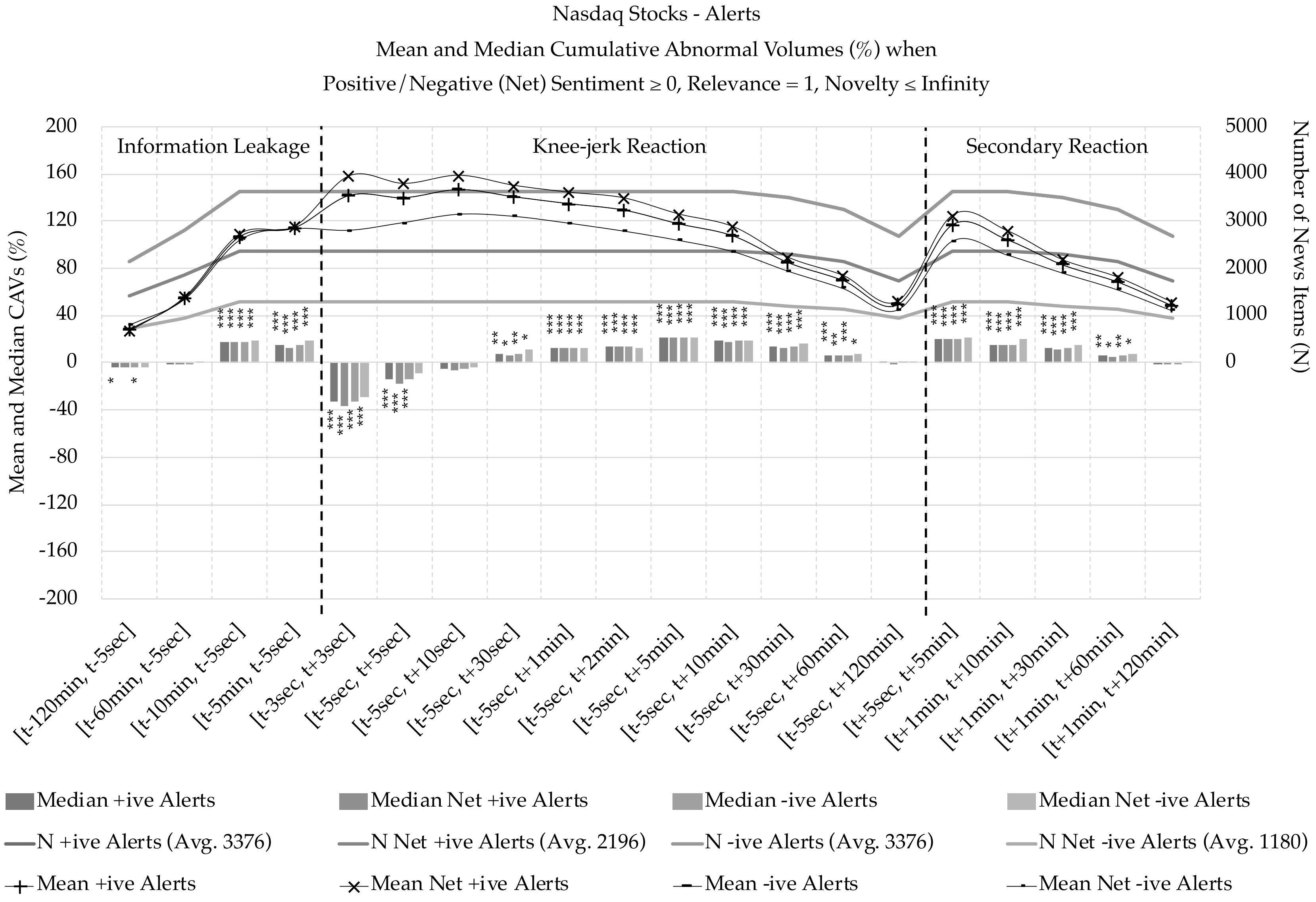


Figure 18a

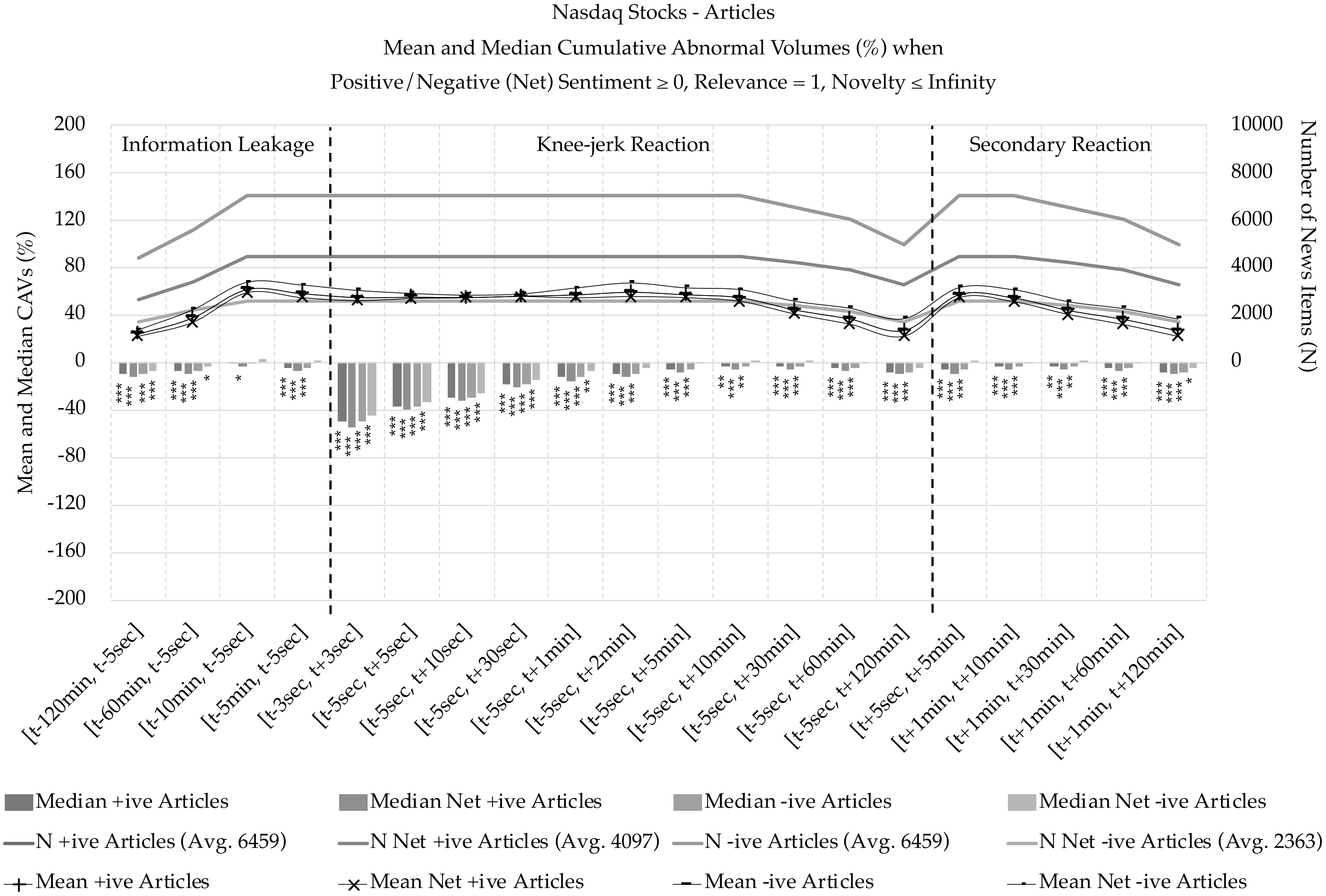
Mean and Median CAVs for Relevant Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18b

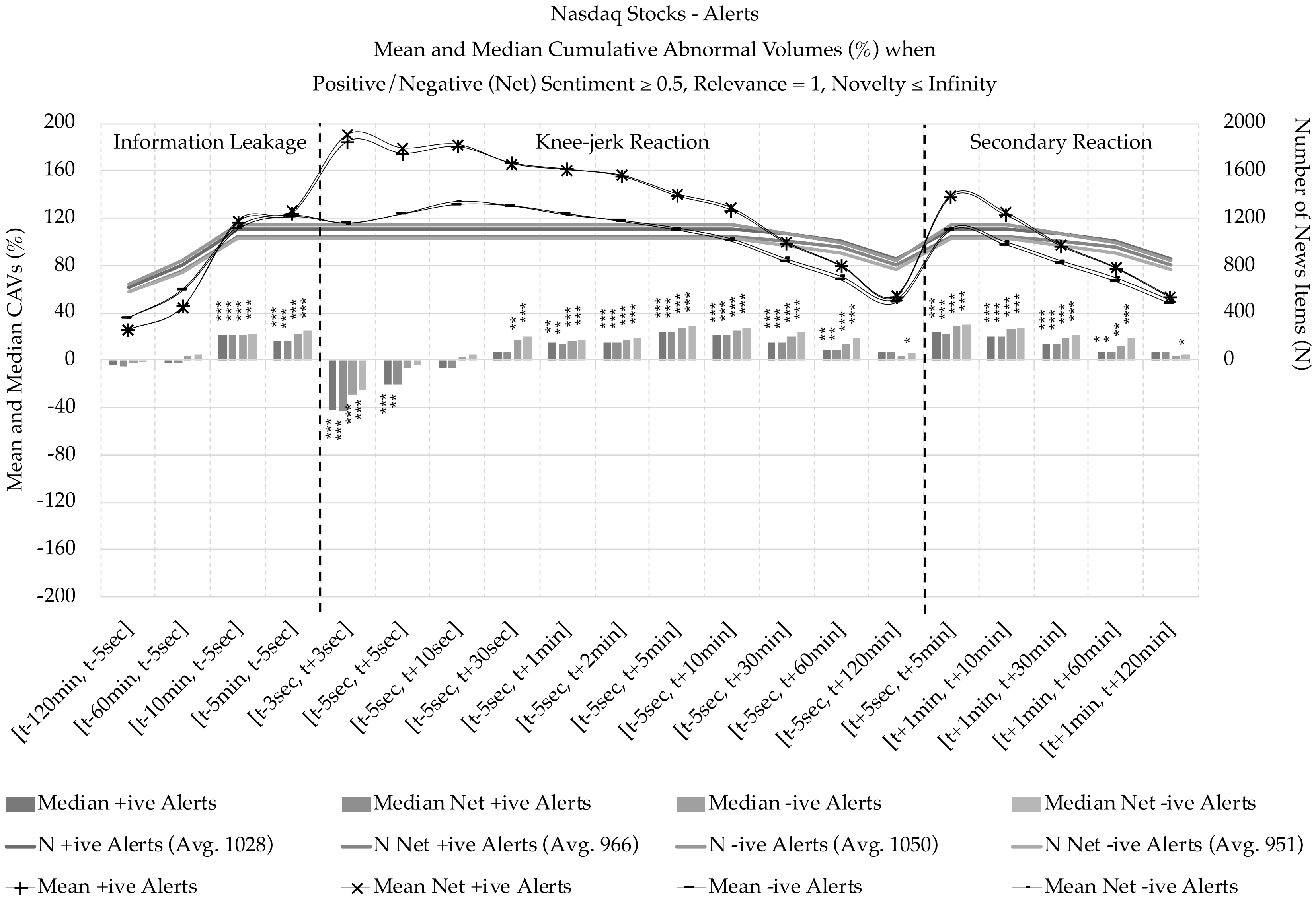
Mean and Median CAVs for Relevant Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18c

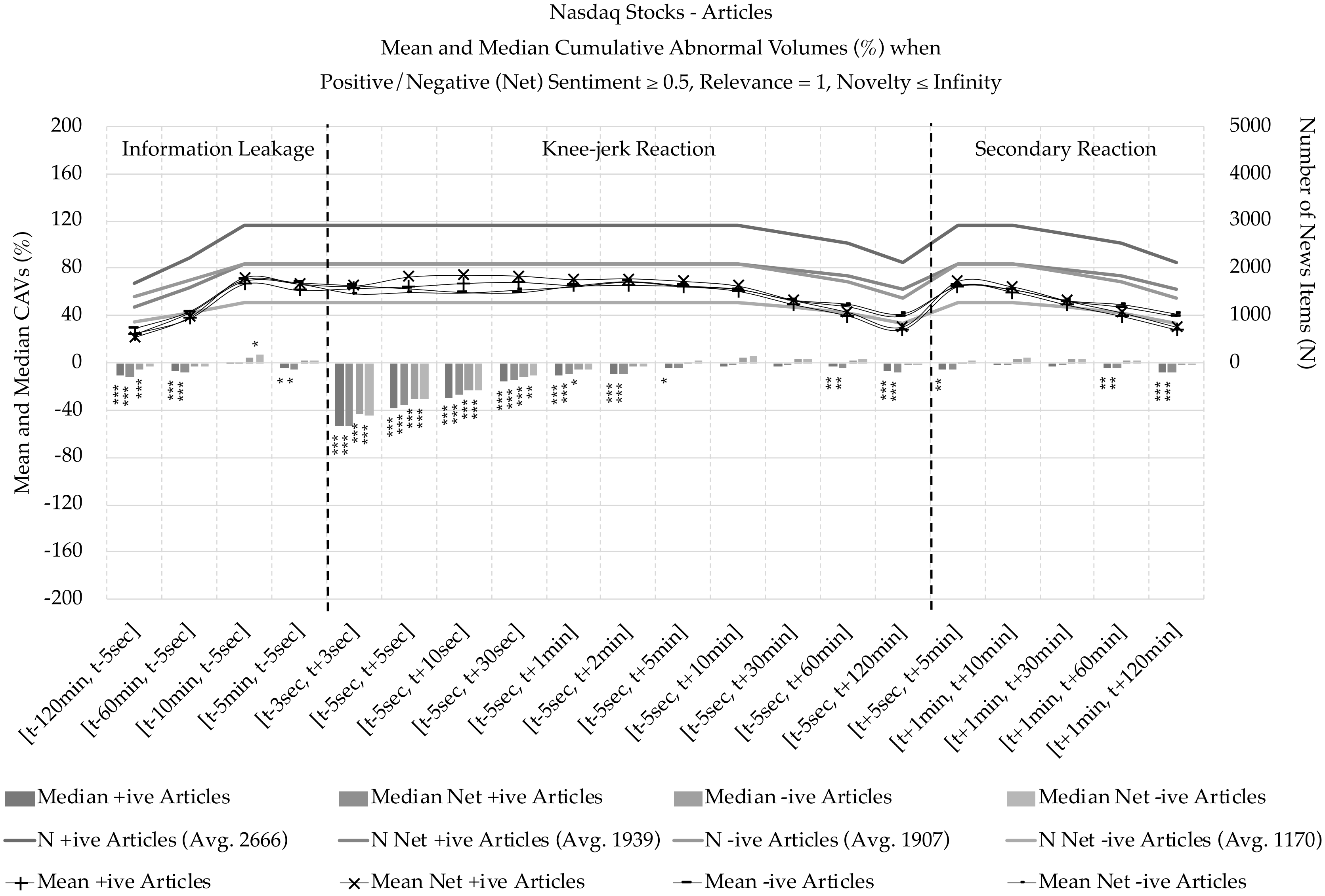
Mean and Median CAVs for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18d

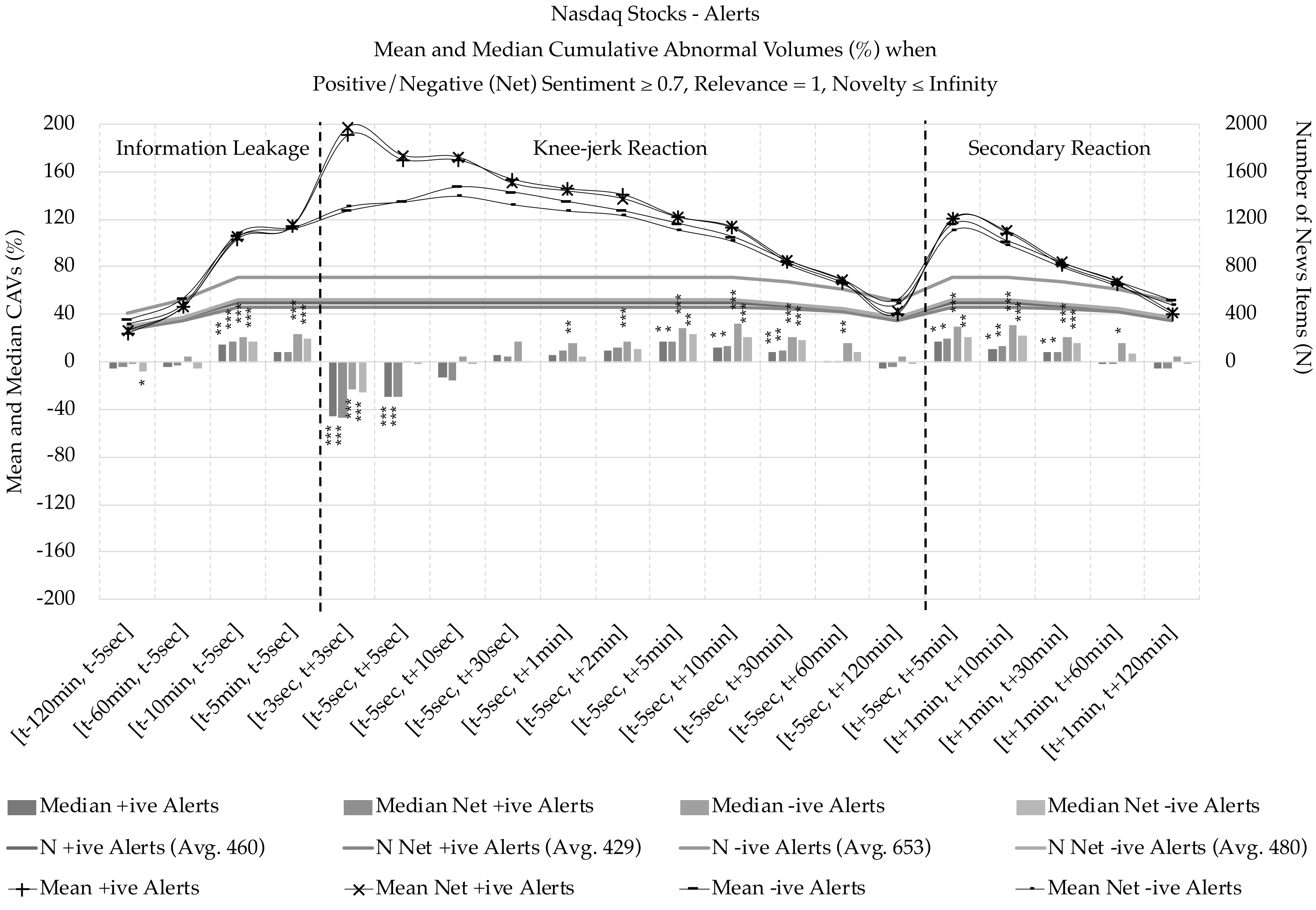
Mean and Median CAVs for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18e

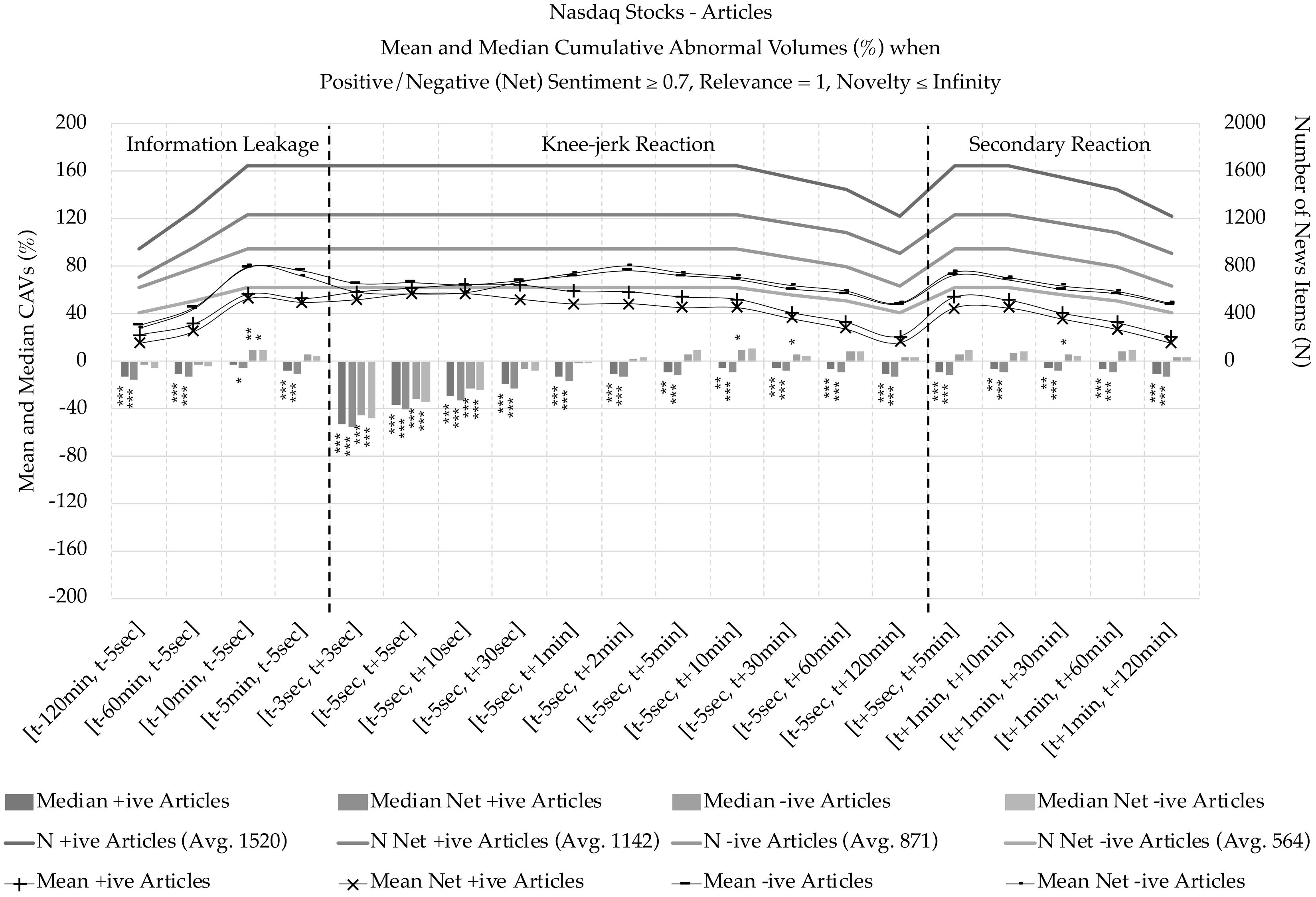
Mean and Median CAVs for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure 18f

Mean and Median CAVs for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

5.5. Deep Dive of Individual Stock Results

Investigating individual news releases in the context of competing newswires and the market microstructure reveals the many complexities of newsflow analysis. I examine the underlying news and market dynamics for a subset of the most novel, relevant, and polarized news items in the sample. These news items generated the most pronounced CAR and CAV results in the first minute's worth of Knee-jerk Reaction windows, allowing me to better demonstrate some of the caveats to analyzing high frequency news data. Although extrapolating a deep dive of this nature to the entire news sample is beyond the scope of my research, this exercise is an opportunity to illustrate the limitations and shortfalls that should be addressed to make future studies more robust.

5.5.1.PACCAR's Share Repurchase Announcement

The most aggressively traded news item in the [t-3sec, t+3sec] window was an Alert announcing a $300M share repurchase by PACCAR on September 23, 2015. The unscheduled announcement incited a 27,788 per cent spike in abnormal volume and 71bps in abnormal returns over the [t-3sec, t+3sec] window, visible in Figure 19. These results were driven by 546 trades that occurred when the news was released at 10:00:00. In terms of its TRNA metrics, the Alert was classified as both relevant and novel, and was attributed a highly positive net tone of 83 per cent, all of which are consistent with the observed knee-jerk reaction. However, upon closer inspection, it appears that Thomson Reuters was actually the third major newswire (that I know of) to report the news, with Bloomberg and Dow Jones releasing headlines 353ms and 294ms earlier, respectively:

Table 17

First Headline from each Major Newswire Reporting on PACCAR's Share Repurchase

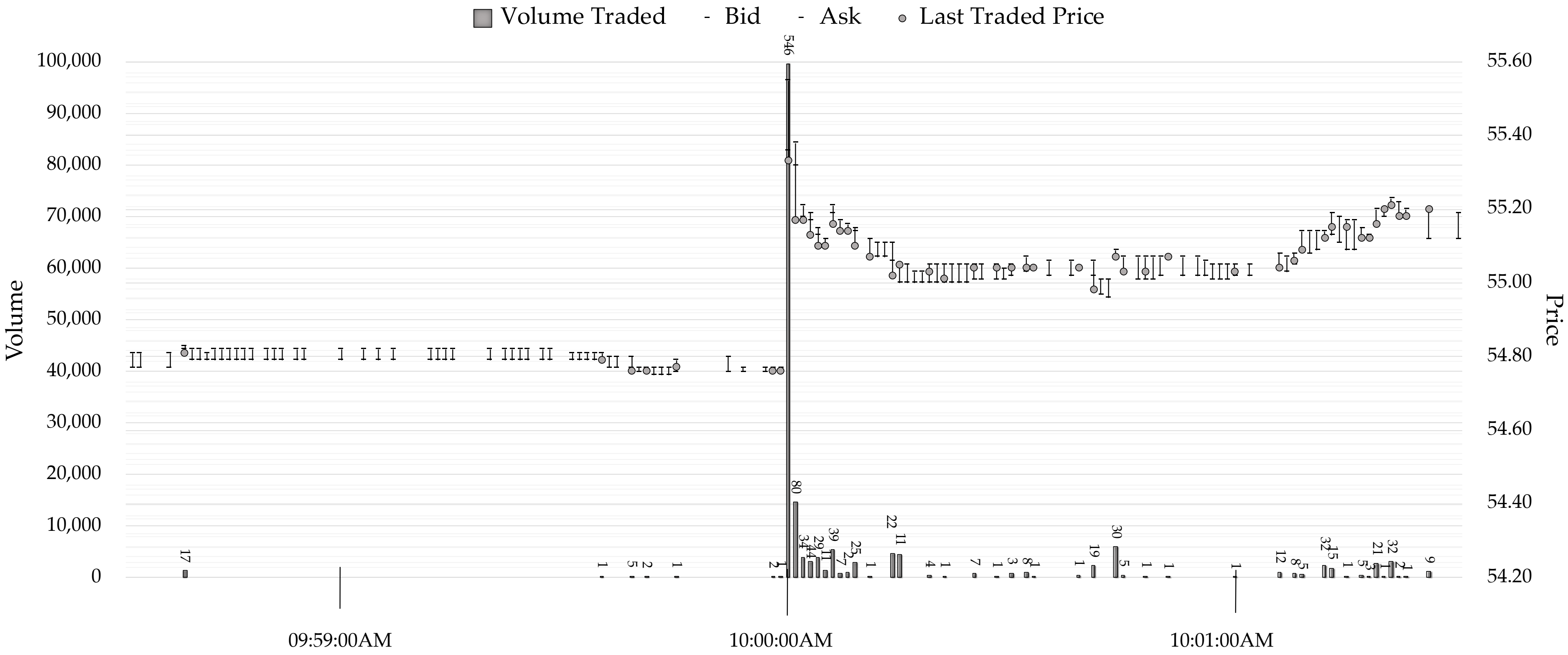
Timestamp (ET)	Source	Headline
2015-09-23 10:00:00.140	Bloomberg	*PACCAR REPORTS $300M SHR REPURCHASE
2015-09-23 10:00:00.209	Dow Jones	DJ PACCAR ANNOUNCES $300 MILLION SHARE REPURCHASE
2015-09-23 10:00:00.493	Thomson Reuters	PACCAR ANNOUNCES $300 MILLION SHARE REPURCHASE

Note timestamps are rounded to the nearest millisecond.

While these differences may seem inconsequentially small, by plotting the competing timestamps against the underlying PACCAR trades in Figure 20, we can see when each individual trade took place between 10:00:00 and 10:00:01 relative to each newswire's headline. As it turns out, 125 trades occurred between the Bloomberg and Dow Jones timestamps, for a total of 4,180 contracts traded. Another 248 trades of 21,912 contracts took place between the Dow Jones and Thomson Reuters timestamps, followed by 128 trades of 33,428 contracts after the Thomson Reuters timestamp until the end of the second. The volume-weighted-average-price (VWAP) of the later trades that took place after the TRNA timestamp are 10 cents or 18.2bps higher than those that occurred immediately following the first Bloomberg headline. In terms of the [t-3sec, t+3sec] CAR, this constitutes a 25 per cent penalty for latency that hasn't been accounted for in my analysis. Hence, as outlined in the limitations section, the concept of novelty is not exhaustive when trading exclusively using the TRNA data, and the results presented are an optimistic oversimplification of the TRNA's true trading potential.

Figure 19

Market Reaction to PACCAR $300M Share Repurchase Announcement

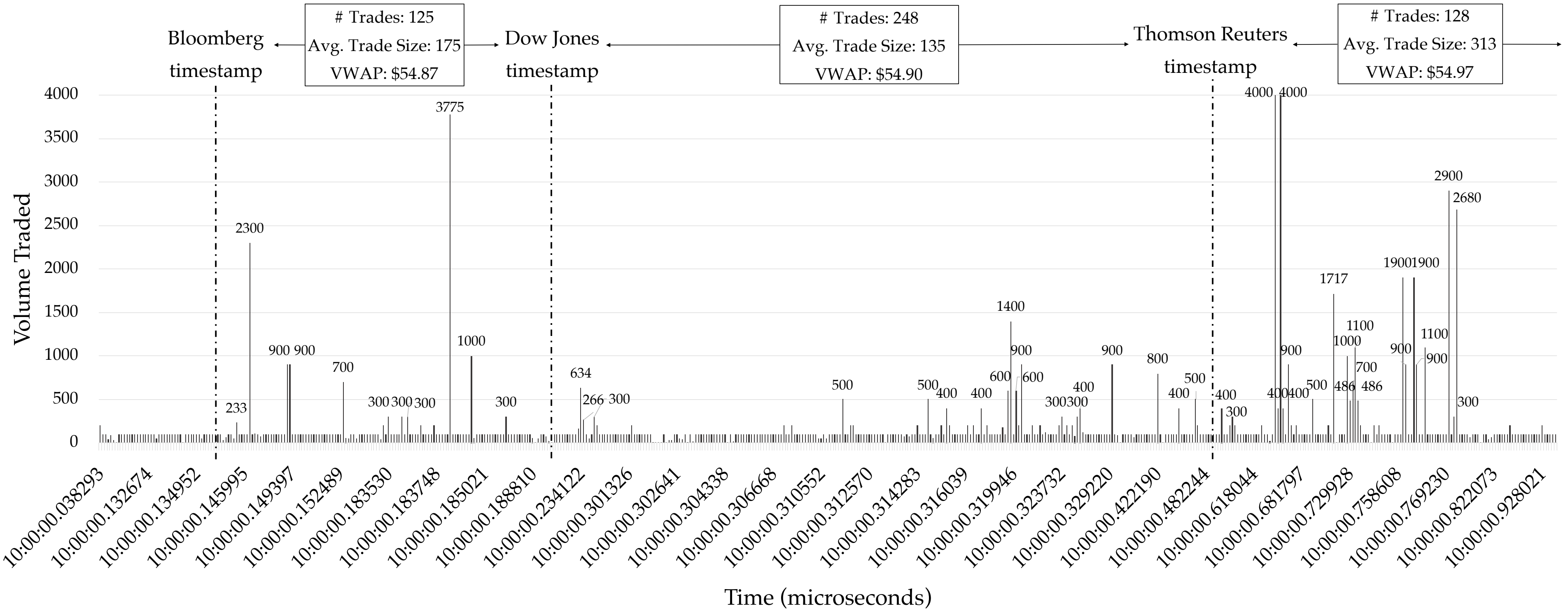


[t0-3sec, t0+3sec] CAV: 27,788%; CAR: 71.06 bps

This chart presents a second-by-second snapshot of the market activity that occurred PACCAR stock in the one and a half minutes before and after the news was released at 10:00:00. The last traded price as well as the best bid and ask are plotted on the right axis, alongside the volume traded on the left axis. Note the labels on the volume bars indicate the number of trades in that second.

Figure 20

Individual Trades of PACCAR Stock from 10:00:00-10:00:01AM



This chart shows all 546 trades in PACCAR stock that occurred between 10:00:00 and 10:00:01 (note time gaps are irregularly spaced). The timestamps of the first Bloomberg, Dow Jones, and Thomson Reuters headlines are overlaid to show the approximate timing of trading activity relative to each newswire. The number of trades, average trade size, and volume-weighted-average-price (VWAP) between each subsequent headline are reported so as to easily compare the latency penalty that results from trading "late" in a high-frequency setting. Note the volume of trades in excess of 200 contracts are labelled.

5.5.2. Vertex's Announcement on Incivek Treatment

A progress update on Vertex's Phase II study for its hepatitis C treatment called Incivek generated the highest CAR values in the event windows spanning from [t-5sec, t+5sec] to [t-5sec, t+30sec] inclusively, generating a CAR of 112bps in the [t-5sec, t+5sec] window, driven by a 2,136 per cent spike in CAV. The Alert in question, shown in Table 18, was tagged as relevant, novel, and was scored a highly net positive tone of 82 per cent, consistent with the direction of the price move observed in Figure 21.

Table 18

Target Headline for Vertex Announcement on Incivek Performance

Timestamp (ET)	Source	Headline
2012-03-06 11:00:10.794	Thomson Reuters	VERTEX PHARMACEUTICALS INC <VRTX.O> SAYS INCIVEK WAS WELL TOLERATED WITH COMMONLY USED ATRIPLA

Note timestamps are rounded to the nearest millisecond.

However, a deep dive reveals that this Alert was in fact almost 10 seconds old, as Table 19 shows the story first broke at 11:00:01 via Dow Jones (DJ1) citing the press release. Oddly enough, the Alert wasn't even the first relevant, novel Thomson Reuters headline on this particular story, as they also published an Article (TR1) at 11:00:01, and a headline (TR2) at 11:00:06, over four seconds before the flagged timestamp from Table 18 (TR3). Interestingly, even though these earlier Thomson Reuters news items both reported qualitatively similar results for Incivek's effectiveness, they received net negative tone scores of -0.59 and -0.67 respectively, which stand in stark contrast to the TR3's highly positive net tone of 0.82. Sentiment scores appear exceedingly sensitive to the exact wording used to describe the news. It is also curious that the TRNA algorithm scored all three news items as novel, despite two of them having the same altID (unique story number). Yet again, novelty proves itself a tricky variable to measure, even within the same news database.

The knee-jerk reaction depicted in Figure 21 shows a clear increase in trading activity and a brief uptick in Vertex's share price following the news. If we accept the market's initial reaction to be the correct interpretation of the news, you could argue that it took the TRNA algorithm three tries to correctly categorize the net tone of this story. From a trading perspective, this level of disagreement implies that if all three news items had been traded according to their respective TRNA scores, the net CAR over the [t-5sec, t+5sec] window would have been roughly -50bps, as the losses from trading the first two negative sentiment scores would have eroded the gains from the third positive score. Ultimately though, the uptick in prices was rather short-lived, begging the question, was the news in fact net neutral to begin with?

The breakdown of trading activity over the first minute is detailed in Figure 22, with each of the headlines from Table 19 overlaid by source. The pattern that emerges shows a wave of trading that begins after the first 15 seconds worth of news comes out, which includes four Thomson Reuters headlines (TR1-TR4), two Dow Jones headlines (DJ1-DJ2), and one Bloomberg headline (BN1). A second wave takes shape after 30 seconds and includes at least two more Bloomberg headlines (BN2-BN3). In this case, it's not clear which newswire(s) actually prompted the spike in trading activity, as the news does not appear to be traded algorithmically.

Table 19

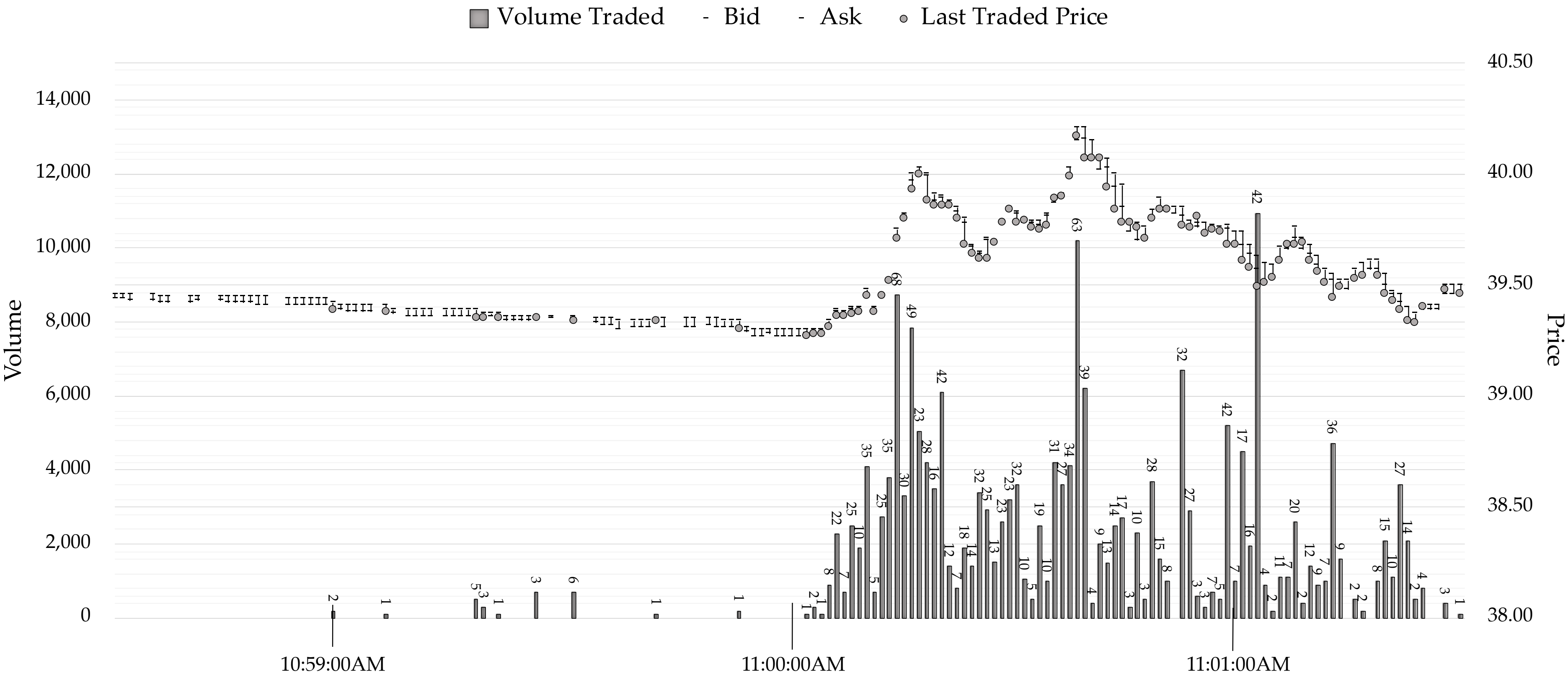
Story Progression for Vertex Announcement on Incivek Performance

Timestamp (ET)	Source (Tag)	Net Tone	Headline
2012-03-06 11:00:01.119	Dow Jones (DJ1)	N/A	PRESS RELEASE: Data from Phase 2 Study of an INCIVEK(R) Combination Regimen Showed 74% of People Co-Infected with Hepatitis C and HIV Had Undetectable Hepatitis C Virus 12 Weeks After Treatment Ended (SVR12)
2012-03-06 11:00:01.128	Thomson Reuters (TR1)	-0.59	Vertex, Merck hepatitis drugs work in HIV patients
2012-03-06 11:00:05.562	Bloomberg (BN1)	N/A	*DATA FROM PHASE 2 STUDY OF AN INCIVEK(R) COMBINATION REGIMEN
2012-03-06 11:00:06.734	Thomson Reuters (TR2)	-0.67	DATA FROM PHASE 2 STUDY OF AN INCIVEK(R) COMBINATION REGIMEN SHOWED 74% OF PEOPLE CO-INFECTED WITH HEPATITIS C AND HIV HAD UNDETECTABLE HEPATITIS C VIRUS 12 WEEKS AFTER TREATMENT ENDED (SVR12)
2012-03-06 11:00:10.794	Thomson Reuters (TR3)	0.82	VERTEX PHARMACEUTICALS INC <VRTX.O> SAYS INCIVEK WAS WELL TOLERATED WITH COMMONLY USED ATRIPLA
2012-03-06 11:00:13.499	Thomson Reuters (TR4)	-0.29	Data from Phase 2 Study of an INCIVEK® Combination Regimen Showed 74% of People Co-Infected with Hepatitis C and HIV Had Undetectable Hepatitis C Virus 12 Weeks After Treatment Ended (SVR12)
2012-03-06 11:00:15.481	Dow Jones (DJ2)	N/A	*DJ VERTEX: DATA FROM PHASE 2 STUDY OF AN INCIVEK(R) COMBINATION REGIMEN SHOWED 74% OF PEOPLE CO-INFECTED WITH HEPATITIS C AND HIV HAD UNDETECTABLE HEPATITIS C VIRUS 12 WEEKS AFTER TREATMENT ENDED (SVR12)
2012-03-06 11:00:25.427	Bloomberg (BN2)	N/A	*VERTEX SAYS DATA SHOWS 74% UNDETECTABLE VIRUS AFTER 12 WEEKS
2012-03-06 11:00:36.169	Bloomberg (BN3)	N/A	*VERTEX SAYS NO PATIENTS EXPERIENCED HIV BREAKTHROUGH :VRTX US

These headlines show the sequence in which the Vertex announcement broke across Bloomberg, Dow Jones, and Thomson Reuters newswires. Note timestamps are rounded to the nearest millisecond.

Figure 21

Market Reaction to Vertex Press Release on Incivek Treatment Performance

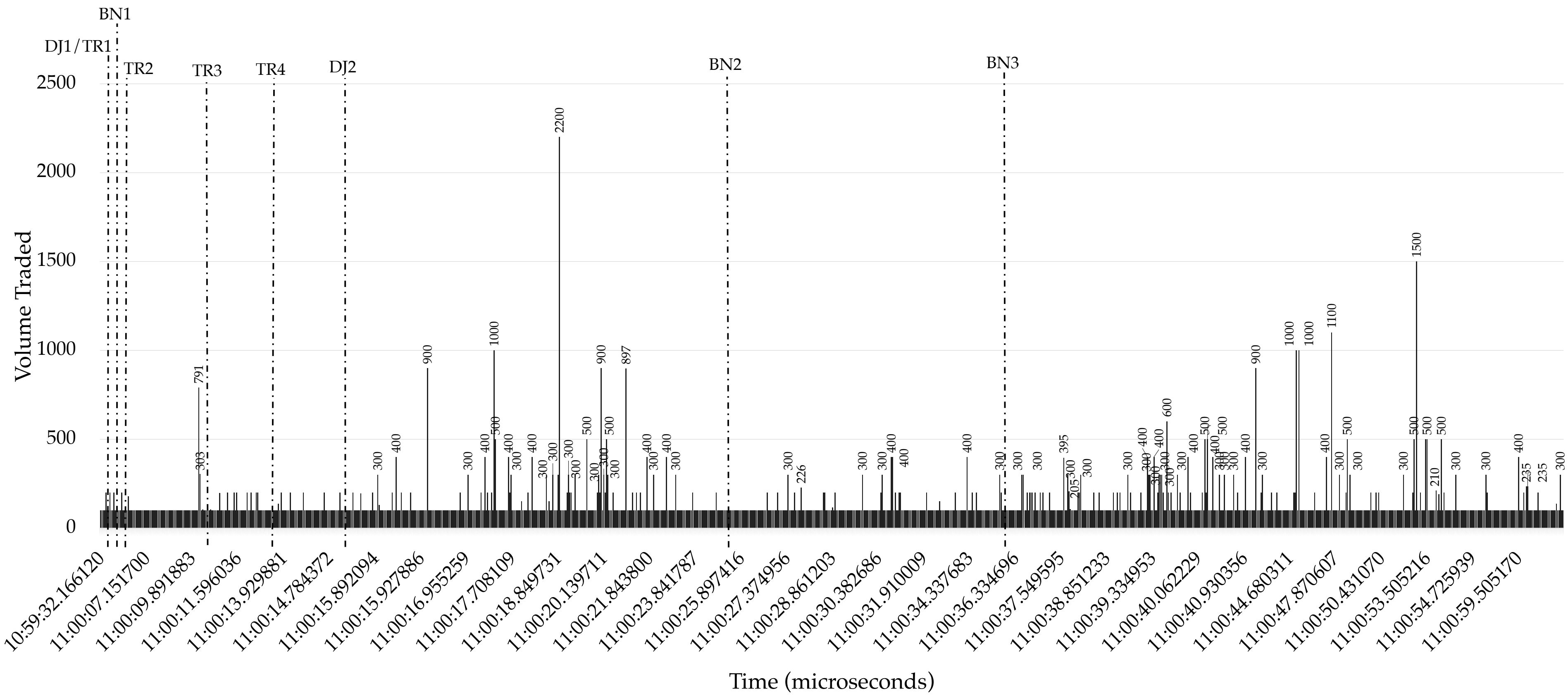


[t0-5sec, t0+5sec] CAV: 2,136%; CAR: 111.76 bps

This chart presents a second-by-second snapshot of the market activity that occurred in Vertex stock one and a half minutes before and after the news was released at 11:00:01. The last traded price as well as the best bid and ask are plotted on the right axis, alongside the volume traded on the left axis. Note the labels on the volume bars indicate the number of trades in that second.

Figure 22

Individual Trades of Vertex Stock from 11:00-11:01AM



This chart shows all Vertex stock trades that occurred between 11:00 and 11:01 (note time gaps are irregularly spaced). The timestamps of the first few Bloomberg (BN), Dow Jones (DJ), and Thomson Reuters (TR) headlines are overlaid to show the approximate timing of trading activity relative to the distribution of news. See Table 19 for the individual headlines. Note the volume of trades in excess of 200 contracts are labelled.

5.5.3. Symantec Hacker Threat

On Friday the 13th, 2012, a Tweet from hacker "Yama Tough" threatened to release the source code for Symantec's flagship Norton Antivirus software. The Tweet found its way into a Thomson Reuters Alert that was flagged as relevant, novel, and highly negative, with a net sentiment score of -68 per cent:

Table 20

First Headline from Symantec Hacker Threat

Timestamp (ET)	Source	Headline
2012-01-13 12:02:47	Thomson Reuters	HACKER 'YAMA TOUGH' SAYS VIA TWITTER WILL RELEASE SOURCE CODE FOR SYMANTEC CORP'S <SYMC.O> NORTON UTILITIES SOFTWARE TODAY

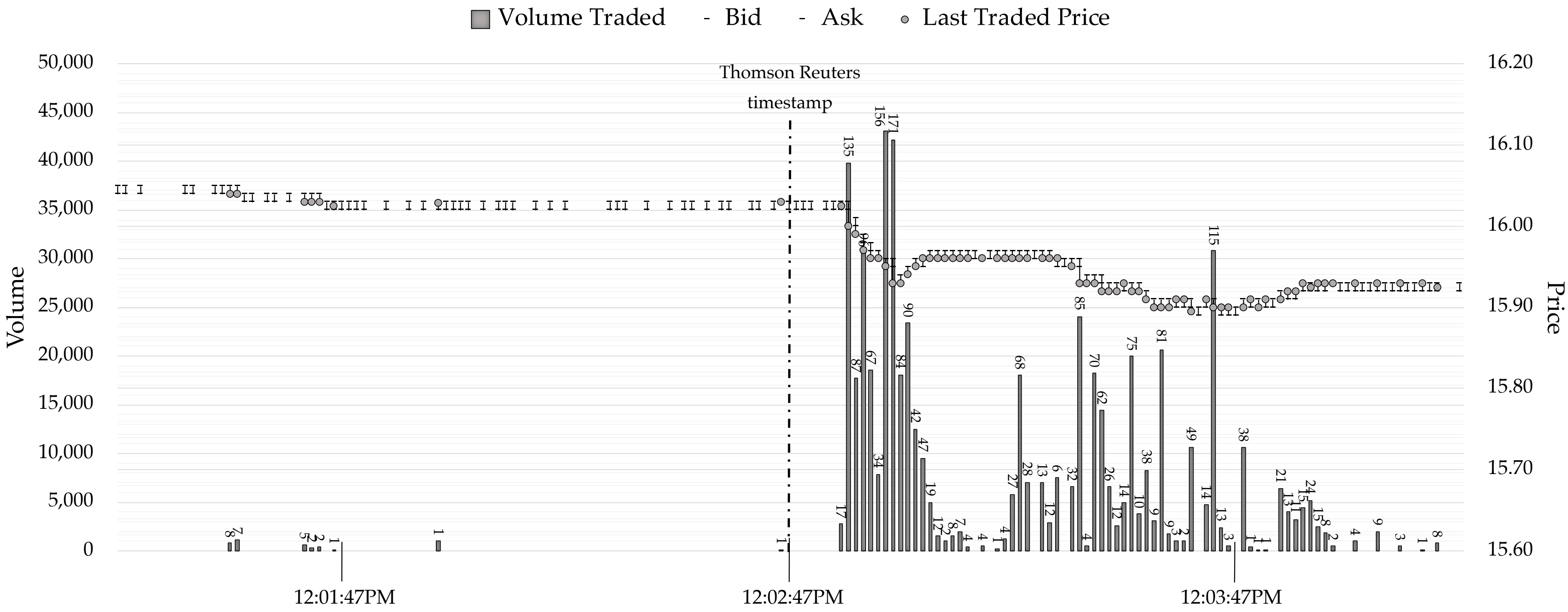
Note timestamps are rounded to the nearest millisecond.

The Alert was among the most heavily traded in the [t-5sec, t+30sec] and [t-5sec, t+1min] windows, generating a CAR of 76bps and a 3,774 per cent surge in CAV during the [t-5sec, t+30sec] window.

Apart from the Tweet itself (whose timestamp could not be traced), Thomson Reuters appears to have had the exclusive on this news, with no other major newswires releasing any headlines about it. The knee-jerk reaction shown in Figure 23 corroborates this suspicion, as it took a full seven seconds for markets to begin selling Symantec stock on the news. The fact that the TRNA metrics were consistent with the ensuing reaction, and that the news was in fact novel and relevant, suggests that this particular news item would have been a good target for an algorithmic trading strategy, at the time. However, it's worth noting that since 2012, market participants have come to treat Twitter as a veritable news source, monitoring it with the upmost scrutiny, which arguably would have taken some of the wind out of this headline had it been released today.

Figure 23

Market Reaction to Symantec Hacker Threat to Release Source Code for Norton Antivirus Software



[t0-5sec, t0+30sec] CAV: 3,774%; CAR: 75.93 bps

This chart presents a second-by-second snapshot of the market activity that occurred in Symantec stock the one and a half minutes before and after the news was released at 12:02:47. The last traded price as well as the best bid and ask are plotted on the right axis, alongside the volume traded on the left axis. Note the labels on the volume bars indicate the number of trades in that second.

5.6. Quantile Regression Results

Results of the quantile regression performed on Apple's long and short abnormal returns are reported in Tables 21 and 22, respectively. Abnormal returns appear to be significantly driven by the Alert, Relevance, and Sentiment variables at the one percent level in the majority of quantiles, while Word Count (a proxy for Articles) is largely insignificant. The Traffic indicators are also highly significant, though in some cases their impact on abnormal returns appears quantitatively small.

Consistent with Hypothesis 2, the coefficients for Alerts and Relevance are upward sloping, as shown in Figures 24a and 24b. This indicates that the basis point impact of Relevant Alerts on abnormal returns increases alongside return quantiles. In the case of Alerts, the slope of this relationship is relatively constant across long and short abnormal returns, however Relevance seems to generate a steeper slope for short abnormal returns. Interestingly, both variables are negatively related to lower the abnormal return quantiles, but steadily become significantly positive as the quantiles increase. The only exception to this is the coefficient for short abnormal returns for Alerts, which becomes less negative. The coefficients for Sentiment shown in Figure 24c are curious in that they seem to have virtually no impact on abnormal long returns, but appear positively correlated to short abnormal returns, which runs contrary to expectations.

As shown in Figures 24d and 24e, the coefficients for Event Counts and Reference Event Counts are quite mixed for long and short abnormal returns, making it difficult to draw conclusions that would address Hypothesis 3. The slope of the Event Count coefficients are fairly flat across quantiles, and appear to suggest that in the case of short-selling, less noise is better, while when buying, some noise is actually mildly positive for abnormal returns. On the other hand, the impact of Reference Event Counts, which is more pronounced than that of Event Counts, tells us that more simultaneously released news about the same company tends to exacerbate stock losses, but diminishes gains. It would be interesting to know whether these simultaneously released news items tend to either agree or disagree with the sentiment expressed in the target news item, as that may help explain this result.

The trend in the Reference Event Count for the Before Window suggests that as abnormal returns increase, they become increasingly negatively related to the volume of news about the target stock that got released in the three hours preceding the target news release, as shown in Figure 24g. This implies that a piece of news tends to be more impactful when less news about the same company has been released in the run-up. This is consistent with the expectations in Hypothesis 3, and makes intuitive sense since it also means the news is more likely to be novel in nature.

The traffic variables that deal with timing and proximity to other news items tend to have a small influence on abnormal returns, despite being highly significant. However, the trend in Figure 24h does point to increases in the number of Seconds Since the Previous Event as partly explaining larger abnormal returns, as if to suggest the target news item benefits from having the spotlight to itself. This is in line with expectations. Interestingly though, the Median Seconds Between Events appears to be negatively correlated to return quantiles, which suggests that more market-moving reactions tend be preceded by bursts of news items clustered in shorter, more congested intervals, rather than being more spaced out.

Day-of-the-week trends, while somewhat difficult to interpret in the context of regression quantiles, are significant, and do exhibit an interesting pattern for Apple in that the largest abnormal short returns tend to be skewed towards occurring at the beginning of the week, on either Monday or Tuesday (see Table 22). Long abnormal returns don't appear to exhibit any of these trends, and neither do the time-of-day dummies.

While Apple's results are insufficient to draw any firm conclusions about the relationship between abnormal returns and the news and traffic variables across the broader sample of Nasdaq stocks, they do raise some interesting questions as to whether or not other stocks tend to display similar patterns, and whether those trends are tied to corporate sector, market capitalization, or propensity for news coverage, since it's reasonable to suspect that the explanatory variables could behave differently for smaller, less visible stocks as compared to a large, highly visible company such as Apple. It would certainly be interesting to extrapolate this type of analysis to a more robust sample of Nasdaq stocks in the future and revisit some of these preliminary observations.

Table 21

Quantile Regression Results for AAPL Abnormal Long Returns

Dependent Variable: ABNORMAL LONG RETURNS	
Usable Observations	1,894,786
Degrees of Freedom	1,894,766
Skipped/Missing (from 1894866)	80

	Stock: AAPL	**Quantile 0.1**		**Quantile 0.3**		**Quantile 0.5**		**Quantile 0.7**		**Quantile 0.9**	
		DW Stat	0.70	DW Stat	1.47	DW Stat	1.72	DW Stat	1.43	DW Stat	0.77
	Variable	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error
1	Constant	-4.61 ***	0.02	-1.38 ***	0.01	-3.30 ***	0.01	-2.60 ***	0.01	-2.45 ***	0.02
2	Alert	-1.37 ***	0.23	0.24	0.16	0.53 ***	0.15	1.11 ***	0.16	2.87 ***	0.22
3	Relevance	-0.46 ***	0.16	0.13	0.11	0.09	0.10	0.20 *	0.11	0.68 ***	0.15
4	Sentiment Log Odds	-0.24 ***	0.05	-0.18 ***	0.03	-0.09 ***	0.03	-0.10 ***	0.03	-0.01	0.04
5	Sentiment Word Count	0.00	0.00	0.00 *	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	Traffic Event Count	0.40 ***	0.00	0.48 ***	0.00	0.46 ***	0.00	0.50 ***	0.00	0.49 ***	0.00
7	Traffic Reference Event Count	-2.41 ***	0.03	-1.95 ***	0.02	-1.42 ***	0.02	-1.25 ***	0.02	-0.97 ***	0.03
8	Traffic Seconds Since Previous Event	0.00 ***	0.00	0.00 ***	0.00	0.00 ***	0.00	0.00 ***	0.00	0.01 ***	0.00
9	Traffic Before Event Count	0.00 ***	0.00	0.00 ***	0.00	0.00 ***	0.00	0.00 ***	0.00	0.00 ***	0.00
10	Traffic Before Reference Event Count	0.04 ***	0.00	0.01 ***	0.00	0.00 ***	0.00	-0.01 ***	0.00	-0.03 ***	0.00
11	Traffic Before Median Seconds Between	0.30 ***	0.00	0.24 ***	0.00	0.18 ***	0.00	0.10 ***	0.00	0.00 **	0.00
12	Abnormal Adjusted Volume (t-1)	0.15 ***	0.00	0.15 ***	0.00	0.13 ***	0.00	0.13 ***	0.00	0.15 ***	0.00
13	Monday	1.96 ***	0.01	0.76 ***	0.01	1.35 ***	0.01	2.02 ***	0.01	3.72 ***	0.01
14	Tuesday	-3.31 ***	0.01	-4.48 ***	0.01	-0.02 **	0.01	0.75 ***	0.01	2.31 ***	0.01
15	Thursday	0.26 ***	0.01	-0.98 ***	0.01	3.54 ***	0.01	4.08 ***	0.01	-0.18 ***	0.01
16	Friday	-2.77 ***	0.01	-1.07 ***	0.01	0.11 ***	0.01	0.68 ***	0.01	4.15 ***	0.01
17	Opening Half Hour	-2.58 ***	0.02	-0.13 ***	0.01	2.27 ***	0.01	2.87 ***	0.01	2.84 ***	0.02
18	Closing Half Hour	-0.26 ***	0.02	0.22 ***	0.01	0.49 ***	0.01	0.52 ***	0.01	1.04 ***	0.02
19	Morning	4.99 ***	0.02	1.64 ***	0.01	1.03 ***	0.01	-1.19 ***	0.01	0.16 ***	0.02
20	Midday	2.93 ***	0.01	4.82 ***	0.01	-2.58 ***	0.01	-2.80 ***	0.01	-0.78 ***	0.01

This table reports the regression coefficient estimates and standard errors across five quantiles for AAPL's long abnormal returns. Significance is confirmed using both T-test and P-value where *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 22

Quantile Regression Results for AAPL Abnormal Short Returns

Dependent Variable: ABNORMAL SHORT RETURNS	
Usable Observations	1,894,786
Degrees of Freedom	1,894,766
Skipped/Missing (from 1894866)	80

Stock: AAPL		Quantile 0.1		Quantile 0.3		Quantile 0.5		Quantile 0.7		Quantile 0.9	
		DW Stat	0.57	DW Stat	1.16	DW Stat	1.44	DW Stat	1.18	DW Stat	0.57
	Variable	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error	Coeff	Std Error
1	Constant	6.83 ***	0.04	6.52 ***	0.03	5.27 ***	0.03	2.04 ***	0.03	16.86 ***	0.04
2	Alert	-3.75 ***	0.56	-2.81 ***	0.39	-1.63 ***	0.35	-1.53 ***	0.39	-0.83	0.55
3	Relevance	-1.76 ***	0.38	-0.38	0.26	-0.63 ***	0.24	0.39	0.26	2.33 ***	0.37
4	Sentiment Log Odds	0.18	0.11	0.22 ***	0.08	0.40 ***	0.07	0.49 ***	0.08	0.71 ***	0.11
5	Sentiment Word Count	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	Traffic Event Count	-0.93 ***	0.01	-1.02 ***	0.01	-0.84 ***	0.01	-0.79 ***	0.01	-0.71 ***	0.01
7	Traffic Reference Event Count	1.82 ***	0.07	2.42 ***	0.05	2.55 ***	0.04	3.00 ***	0.05	3.69 ***	0.07
8	Traffic Seconds Since Previous Event	-0.01 ***	0.00	0.00 ***	0.00	0.00	0.00	0.00 ***	0.00	0.01 ***	0.00
9	Traffic Before Event Count	-0.01 ***	0.00	0.00 ***	0.00	0.00	0.00	0.00 ***	0.00	0.01 ***	0.00
10	Traffic Before Reference Event Count	0.09 ***	0.00	0.03 ***	0.00	0.01 ***	0.00	-0.02 ***	0.00	-0.09 ***	0.00
11	Traffic Before Median Seconds Between	0.50 ***	0.01	0.31 ***	0.00	0.17 ***	0.00	0.00	0.00	-0.16 ***	0.01
12	Abnormal Adjusted Volume (t-1)	-1.45 ***	0.01	-1.32 ***	0.00	-1.33 ***	0.00	-1.32 ***	0.00	-1.29 ***	0.01
13	Monday	3.11 ***	0.03	8.62 ***	0.02	3.16 ***	0.02	22.55 ***	0.02	19.75 ***	0.03
14	Tuesday	6.39 ***	0.03	11.56 ***	0.02	4.44 ***	0.02	5.66 ***	0.02	17.66 ***	0.03
15	Thursday	-17.13 ***	0.03	-8.20 ***	0.02	4.25 ***	0.02	0.51 ***	0.02	0.61 ***	0.03
16	Friday	-4.64 ***	0.03	-4.99 ***	0.02	-14.36 ***	0.02	-17.82 ***	0.02	3.93 ***	0.03
17	Opening Half Hour	7.95 ***	0.04	4.98 ***	0.03	0.22 ***	0.03	-6.28 ***	0.03	6.71 ***	0.04
18	Closing Half Hour	0.41 ***	0.04	0.29 ***	0.03	0.35 ***	0.02	0.73 ***	0.03	1.64 ***	0.04
19	Morning	3.73 ***	0.04	3.46 ***	0.03	6.26 ***	0.03	8.45 ***	0.03	-6.63 ***	0.04
20	Midday	-6.01 ***	0.03	-4.71 ***	0.02	-1.44 ***	0.02	0.99 ***	0.02	6.09 ***	0.03

This table reports the regression coefficient estimates and standard errors across five quantiles for AAPL's short abnormal returns. Significance is confirmed using both T-test and P-value where *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Figure 24

Quantile Regression Estimates for Long and Short Abnormal Returns

Fig. 24a

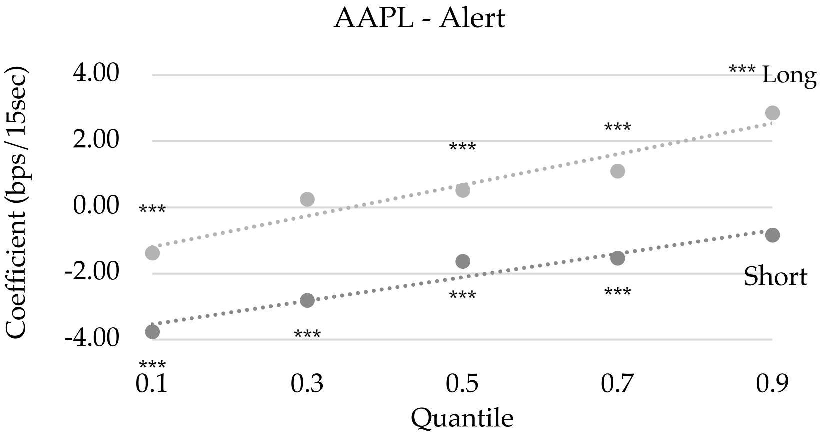


Fig. 24b

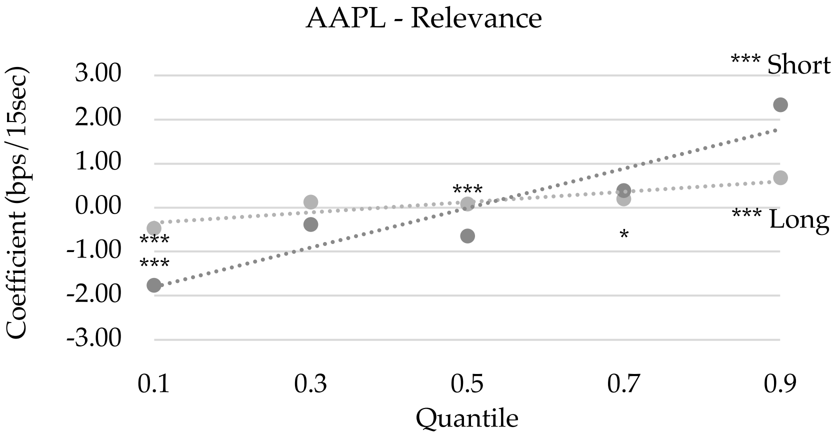
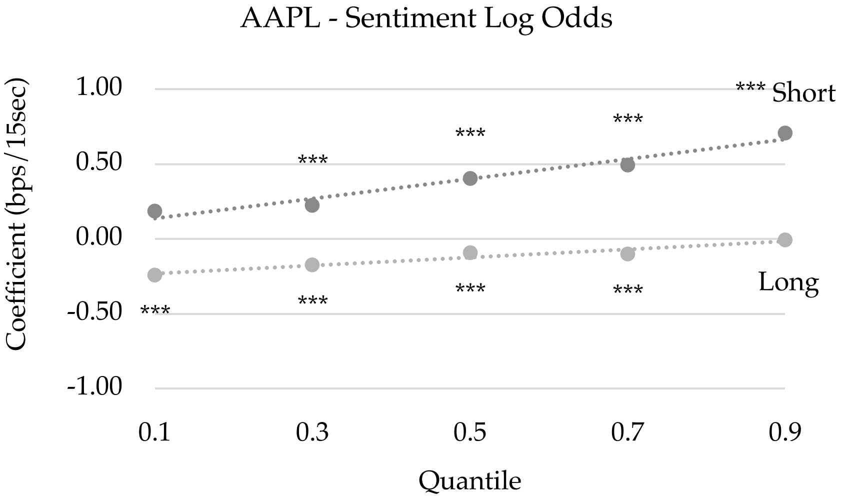


Fig. 24c



These figures show the coefficient estimates for various explanatory variables across five quantiles for both long and short abnormal returns for the individual stock. Note insignificant variables have been excluded. Significance is confirmed using both T-test and P-value where *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Figure 24 – continued

Quantile Regression Estimates for Long and Short Abnormal Returns

Fig. 24d

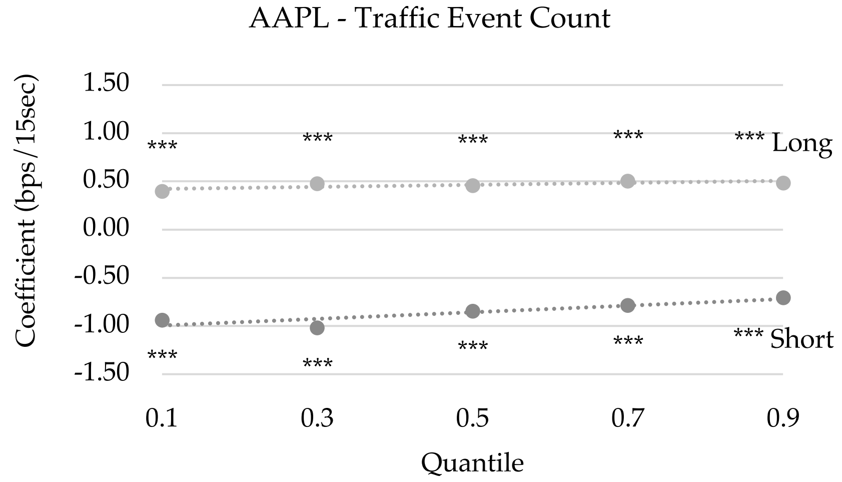


Fig. 24e

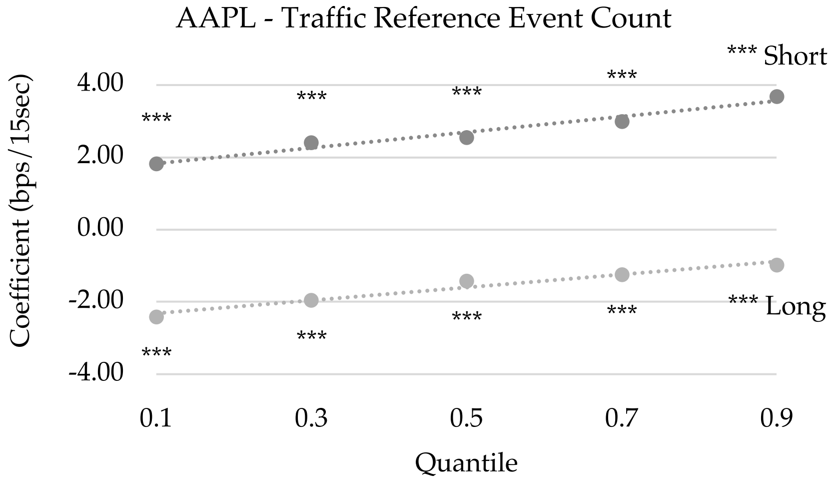


Fig. 24f

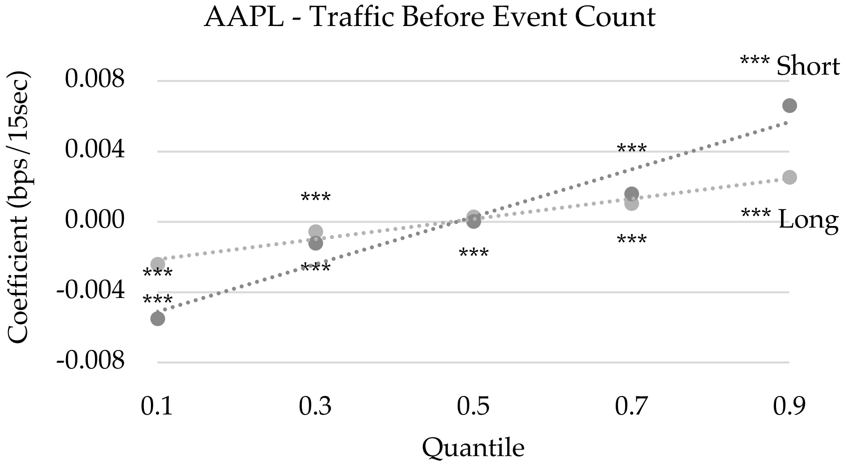
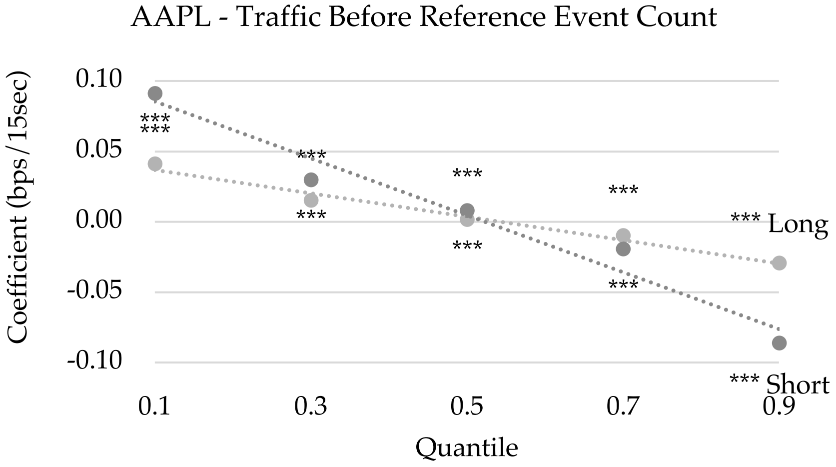


Fig. 24g



These figures show the coefficient estimates for various explanatory variables across five quantiles for both long and short abnormal returns for the individual stock. Note insignificant variables have been excluded. Significance is confirmed using both T-test and P-value where *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Figure 24 – continued

Quantile Regression Estimates for Long and Short Abnormal Returns

Fig. 24h

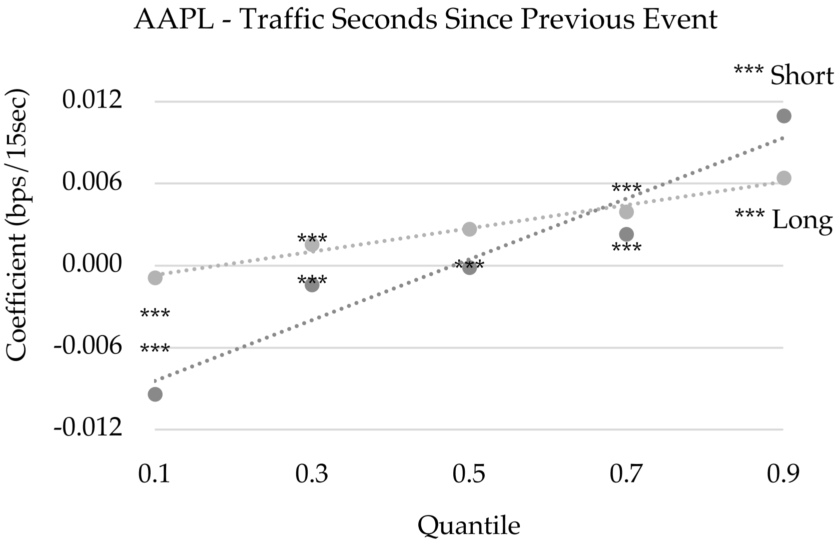
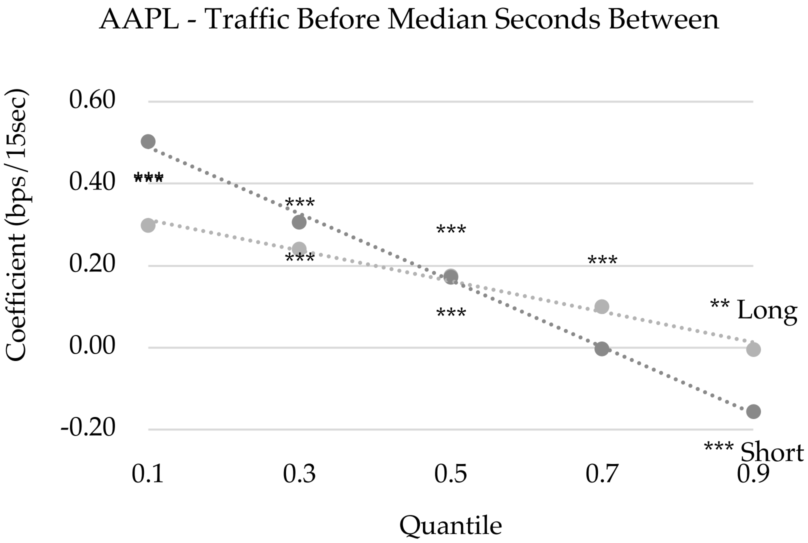


Fig. 24i



These figures show the coefficient estimates for various explanatory variables across five quantiles for both long and short abnormal returns for the individual stock. Note insignificant variables have been excluded. Significance is confirmed using both T-test and P-value where *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

6. Conclusion

According to the results of the CAR and CAV analyses, abnormal trading activity appears most significant and pronounced in the first few minutes around a news release, with volume more sensitive to the TRNA metrics than returns, consistent with the findings of Groß-Klußmann and Hautsch (2011), and Smales (2014b). In the five to ten minutes pre-release, both median abnormal volume and returns are significantly positive at 20 per cent and five basis points respectively, and during the first three seconds post-release, mean abnormal volume spikes over 100 per cent higher while abnormal returns shrink to under two basis points.

Both series of tests point to Alerts as having a bigger impact on market dynamics than Articles on the knee-jerk. On average, the CAV analysis shows that positive news is traded more aggressively than negative news in the first few minutes, but also experiences sharper reversals in abnormal returns in the hours following, consistent with Dzielinski (2012), and Borokova and Mahakena (2015). On the other hand, the median CARs for negative news continue to drift higher, echoing Hong et al. (2000) and Smales (2014c, 2015b) who find negative news to be more informative than positive news. Overall, mean volume appears to fluctuate proportionally with the TRNA metrics, suggesting some stocks may be more prone to algorithmic trading than others, especially on positive news. This premise is consistent with Barber and Odean (2008) and Ferguson et al. (2015), who claim "investors are net buyers of attention-grabbing stocks".

Similar to Groß-Klußmann and Hautsch (2011), above-average trading activity was detected in the hours preceding a news release, with CARs spiking as high as 20bps. While information leaking out from other sources may be to partly blame, the underlying headlines also point to a probable feedback loop being created when reporters write about recent stock moves, a feature that has yet to be addressed in the related research.

The distributions of CARs and CAVs across the Relevance, Novelty, and Sentiment indicators were largely similar, with trading activity increasing as thresholds became stricter. The clearest relationships were observed for Relevance and Novelty, with weaker trends detected in the Sentiment scores, similar to the quantile regression results. However, when these thresholds were combined, high Sentiment scores were clearly a common denominator in generating high median CARs, while Novelty was a key factor in all but negative Alerts. Relevance was important to both Alerts and Articles, but appeared more powerful in the latter, given its highly positive correlation to Alerts. The preliminary quantile regression results also point to news traffic as being highly significant in explaining abnormal returns, though for some variables the impact appears quite small.

While these results are not without their caveats, in many ways they do appear consistent with the proposed hypotheses. However, in terms of their performance in a trading strategy, simulated returns using TRNA signals were negative, with median HPRs hovering between minus two to four basis points. A deep dive into some of these news releases reveals the many complexities involved in algorithmically trading the news, with the power of Novelty proving to be inherently limited, and the crux of the trade dependent on the Sentiment indicator correctly predicting the direction of the news. Although there are legitimate trading opportunities to be had among the CAV and CAR outliers, they appear to be few and far between, and would likely require a wide breadth of single-name stock coverage and a carefully targeted scope of tradable opportunities to be profitable.

References

Allen, D.E., McAleer, M., Singh, A.K., 2013b. An entropy based analysis of the relationship between the Dow Jones Index and the TRNA Sentiment series. Working paper, Retrieved from http://www.qminitiative.org/quantitative- management-initiative.html.

Allen, D. E.; McAleer, M.; Singh, A.K., 2014. Machine News and Volatility: The Dow Jones Industrial Average and the TRNA Sentiment Series. Tinbergen Institute Discussion Paper 14-014/III.

Allen, D. E.; McAleer, M.; Singh, A.K., 2015. Daily Market News Sentiment and Stock Prices. Tinbergen Institute Discussion Paper 15-090/III.

Antweiler, W., Frank, M., 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. Journal of Finance 59, 1259–1294.

Antweiler, W., Frank, M., 2006. Do U.S. stock markets typically overreact to corporate news stories? Working paper, University of British Columbia.

Barber, B. M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

Boehmer, E., Jones, C. M., Zhang, X., 2008. Which Shorts Are Informed? The Journal of Finance 62(2).

Boehmer, E, Huszar, Z.R., Wang, Y., Zhang, X., 2016. Are Shorts Equally Informed? A Global Perspective. Working paper.

Bollen, J., Mao, H., Pepe, A., 2011a. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. Paper read at Fifth International AAAI Conference on Weblogs and Social Media.

Bollen J, Mao, H., Zeng, X.J., 2011b. Twitter mood predicts the stock market. Journal of Computational Science 2(1), 1–8.

Borovkova, S., Mahakena, D., 2015. News, volatility and jumps: the case of natural gas futures. Quantitative Finance 15(7), 1217–1242.

Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2012. Which news moves stock prices? A textual analysis. NBER Working Paper.

Brav, A., Heaton, J.B., 2002. Competing theories of financial anomalies. Review of Financial Studies 15, 575–606.

Cahan, R., Jussa, J., Luo, Y., 2009. Breaking news: How to use news sentiment to pick stocks. Macquarie US Equity Research.

Cao, S., Jiang, W., Yang, B., Zhang, A., 2020. How to Talk When a Machine is Listening: Corporate Disclosure in the Age of AI. NBER Working Paper Series

Chan, W.S., 2003. Stock Price reaction to news and no-news: drift and reversal after headlines. Journal of Financial Economics 70, 223–260.

Chen, J., 2004. Information Theory and Market Behavior. Retrieved from https://ssrn.com/abstract=622901 or http://dx.doi.org/10.2139/ssrn.622901

Clark, P.K., 1973. A subordinate stochastic process model with finite variance for speculative prices. Econometrica 41, 135–155.

Dacorogna, M. M., Gencay, R., Muller, U., Olsen, R. B., Pictet, O. V., 2001. An Introduction to High-Frequency Finance, Academic Press, London.

Daniel, K.D., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and over-reactions. Journal of Finance 53, 839–886.

Sanjiv, D.R., Chen, M.Y., 2007. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. Management Science 53(9), 1375–1388. Retrieved from https://doi.org/10.1287/mnsc.1070.0704

DellaVigna, S., Pollet, J.M., 2009. Investor inattention and Friday earnings announcements. Journal of Finance 64, 709–749.

Demers, E.A., Vega, C., 2010. Soft information in earnings announcements: news or noise? INSEAD & Board of Governors of the Federal Reserve System, Washington, DC.

Dzielinski, M., 2012. Which news resolves asymmetric information? Available at SSRN 2102512.

Engelberg, J., 2008. Costly information processing: evidence from earnings announcements. In AFA 2009 San Francisco Meetings.

Engelberg J., Reed, A.V., Ringgenberg, M.C., 2012. How are shorts informed? Short sellers, news, and information processing. Journal of Financial Economics 105(2), 260–278.

Feldman, R., Govindaraj, S., Livnat, J., Segal, B., 2010. Management's Tone Change, Post Earnings Announcement Drift and Accruals. Review of Accounting Studies 15 (4), 915–953.

Ferguson, N. J., Philip, D., Lam, H.Y.T., Guo, M., 2015. Media content and stock returns : the predictive power of press. Multinational Finance Journal 19(1), 1–31.

Field, L.C., Hanka, G., 2001. The expiration of share lockups. Journal of Finance 2, 471–500.

Fisher, I.E., Garnsey, M., Hughes, M.E., 2016. Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research. Intelligent Systems in Accounting, Finance and Management. 10.1002/isaf.1386.

Foucault, T., Hombert, J., Roşu, I., 2016. News Trading and Speed. Journal of Finance 71(1), 335–382.

Foster, D.F., Viswanathan, S., 1996. Strategic Trading When Agents Forecast the Forecasts of Others. Journal of Finance 51, 1437–1478.

Garcia, D., 2013. Sentiment during recessions. Journal of Finance 68(3).

Groß-Klußmann, A., Nikolaus H., 2011. When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. Journal of Empirical Finance 18, 321–340.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Hafez, P., Xie, J., 2012a. Ravenpack sentiment and its relationship with macro-economic indicators. RavenPack International SL.

Hakim, T., Lypny, G., Bhabra, H.S., 2012. IPO lockup expiration in the Middle East and North Africa. Journal of Multinational Financial Management 22, 252–262.

Handbook of Journalism – Other Common Story Forms. Retrieved from http://handbook.reuters.com/index.php?title=Other_Common_Story_Forms

Handbook of Journalism – The Drill for Breaking News. Retrieved from http://handbook.reuters.com/index.php?title=The_Drill_for_Breaking_News

Hansen, P., Lunde, A., 2006. Realized variance and microstructure noise. Journal of Business Economic Statistics 24, 127–161.

Healy, A.D., Lo, A.W., 2011. Managing real-time risks and returns: The Thomson Reuters NewsScope event indicies. Handbook of News Analytics in Finance, John Wiley & Sons, 73–108.

Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Hong, H., Lim, T., Stein, J.C., 2000. Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. Journal of Finance 55(1), 265–295.

Huang, A.H., Zang, A.Y., Zheng, R., 2014. Evidence on the Information Content of Text in Analyst Reports. The Accounting Review 89(6), 2151–2180.

Huynh, T.D., Smith, D.R., 2013. News sentiment and momentum. FIRN Research Paper.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47(2), 263. ISSN 0012-9682. doi:10.2307/1914185.

Kahneman, D., Frederick, S., 2001. Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In Thomas Gilovich; Dale Griffin; Daniel Kahneman. Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge: Cambridge University Press, 49–81. ISBN 978-0-521-79679-8. OCLC 47364085.

Kalev, P.S., Liu, W.S., Pham, P.K., Jarnecic, E., 2004. Public information arrival and volatility of intraday stock returns. Journal of Banking and Finance 28, 1441–1467.

Kandel, E., Pearson, N., 1995. Differential Interpretation of Public Signals and Trade in Speculative Markets. Journal of Political Economy 103, 831–872.

Kearney, C., Liu, S., 2014. Textual Sentiment in Finance: A Survey of Methods and Models. International Review of Financial Analysis 33, 171–185.

Koenker, R., Bassett, G., 1978. Regression Quantiles. Econometrica, 46:1, 33–50.

Koenker, R., Hallock, F., 2001. Quantile Regression. Journal of Economic Perspectives, 15(4), 143–156.

Kothari, S.P., Li, X., Short, J.E., 2009. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. The Accounting Review 84(5), 1639–1670.

Kyle, A.S., 1985. Continuous Auctions and Insider Trading. Econometrica 53, 1315–1335.

Leinweber, D., Sisk, J., 2011. Event driven trading and the 'new news'. Journal of Portfolio Management 38, 110–124.

Li, F., 2010. The determinants and information content of forward-looking statements in corporate filings— A Naive Bayesian machine learning algorithm approach. Journal of Accounting Research 48, 1049–1102.

Liddy, E.D., 2001. Natural Language Processing. In Encyclopedia of Library and Information Science 2nd Edition, NY. Marcel Decker, Inc.

Loughran, T., McDonald, B., 2011a. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance 66(1), 35–65.

NASDAQ OMX, Global Data Products Combined Holidays. 2013. Retrieved from www.nasdaqtrader.com/content/technicalsupport/dataproducts/GDPcombinedholidays.xls

Nasukawa, T., Nagano, T., 2001. Text analysis and knowledge mining system. IBM Systems Journal 40, 967–984.

Natural Language Processing Market by Type (Rule-Based, Statistical, and Hybrid), Technologies (Recognition, IVR, OCR, Speech Recognition, Text Processing, Pattern & Image Recognition), by Deployment Type, Vertical & by Region - Global Forecast to 2021. Retrieved from http://www.marketsandmarkets.com/Market-Reports/natural-language-processing-nlp-825.html

Natural Language Processing Market - Growth, Trends, and Forecast (2020 - 2025). 2020, March. Retrieved from https://www.reportlinker.com/p05877815/Natural-Language-Processing-Market-Growth-Trends-and-Forecast.html?utm_source=PRN

Nemschoff, M., 2014, June 28. A Quick Guide to Structured and Unstructured Data. Retrieved from http://www.smartdatacollective.com/quick-guide-structured-and-unstructured-data/

Petersen, M.A., 2004. Information: Hard and Soft. Working Paper.

Rise of the news-reading machines. 2010, January 26. Retrieved from https://ftalphaville.ft.com/2010/01/26/134561/rise-of-the-news-reading-machines/

Sedlak, M., 2016, June 22. How Natural Language Processing is transforming the financial industry. Retrieved from https://www.ibm.com/blogs/watson/2016/06/natural-language-processing-transforming-financial-industry-2/

Shi, Y., Liu, W., Ho, K., 2016. Public news arrival and the idiosyncratic volatility puzzle. Journal of Empirical Finance 37, 159–172.

Shiller, R., 2005. Irrational Exuberance. Princeton NJ: Princeton University Press, 2nd Edition.

Simon, H.A., 1955. A behavioral model of rational choice. The Quarterly Journal of Economics 69(1), 99–118.

Simon, H.A., 1957. Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting. John Wiley and Sons, New York.

Sinha, N.R., 2016. Underreaction to news in the US stock market. Quarterly Journal of Finance 6(2), 1–46.

Smales, L.A., 2013. Impact of macroeconomic announcements on interest rate futures: high-frequency evidence from Australia. Journal Financial Research 36, 371–388.

Smales, L.A., 2014a. News sentiment and the investor fear gauge. Finance Research Letters 11(2), 122–130.

Smales, L.A., 2014b. Non-scheduled news arrival and high-frequency stock market dynamics: evidence from the Australian stock exchange. Research in International Business and Finance 32, 122–138.

Smales, L.A., 2014c. News Sentiment in the Gold Futures Market. Journal of Banking and Finance 49, 275–286.

Smales, L.A., 2015a. Time-variation in the impact of news sentiment. Asymmetric Volatility Response to news sentiment in gold futures. International Review of Financial Analysis 37, 40–50.

Smales, L.A., 2015b. Asymmetric Volatility Response to news sentiment in gold futures. Journal of International Financial Markets, Institutions and Money 34, 161–172.

Solomon, D., 2012. Selective Publicity and Stock Prices. Journal of Finance 67, 599– 638.

Storkenmaier, A., Wagener, M., Weinhardt, C., 2012. Public information in fragmented markets. Financial Markets and Portfolio Management 26(2), 179–215.

Tauchen, G.E., Pitts, M., 1983. The Price Variability-Volume Relationship on Speculative Markets. Econometrica 51, 485–505.

Tetlock, P.C., 2007. Giving content to investor sentiment: the role of media in the stock market. The Journal of Finance 62(3), 1139–1168.

Tetlock, P.C., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: quantifying language to measure firms' fundamentals. The Journal of Finance 63, 1437–1467.

Tetlock, P.C., 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23(9), 3520–3557.

Tetlock, P.C., 2011. All the news that's fit to reprint: do investors react to stale information? Review of Financial Studies 24(5), 1481–1512.

Tumarkin, R., Whitelaw, R.F., 2001. News or noise? Internet postings and stock prices. Financial Analysts Journal 57(3), 41-51.

Vega, C., 2006. Stock price reaction to public and private information. Journal of Financial Economics 82, 103–133.

Vu, T.T., Chang, S., Ha, Q.T., Collier, N., 2012. An experiment in integrating sentiment features for tech stock prediction in Twitter. Paper read at Workshop on Information Extraction and Entity Analytics on Social Media Data, Mumbai, India.

Wasserfallen, W., Zimmermann, H., 1985. The behavior of intraday exchange rates. Journal of Banking and Finance 9, 55–72.

Varian, H., 1985. Divergence of Opinion in Complete Markets: A Note. The Journal of Finance, 40(1), 309-317. doi:10.2307/2328062

von Beschwitz, B., Keim, D.B., Massa, M., 2015. First to "Read" the News: News Analytics and High Frequency Trading. Paris December 2015 Finance Meeting EUROFIDAI – AFFI.

Yu, Y., Duan, W., Cao, Q., 2013. The impact of social and conventional media on firm equity value: a sentiment analysis approach. Decision Support Systems 55(4), 919–926.

Zhang, S.S., 2017. Need for speed: Hard information processing in a high-frequency world. Journal of Futures Markets 10.1002/fut.21861.

Appendix A

The need for speed and a brief history of Thomson Reuters

The race to be the fastest can be traced back as far as the 1850s, when Paul Julius Reuter ran a financial information service using a fleet of 45 homing pigeons to carry news and stock prices between Brussels and Aachen [10,11]. In 1851, he moved to London to establish the Reuters News Agency – known, as of 2008, as Thomson Reuters – and over time built up an impressive network of telegraph cables and mail steamers that landed him his first major scoop in 1865: breaking the news of President Lincoln's assassination to Europe [12].

By 1866, Reuters had secured strategic access to the first successfully laid transatlantic telegraph cable [13], the likes of which still serve as critical information arteries today. An estimated 428 fiberoptic cables are currently in operation (fiberoptic became the industry standard circa 1988, when the TAT-8 was lit), transmitting data at rates approaching 160Tbps [14,15].

In the 1900s, the invention of the first computers allowed us to entertain the notion of electronic markets. Reuters helped pave the way in transmitting market quotes digitally, launching their Monitor terminal in 1973, and a video terminal in 1981, well before the Bloomberg Terminal swept the market in 1982 [16,17]. It wasn't until the 1990's, however, that the advent of algorithmic trading really took hold, and speed competition reached imperceptible levels, with nanoseconds making all the difference.

As the value of data increased in conjunction with how fast it could be transmitted, financial news vendors began charging premiums for real-time, low latency access to their quantitative and qualitative news data feeds. The realm of quantitative data is straightforward by nature – in the sense that algorithms have no trouble trading on clearly defined numerical cues – however, its qualitative counterpart still relies heavily on human interpretation. Enter: natural language processors (NLPs).

---

[10] Cutler, D., *Company History – Highlights from across Thomson Reuters*, https://www.thomsonreuters.com/en/about-us/company-history.html

[11] Entwisle, J., *Paul Julius Reuter's startup: From Brussels to Aix*, 2016, https://blogs.thomsonreuters.com/answerson/reuters-startup-brussels-aix/

[12] The Editors of Encyclopaedia Britannica, *Thomson Reuters – Canadian Company*, 2018, https://www.britannica.com/topic/Reuters

[13] Hills, J. *The Struggle for Control of Global Communication: THE FORMATIVE CENTURY*, 2002, pp. 55-56, shorturl.at/acmMN

[14] Burns, B, *Submarine Cable History*, 2012, https://www.submarinecablesystems.com/history

[15] TeleGeography, *Submarine Cable 101*, shorturl.at/mBT69

[16] *Thomson Reuters FX Trading*, 2016, https://www.refinitiv.com/content/dam/marketing/en_us/documents/fact-sheets/fx-trading-fact-sheet.pdf

[17] McCracken, H., *How the Bloomberg Terminal Made History – And Stays Ever Relevant*, 2015, https://www.fastcompany.com/3051883/the-bloomberg-terminal

Appendix B

Summary of the Thomson Reuters Handbook of Journalism Classifications for News

Category	Description	Characteristics	Story Type Topic Code	Notes
ALERT aka “snap” aka “bulletin”	- Highest priority, potentially market-moving news - Only presents the essential facts - Usually no more than a series of 5 alerts for a given story - About 100 characters in length	- Written entirely in upper case (except for lower case letters in RICs)		- ALERTS aren’t always tagged at the company level, which is problematic since they are typically the first type of news on a particular story to hit the market
NEWSBREAK aka "Urgent" aka “Rush” aka “Cover”	- Short breaking news story that puts facts into context - Usually follows an Alert(s) within 10 minutes	- If no BRIEF exists, will have same USN as Alert - No other distinguishing factors		- Hard to identify - If a BRIEF exists, the Newsbreak won’t have the same USN as the Alert - If a Newsbreak is updated, USN will change and headline will contain UPDATE 1 (new material included at the bottom)
UPDATE	- Story that adds content, analysis, details, quotes, background to an earlier story - Typically follows an “Urgent”/Newsbreak but is a different story	- Headline tag (begins with): UPDATE #, RPT-UPDATE #, CORRECTED-UPDATE #, CORRECTED-OFFICIAL-UPDATE #, REFILE-UPDATE #, INTERVIEW-UPDATE #, EXCLUSIVE-UPDATE #, RPT-EXCLUSIVE-UPDATE #- - Subsequent UPDATEs are numbered (UPDATE 1, 2, etc.) and have the same USN through the rest of the 24-hour news cycle	- No code according to RTRS guidelines	- UPDATE 1 to that newsbreak takes a new USN and all subsequent updates retain that USN through the rest of the 24-hour news cycle - When a story begins without a snap or a newsbreak, the USN remains the same on all subsequent updates through the rest of the 24-hour news cycle. This applies to all types of story.
CORRECTED	- Substantive, factual error appears in a story or table - Could alter the meaning or significance of the story or undermine its credibility. Ex: wrong RIC, incorrect numbers or proper names	- Headline tag (begins with): CORRECTED-, CORRECTED-(OFFICIAL), CORRECTED-OFFICIAL		- Will overwrite erroneous Alerts/Stories (they are supposed to disappear from the data feed), but appear to use new timestamps - CORRECTED alert will have a new USN in some cases - CORRECTED story will have the same USN
INTERVIEW	- Conveys Reuters exclusivity - Not usually updated	- Headline tag (begins with): INTERVIEW-, RPT-INTERVIEW-, UPDATE #-INTERVIEW-, CORRECTED-INTERVIEW-, INTERVIEW-UPDATE #, PDAC-INTERVIEW-, REFILE-INTERVIEW-, CORRECTED-(OFFICIAL)-(June 22)-RPT-UPDATE #-INTERVIEW, CORRECTED-RPT-UPDATE #-INTERVIEW-	- INTER	
EXCLUSIVE (INTERVIEW)	- Outstanding interview of exceptional significance with a major figure	- Headline tag (begins with): EXCLUSIVE-, EXCLUSIVE-UPDATE #, UPDATE 4-EXCLUSIVE-, CORRECTED-UPDATE #-EXCLUSIVE-, RPT-EXCLUSIVE-UPDATE #-	- EXCLSV	
WITNESS	- Eyewitness accounts on significant news events	- Headline tag (ends with): -REUTERS WITNESS, -RTRS WITNESS, - REUTERS EYEWITNESS, - witness, -witness	- No code according to RTRS guidelines	- Not tagged at the beginning as handbook suggests
REUTERS SUMMIT	- Stories generated by Reuters Summits on a specific topic	- Headline tag (begins with): RPT-Reuters Summit-PREVIEW-, RPT-REUTERS SUMMIT-, REUTERS SUMMIT-, Reuters Summit-, Reuters Insider - Reuters Summit:, Reuters Summit-UPDATE #-, CORRECTED-Reuters Summit-, REFILE-Reuters Summit-, BRIEF-Reuters Summit-, UPDATE #-Reuters Summit-, CORRECTED-(OFFICIAL)-Reuters Summit-	- RSUM	
DEALTALK	- News about IPOs, M&As, SEOs, and ongoing transactions	- Should be used in place of the old and rarely used IPO VIEW headline tag - Headline tag (begins with): IPO VIEW-, RPT-IPO VIEW-, UPDATE #-IPO VIEW-, CORRECTED-IPO VIEW-, DEALTALK-, CORRECTED-DEALTALK-, RPT-DEALTALK-, UPDATE #-DEALTALK-, REFILE-DEALTALK-	- DLTK	
MOVES	- Short pieces about key people moving between banks and brokerages	- Headline tag (begins with): MOVES-, UPDATE #-MOVES-, REFILE-MOVES-, CORRECTED-MOVES-, RPT-MOVES-	- INVB - BISV - BACT	- Could be relevant on a company-specific level, though not sure how important these people are - Not sure at what point a person is important enough to make it to an ALERT as well
ADVISORY	- A notice to subscribers about specific stories	- Headline tag (begins with): ADVISORY-	- ADVS	
WITHDRAWAL	- When a story is killed - Reserved for stories that are totally wrong or so fundamentally flawed that a conventional correction isn’t possible	- Headline tag (begins with): DELETE [only this word], Delete [only this word], ADVISORY- [+ ignore or withdraw*]		

Appendix B – continued

Summary of the Thomson Reuters Handbook of Journalism Classifications for News

Category	Description	Characteristics	Story Type Topic Code	Notes
BRIEF	- Repeats alerts verbatim (can have >5 bullets) - Allows clients who can't receive Alerts to get headlines	- Same USN as Alert - Usually says BRIEF in headline (begins with): BRIEF-, BRIEF -, RPT-BRIEF-, CORRECTED-BRIEF-, REFILE-BRIEF-		- A BRIEF may have a company tag while and earlier ALERT does not (may need to check USN against document level scores)
REPEAT	- Repeated exactly as it first appeared (no corrections or changes) - Repeated when it needs to be distributed more widely to extra product codes	- Headline tag (begins with): REPEAT:, PRESS RELEASE - REPEAT-BMO STUDY:, REPEAT-, RPT-		- Sometimes the repeat will be the first iteration in our data series –I'm not sure why - Seems like USN changes once a story is repeated
HIGHLIGHTS	- Highlights from a live event - Bullet form	- Headline tag (begins with): HIGHLIGHTS-, POLL-HIGHLIGHTS-, UPDATE #-INVESTOR HIGHLIGHTS-	- HLGT	- Contains multiple bullet points, not necessarily all related to the company in question, therefore the sentiment scores may be, to some extent, contaminated by other bullet points - Earlier versions get overwritten, I'm not sure if that's true in our dataset as well
ANALYSIS	- In-depth look at an issue - Adds content that explains the significance of the news - Looks at possible future developments - Not used to break news - Typically not updated	- Headline tag (begins with): ANALYSIS -, RPT-ANALYSIS-, REFILE-ANALYSIS-, IFR-ANALYSIS-, CORRECTED-ANALYSIS-, LMEWEEK-ANALYSIS-	- ANV	- Doesn't really break any news - Adds content and colour, but is the USN even traceable?
SPECIAL REPORT/ INSIGHT/ SPECIAL REPORTING	- Investigative stories that are in depth and revelatory on a current news topic or theme - Done on initiative and may not have otherwise come to light - Special reports are magazine length (2000-4000 words) - Insights are shorter (max 1500 words)	- Headline tag (begins with): INSIGHT-, RPT-INSIGHT-, REFILE-INSIGHT-, UPDATE #-INSIGHT-, CORRECTED-INSIGHT-, UPDATE #-INSIGHT:, SPECIAL REPORT-, RPT-SPECIAL REPORT-, REFILE-SPECIAL REPORT-, CORRECTED-SPECIAL REPORT-, UPDATE #-SPECIAL REPORT:, RPT-SPECIAL REPORT:	- SREP - EREP	- Subjective reporting – not breaking news that other wires will necessarily cover, but could still sway sentiment.
SNAPSHOT	- Short format summary for Eikon users only - not meant for media clients - Compiled manually – unlike a BRIEF	- Headline tag (begins with): US STOCKS SNAPSHOT-, SNAPSHOT-	- No code according to RTRS guidelines	- Can be used to cover an ALERT, or stand alone
NEWSMAKER	- Stories profiling an individual in the news - Subjects may be leading figures in business, politics, sports, arts and entertainment, etc.	- Headline tag (begins with): NEWSMAKER-, CORRECTED-NEWSMAKER-, REFILE-NEWSMAKER-, UPDATE #-NEWSMAKER-, RPT-NEWSMAKER-, hold-NEWSMAKER-	- NMKR	- Likely background news
OBITUARY	- Stories profiling a newsworthy person that has just died - Subjects may be leading figures in business, politics, sports, arts and entertainment, etc.	- Headline tag (begins with): OBITUARY-	- OBIT	- Likely background news
RESEARCH ALERT	- Highlights broker research on particular companies or sectors.	- Headline tag (begins with): BRIEF-RESEARCH ALERT-	- RCH	
WRAPUP	- Combines elements of previously separate stories into one trunk or main story	- Headline tag (begins with): WRAPUP #, CORRECTED-WRAPUP # - WRAPUPs are numbered in sequence (WRAPUP 1, 2, etc.) and have the same USN	- No code according to RTRS guidelines	- Delayed - Can't match USN to original story
RESEARCH ROUNDUP	- Summarizes earlier research	- Headline tag (begins with): COPPER ROUNDUP-, RPT- COPPER ROUNDUP-, REFILE- COPPER ROUNDUP-		
PREVIEW	- Curtain-raiser, usually filed 2-3 days before an event (not routine)	- Headline tag (begins with): PREVIEW-, RPT-PREVIEW-, RPT-Reuters Summit-PREVIEW-, CORRECTED-PREVIEW-, PREVIEW:, PREVIEW :, AUTOSHOW-PREVIEW-	- PRE	
WEEKAHEAD	-Outlines events for the upcoming week (similar to DIARY)	- Headline tag (begins with): WEEKAHEAD-, MIDEAST WEEKAHEAD-, MIDEAST MARKETS WEEKAHEAD-, Reuters Insider - Europe Weekahead:, UPDATE #-WEEKAHEAD-, RPT-WEEKAHEAD-	- No code according to RTRS guidelines	
POLL	- Exclusively run Reuters poll - Establishes a consensus view ahead of events	- Headline tag (begins with): RPT-POLL-, CORRECTED-EARNINGS POLL-, RPT-EARNINGS POLL-, REFILE-EARNINGS POLL-, UPDATE #-EARNINGS POLL-, RPT-TABLE-POLL-, EARNINGS POLL-	- RPOLL	

Appendix B – continued

Summary of the Thomson Reuters Handbook of Journalism Classifications for News

Category	Description	Characteristics	Story Type Topic Code	Notes
REFILE	- Corrects minor errors without unnecessarily alarming readers - Error has no bearing on investment decision or understanding of the news. Ex: typos, dropped or duplicated words	- Headline tag (begins with): REFILE-		
INSTANT/ ANALYST VIEW	- Provides bullet point reactions from analysts and major decision makers to breaking news - Usually appears within 15 minutes of the release, and wraps up within 30 minutes	- Headline tag (begins with): INSTANT VIEW #, UPDATE #-INSTANT VIEW –, CORRECTED-INSTANT VIEW #, INSTANT VIEW-NASDAQ, RPT-INSTANT VIEW #-, ANALYST VIEW-, ANALYST VIEW #-, CORRECTED-ANALYST VIEW-	- INVU	- Seems to have a different USN from the original story - More of an opinion piece than news
FEATURE	- Story aimed at enlightening the reader on a trend, issues or personality - Must be accompanied by illustrative material - Not time sensitive	- Headline tag (begins with): FEATURE-, REFILE-FEATURE-, CORRECTED-FEATURE-, RPT-FEATURE-	- FEA	- Not sure to what extent these pieces sway market sentiment
COLUMN	- Opinion piece	- Headline tag (begins with): COLUMN-, REFILE-COLUMN-, RPT-COLUMN-, CORRECTED-RPT-COLUMN-, CORRECTED-COLUMN-	- CLM	- Headline tag: COLUMN- unless there is a product name – I don't know what this means or looks like
DIARY	- List of upcoming economic events	- Headline tag (begins with): STOCK NEWS US- & diary or diaries	- DIARY	- Watch out for "subsidiary" - Harder to filter out
FACTBOX	- Can use Q&A format to explain complex issue - Can use scenario format to speculate on possible outcomes for a situation	- Headline tag (begins with): FACTBOX-, UPDATE-FACTBOX-, RPT-FACTBOX-, REFILE-FACTBOX-	- FBOX	
TIMELINE/ CHRONOLOGY	- Timeline of key events pertaining to an issue	- Headline tag (begins with): TIMELINE-, TIMELINE-Chronology, UPDATE # TIMELINE-, TIMELINE-UPDATE #-	- TMLN	
BRIGHTS/ ODDS	- Funny, quirky, bizarre story - Wouldn't be covered were it not so unusual	- No distinguishing features	- ODD	
TABLE	- Table presented after an alert	- Headline tag (begins with): TABLE-, RPT-TABLE-POLL-, RPT-TABLE-, CORRECTED-TABLE-, UPDATE #-TABLE-, REFILE-TABLE-, CORRECTED - TABLE-, CORRECTED-OFFICIAL- TABLE-, RPT- TABLE-	- No code according to RTRS guidelines	
TEXT/ TRANSCRIPT	- The entire original text or verbatim transcript of a statement or speech. - Reported after the story	- Headline tag (begins with): TEXT-, TEXT:, OFFICIAL-TEXT-, RPT-TEXT-, CORRECTED-TEXT-, REFILE-TEXT-	- TXT	
WIDER IMAGE	- Photo essay	- Headline tag: WIDER IMAGE		
TOP NEWS SUMMARIES	- Highlights main stories and headlines, typically no more than 12	- Headline tag: *TOP NEWS*	- TOP - GLANCE - XREF	- This repeats old news to help subscribers catch up
PRESS DIGEST	- Summaries of news reported by other significant newswires	- Headline tag (begins with): PRESS DIGEST –, PRESS DIGEST-, PRESS DIGEST-Australian Business News –, PRESS DIGEST - New York Times business news –, PRESS DIGEST - Indian Business News -, PRESS DIGEST - Financial Times –, PRESS DIGEST - MALAYSIA –, PRESS DIGEST-New Zealand newspapers –	- PRESS	
BREAKINGVIEWS	- Commentary on current events by Reuters columnists	- Headline tag (begins with): BREAKINGVIEWS-, CORRECTED-BREAKINGVIEWS-, RPT-BREAKINGVIEWS-, Reuters Insider - Breakingviews:	- BRV	
DAY AHEAD	- Market preview for the following day	- Headline tag (begins with): Reuters Insider - Europe Day Ahead:, Reuters Insider - U.S. Day Ahead:	- No code according to RTRS guidelines	
REVIEW	- A critique of a consumer product or service including books, films, restaurants, shows	- Couldn't find any observations	- REV	
TAKE A LOOK	- At-a-glance index of stories on a particular news issue.	- Headline tag (begins with): TAKE A LOOK-, REFILE-TAKE A LOOK-, AIRSHOW-TAKE A LOOK-, RPT-TAKE A LOOK-	- TAL	
TECHNICALS	- Predictive statistical analysis of markets based on historical price movements.	- Headline tag (begins with): TECHNICALS-	- INSI	
TRAVEL POSTCARD	- Tips for short trips to cities and countries from Reuters correspondents.	- Couldn't find any observations	- PCARD	

Appendix C

Nasdaq Constituents between 2011-2015

Month/Stock	AAL	AAPL	ADBE	ADI	ADP	ADSK	AKAM	ALTR	ALXN	AMAT	AMGN	AMZN	APOL	ATVI	AVGO	BBBY	BIDU	BIIB	BMC
2010-12-31		X	X		X	X	X	X		X	X	X	X	X		X	X	X	X
2011-01-31		X	X		X	X	X	X		X	X	X	X	X		X	X	X	X
2011-02-28		X	X		X	X	X	X		X	X	X	X	X		X	X	X	X
2011-03-31		X	X		X	X	X	X		X	X	X	X	X		X	X	X	X
2011-04-30		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-05-31		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-06-30		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-07-31		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-08-31		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-09-30		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-10-31		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-11-30		X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X
2011-12-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-01-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-02-29		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-03-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-04-30		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-05-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-06-30		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-07-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-08-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-09-30		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-10-31		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-11-30		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2012-12-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-01-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-02-28		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-03-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-04-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-05-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-06-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X
2013-07-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2013-08-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2013-09-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2013-10-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2013-11-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2013-12-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-01-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-02-28		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-03-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-04-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-05-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-06-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-07-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-08-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-09-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-10-31		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-11-30		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2014-12-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-01-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-02-28	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-03-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-04-30	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-05-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-06-30	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-07-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-08-31	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-09-30	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
2015-10-31	X	X	X	X	X	X	X		X	X	X	X		X	X	X	X	X	
2015-11-30	X	X	X	X	X	X	X		X	X	X	X		X	X	X	X	X	
2015-12-31	X	X	X	X	X	X	X		X	X	X	X		X	X	X	X	X	

Month/Stock	BMRN	BRCM	CA	CELG	CEPH	CERN	CHKP	CHRW	CHTR	CMCSA	COST	CSCO	CTAS	CTRP	CTRX	CTSH	CTXS	DELL	DISCA
2010-12-31		X	X	X	X	X	X	X		X	X	X	X	X		X	X	X	
2011-01-31		X	X	X	X	X	X	X		X	X	X	X	X		X	X	X	
2011-02-28		X	X	X	X	X	X	X		X	X	X	X	X		X	X	X	
2011-03-31		X	X	X	X	X	X	X		X	X	X	X	X		X	X	X	
2011-04-30		X	X	X	X	X	X	X		X	X	X		X		X	X	X	
2011-05-31		X	X	X	X	X	X	X		X	X	X		X		X	X	X	
2011-06-30		X	X	X	X	X	X	X		X	X	X		X		X	X	X	
2011-07-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2011-08-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2011-09-30		X	X	X		X	X	X		X	X	X		X		X	X	X	
2011-10-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2011-11-30		X	X	X		X	X	X		X	X	X		X		X	X	X	
2011-12-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-01-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-02-29		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-03-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-04-30		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-05-31		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-06-30		X	X	X		X	X	X		X	X	X		X		X	X	X	
2012-07-31		X	X	X		X	X	X		X	X	X				X	X	X	
2012-08-31		X	X	X		X	X	X		X	X	X				X	X	X	
2012-09-30		X	X	X		X	X	X		X	X	X				X	X	X	
2012-10-31		X	X	X		X	X	X		X	X	X				X	X	X	
2012-11-30		X	X	X		X	X	X		X	X	X				X	X	X	
2012-12-31		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-01-31		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-02-28		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-03-31		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-04-30		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-05-31		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-06-30		X	X	X		X	X	X		X	X	X			X	X	X	X	X
2013-07-31		X	X	X		X	X	X	X	X	X	X			X	X	X	X	X
2013-08-31		X	X	X		X	X	X	X	X	X	X			X	X	X	X	X
2013-09-30		X	X	X		X	X	X	X	X	X	X			X	X	X	X	X
2013-10-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2013-11-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2013-12-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-01-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-02-28		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-03-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-04-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-05-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-06-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-07-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-08-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-09-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-10-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-11-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2014-12-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-01-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-02-28		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-03-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-04-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-05-31		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-06-30		X	X	X		X	X	X	X	X	X	X			X	X	X		X
2015-07-31	X	X	X	X		X	X	X	X	X	X	X				X	X		X
2015-08-31	X	X	X	X		X	X	X	X	X	X	X				X	X		X
2015-09-30	X	X	X	X		X	X	X	X	X	X	X				X	X		X
2015-10-31	X	X	X	X		X	X	X	X	X	X	X				X	X		X
2015-11-30	X		X	X		X	X	X	X	X	X	X				X	X		X
2015-12-31	X		X	X		X	X		X	X	X	X		X		X	X		X

Appendix C – continued

Nasdaq Constituents between 2011-2015

Month/Stock	DISH	DLTR	DTV	EA	EBAY	ENDP	EQIX	ESRX	EXPD	EXPE	FAST	FB	FFIV	FISV	FLEX	FLIR	FOSL	FOXA	FSLR
2010-12-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-01-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-02-28		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-03-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-04-30		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-05-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-06-30		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-07-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-08-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-09-30		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-10-31		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-11-30		X	X	X	X			X	X	X	X		X	X	X	X		X	X
2011-12-31		X	X	X	X			X	X	X	X		X	X	X		X	X	X
2012-01-31		X	X	X	X			X	X	X	X		X	X	X		X	X	X
2012-02-29		X	X	X	X			X	X	X	X		X	X	X		X	X	X
2012-03-31		X	X	X	X			X	X	X	X		X	X	X		X	X	X
2012-04-30		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-05-31		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-06-30		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-07-31		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-08-31		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-09-30		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-10-31		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-11-30		X	X	X	X			X	X	X	X		X	X	X		X	X	
2012-12-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-01-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-02-28		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-03-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-04-30		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-05-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-06-30		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-07-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-08-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-09-30		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-10-31		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-11-30		X	X		X		X	X	X	X	X	X	X	X			X	X	
2013-12-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-01-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-02-28	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-03-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-04-30	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-05-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-06-30	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-07-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-08-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-09-30	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-10-31	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-11-30	X	X	X		X		X	X	X	X	X	X	X	X				X	
2014-12-31	X	X	X	X	X		X	X	X		X	X		X				X	
2015-01-31	X	X	X	X	X		X	X	X		X	X		X				X	
2015-02-28	X	X	X	X	X		X	X	X		X	X		X				X	
2015-03-31	X	X	X	X	X			X	X		X	X		X				X	
2015-04-30	X	X	X	X	X			X	X		X	X		X				X	
2015-05-31	X	X	X	X	X			X	X		X	X		X				X	
2015-06-30	X	X	X	X	X			X	X		X	X		X				X	
2015-07-31	X	X		X	X			X	X		X	X		X				X	
2015-08-31	X	X		X	X			X	X		X	X		X				X	
2015-09-30	X	X		X	X			X	X		X	X		X				X	
2015-10-31	X	X		X	X			X	X		X	X		X				X	
2015-11-30	X	X		X	X			X	X		X	X		X				X	
2015-12-31	X	X		X	X	X		X		X	X	X		X				X	

Month/Stock	GENZ	GILD	GMCR	GOLD	GOOGL	GRMN	HSIC	ILMN	INCY	INFY	INTC	INTU	ISRG	JD	JOYG	KHC	KLAC	LBTYA	LIFE
2010-12-31	X	X			X	X	X	X		X	X	X	X		X		X		X
2011-01-31	X	X			X	X	X	X		X	X	X	X		X		X		X
2011-02-28	X	X			X	X	X	X		X	X	X	X		X		X		X
2011-03-31	X	X			X	X	X	X		X	X	X	X		X		X		X
2011-04-30	X	X			X	X	X	X		X	X	X	X		X		X		X
2011-05-31		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-06-30		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-07-31		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-08-31		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-09-30		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-10-31		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-11-30		X	X		X	X	X	X		X	X	X	X		X		X		X
2011-12-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-01-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-02-29		X	X	X	X	X	X			X	X	X	X				X		X
2012-03-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-04-30		X	X	X	X	X	X			X	X	X	X				X		X
2012-05-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-06-30		X	X	X	X	X	X			X	X	X	X				X		X
2012-07-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-08-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-09-30		X	X	X	X	X	X			X	X	X	X				X		X
2012-10-31		X	X	X	X	X	X			X	X	X	X				X		X
2012-11-30		X	X	X	X	X	X			X	X	X	X				X		X
2012-12-31		X		X	X	X	X				X	X	X				X	X	X
2013-01-31		X		X	X	X	X				X	X	X				X	X	X
2013-02-28		X		X	X	X	X				X	X	X				X	X	X
2013-03-31		X		X	X	X	X				X	X	X			X	X	X	X
2013-04-30		X		X	X	X	X				X	X	X			X	X	X	X
2013-05-31		X		X	X	X	X				X	X	X			X	X	X	X
2013-06-30		X		X	X	X	X				X	X	X			X	X	X	X
2013-07-31		X		X	X	X	X				X	X	X			X	X	X	X
2013-08-31		X	X	X	X	X	X				X	X	X			X	X	X	
2013-09-30		X	X	X	X	X	X				X	X	X			X	X	X	
2013-10-31		X	X	X	X	X	X				X	X	X			X	X	X	
2013-11-30		X	X		X	X	X				X	X	X			X	X	X	
2013-12-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-01-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-02-28		X	X		X	X	X	X			X	X	X			X	X	X	
2014-03-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-04-30		X	X		X	X	X	X			X	X	X			X	X	X	
2014-05-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-06-30		X	X		X	X	X	X			X	X	X			X	X	X	
2014-07-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-08-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-09-30		X	X		X	X	X	X			X	X	X			X	X	X	
2014-10-31		X	X		X	X	X	X			X	X	X			X	X	X	
2014-11-30		X	X		X	X	X	X			X	X	X			X	X	X	
2014-12-31		X	X		X	X	X	X			X	X	X			X	X	X	
2015-01-31		X	X		X	X	X	X			X	X	X			X	X	X	
2015-02-28		X	X		X	X	X	X			X	X	X			X	X	X	
2015-03-31		X	X		X	X	X	X			X	X	X			X	X	X	
2015-04-30		X	X		X	X	X	X			X	X	X			X	X	X	
2015-05-31		X	X		X	X	X	X			X	X	X			X	X	X	
2015-06-30		X	X		X	X	X	X			X	X	X			X	X	X	
2015-07-31		X	X		X	X	X	X			X	X	X	X		X	X	X	
2015-08-31		X	X		X	X	X	X			X	X	X	X		X	X	X	
2015-09-30		X	X		X	X	X	X			X	X	X	X		X	X	X	
2015-10-31		X	X		X	X	X	X	X		X	X	X	X		X	X	X	
2015-11-30		X	X		X	X	X	X	X		X	X	X	X		X	X	X	
2015-12-31		X			X		X	X	X		X	X	X	X		X	X	X	

Appendix C – continued

Nasdaq Constituents between 2011-2015

Month/Stock	LLTC	LMCA	LRCX	MAR	MAT	MCHP	MDLZ	MICC	MNST	MRVL	MSFT	MU	MXIM	MYL	NCLH	NFLX	NIHD	NTAP	NUAN
2010-12-31	X		X		X	X		X		X	X	X	X	X		X	X	X	
2011-01-31	X		X		X	X		X		X	X	X	X	X		X	X	X	
2011-02-28	X		X		X	X		X		X	X	X	X	X		X	X	X	
2011-03-31	X		X		X	X		X		X	X	X	X	X		X	X	X	
2011-04-30	X		X		X	X		X		X	X	X	X	X		X	X	X	
2011-05-31	X		X		X	X				X	X	X	X	X		X	X	X	
2011-06-30	X		X		X	X				X	X	X	X	X		X	X	X	
2011-07-31	X		X		X	X				X	X	X	X	X		X	X	X	
2011-08-31	X		X		X	X				X	X	X	X	X		X	X	X	
2011-09-30	X		X		X	X				X	X	X	X	X		X	X	X	
2011-10-31	X		X		X	X				X	X	X	X	X		X	X	X	
2011-11-30	X		X		X	X				X	X	X	X	X		X	X	X	
2011-12-31	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-01-31	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-02-29	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-03-31	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-04-30	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-05-31	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-06-30	X		X		X	X			X	X	X	X	X	X		X		X	X
2012-07-31	X		X		X	X	X		X	X	X	X	X	X		X		X	X
2012-08-31	X		X		X	X	X		X	X	X	X	X	X		X		X	X
2012-09-30	X		X		X	X	X		X	X	X	X	X	X		X		X	X
2012-10-31	X		X		X	X	X		X	X	X	X	X	X		X		X	X
2012-11-30	X		X		X	X	X		X	X	X	X	X	X		X		X	X
2012-12-31	X	X			X	X	X		X		X	X	X	X				X	X
2013-01-31	X				X	X	X		X		X	X	X	X				X	X
2013-02-28	X				X	X	X		X		X	X	X	X				X	X
2013-03-31	X				X	X	X		X		X	X	X	X				X	X
2013-04-30	X				X	X	X		X		X	X	X	X				X	X
2013-05-31	X				X	X	X		X		X	X	X	X				X	X
2013-06-30	X	X			X	X	X		X		X	X	X	X		X		X	X
2013-07-31	X	X			X	X	X		X		X	X	X	X		X		X	X
2013-08-31	X	X			X	X	X		X		X	X	X	X		X		X	X
2013-09-30	X	X			X	X	X		X		X	X	X	X		X		X	X
2013-10-31	X	X			X	X	X		X		X	X	X	X		X		X	X
2013-11-30	X	X		X	X	X	X		X		X	X	X	X		X		X	X
2013-12-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-01-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-02-28	X	X		X	X		X		X		X	X	X	X		X		X	
2014-03-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-04-30	X	X		X	X		X		X		X	X	X	X		X		X	
2014-05-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-06-30	X	X		X	X		X		X		X	X	X	X		X		X	
2014-07-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-08-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-09-30	X	X		X	X		X		X		X	X	X	X		X		X	
2014-10-31	X	X		X	X		X		X		X	X	X	X		X		X	
2014-11-30	X	X		X	X		X		X		X	X	X	X		X		X	
2014-12-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-01-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-02-28	X	X	X	X	X		X		X		X	X		X		X		X	
2015-03-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-04-30	X	X	X	X	X		X		X		X	X		X		X		X	
2015-05-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-06-30	X	X	X	X	X		X		X		X	X		X		X		X	
2015-07-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-08-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-09-30	X	X	X	X	X		X		X		X	X		X		X		X	
2015-10-31	X	X	X	X	X		X		X		X	X		X		X		X	
2015-11-30	X	X	X	X	X		X		X		X	X		X		X		X	
2015-12-31	X	X	X	X	X		X		X		X	X	X	X	X	X		X	

Month/Stock	NVDA	NXPI	ORCL	ORLY	PAYX	PCAR	PCLN	PRGO	PYPL	QCOM	QGEN	QVCA	REGN	RIMM	ROST	SBAC	SBUX	SHLD	SIAL
2010-12-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-01-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-02-28	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-03-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-04-30	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-05-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-06-30	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-07-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-08-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-09-30	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-10-31	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-11-30	X		X	X	X	X	X			X	X	X		X	X		X	X	X
2011-12-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-01-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-02-29	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-03-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-04-30	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-05-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-06-30	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-07-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-08-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-09-30	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-10-31	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-11-30	X		X	X	X	X	X	X		X		X		X	X		X	X	X
2012-12-31	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-01-31	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-02-28	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-03-31	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-04-30	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-05-31	X		X	X	X	X	X	X		X		X	X		X	X	X	X	X
2013-06-30	X		X	X	X	X	X			X		X	X		X	X	X	X	X
2013-07-31	X			X	X	X	X			X		X	X		X	X	X	X	X
2013-08-31	X			X	X	X	X			X		X	X		X	X	X	X	X
2013-09-30	X			X	X	X	X			X		X	X		X	X	X	X	X
2013-10-31	X			X	X	X	X			X		X	X		X	X	X	X	X
2013-11-30	X			X	X	X	X			X		X	X		X	X	X	X	X
2013-12-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-01-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-02-28	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-03-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-04-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-05-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-06-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-07-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-08-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-09-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-10-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-11-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2014-12-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-01-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-02-28	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-03-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-04-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-05-31	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-06-30	X	X		X	X	X	X			X		X	X		X	X	X		X
2015-07-31	X	X		X	X	X	X			X		X	X		X	X	X		
2015-08-31	X	X		X	X	X	X			X		X	X		X	X	X		
2015-09-30	X	X		X	X	X	X			X		X	X		X	X	X		
2015-10-31	X	X		X	X	X	X			X		X	X		X	X	X		
2015-11-30	X	X		X	X	X	X		X	X		X	X		X	X	X		
2015-12-31	X	X		X	X	X	X		X	X		X	X		X	X	X		

Appendix C – continued

Nasdaq Constituents between 2011-2015

Month/Stock	SIRI	SNDK	SPLS	SRCL	STRZA	STX	SWKS	SYMC	TEVA	TMUS	TRIP	TSCO	TSLA	TXN	ULTA	URBN	VIAB	VIP	VMED
2010-12-31		X	X	X		X		X	X							X			X
2011-01-31		X	X	X		X		X	X							X			X
2011-02-28		X	X	X		X		X	X							X			X
2011-03-31		X	X	X		X		X	X							X			X
2011-04-30		X	X	X		X		X	X							X			X
2011-05-31		X	X	X		X		X	X							X			X
2011-06-30		X	X	X		X		X	X							X			X
2011-07-31	X	X	X	X		X		X	X							X			X
2011-08-31	X	X	X	X		X		X	X							X			X
2011-09-30	X	X	X	X		X		X	X							X			X
2011-10-31	X	X	X	X		X		X	X							X			X
2011-11-30	X	X	X	X		X		X	X							X			X
2011-12-31	X	X	X	X		X		X	X										X
2012-01-31	X	X	X	X		X		X	X										X
2012-02-29	X	X	X	X		X		X	X										X
2012-03-31	X	X	X	X		X		X	X										X
2012-04-30	X	X	X	X		X		X	X					X					X
2012-05-31	X	X	X	X		X		X						X			X		X
2012-06-30	X	X	X	X		X		X						X			X		X
2012-07-31	X	X	X	X		X		X						X			X		X
2012-08-31	X	X	X	X		X		X						X			X		X
2012-09-30	X	X	X	X		X		X						X			X		X
2012-10-31	X	X	X	X		X		X						X			X		X
2012-11-30	X	X	X	X		X		X						X			X		X
2012-12-31	X	X	X	X		X		X						X			X		X
2013-01-31	X	X	X	X	X	X		X						X			X		X
2013-02-28	X	X	X	X	X	X		X						X			X		X
2013-03-31	X	X	X	X		X		X						X			X		X
2013-04-30	X	X	X	X		X		X						X			X		X
2013-05-31	X	X	X	X		X		X						X			X		X
2013-06-30	X	X	X	X		X		X						X			X		
2013-07-31	X	X	X	X		X		X					X	X			X		
2013-08-31	X	X	X	X		X		X					X	X			X		
2013-09-30	X	X	X	X		X		X					X	X			X		
2013-10-31	X	X	X	X		X		X					X	X			X	X	
2013-11-30	X	X	X	X		X		X					X	X			X	X	
2013-12-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-01-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-02-28	X	X	X	X		X		X			X	X	X	X			X	X	
2014-03-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-04-30	X	X	X	X		X		X			X	X	X	X			X	X	
2014-05-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-06-30	X	X	X	X		X		X			X	X	X	X			X	X	
2014-07-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-08-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-09-30	X	X	X	X		X		X			X	X	X	X			X	X	
2014-10-31	X	X	X	X		X		X			X	X	X	X			X	X	
2014-11-30	X	X	X	X		X		X			X	X	X	X			X	X	
2014-12-31	X	X	X	X		X		X			X	X	X	X			X	X	
2015-01-31	X	X	X	X		X		X			X	X	X	X			X	X	
2015-02-28	X	X	X	X		X		X			X	X	X	X			X	X	
2015-03-31	X	X	X	X		X		X			X	X	X	X			X	X	
2015-04-30	X	X	X	X		X		X			X	X	X	X			X	X	
2015-05-31	X	X	X	X		X		X			X	X	X	X			X	X	
2015-06-30	X	X	X	X		X		X			X	X	X	X			X	X	
2015-07-31	X	X	X	X		X		X			X	X	X	X			X	X	
2015-08-31	X	X	X	X		X	X	X			X	X	X	X			X	X	
2015-09-30	X	X	X	X		X	X	X			X	X	X	X			X	X	
2015-10-31	X	X	X	X		X	X	X			X	X	X	X			X	X	
2015-11-30	X	X	X	X		X	X	X			X	X	X	X			X	X	
2015-12-31	X	X		X		X	X	X		X	X	X	X	X	X		X		

Month/Stock	VOD	VRSK	VRSN	VRTX	WBA	WCRX	WDC	WFM	WYNN	XLNX	XRAY	YHOO
2010-12-31	X		X	X		X		X	X	X	X	X
2011-01-31	X		X	X		X		X	X	X	X	X
2011-02-28	X		X	X		X		X	X	X	X	X
2011-03-31	X		X	X		X		X	X	X	X	X
2011-04-30	X		X	X		X		X	X	X	X	X
2011-05-31	X		X	X		X		X	X	X	X	X
2011-06-30	X		X	X		X		X	X	X	X	X
2011-07-31	X		X	X		X		X	X	X	X	X
2011-08-31	X		X	X		X		X	X	X	X	X
2011-09-30	X		X	X		X		X	X	X	X	X
2011-10-31	X		X	X		X		X	X	X	X	X
2011-11-30	X		X	X		X		X	X	X	X	X
2011-12-31	X		X	X		X		X	X	X	X	X
2012-01-31	X		X	X		X		X	X	X	X	X
2012-02-29	X		X	X		X		X	X	X	X	X
2012-03-31	X		X	X		X		X	X	X	X	X
2012-04-30	X		X	X		X		X	X	X	X	X
2012-05-31	X		X	X		X		X	X	X	X	X
2012-06-30	X		X	X		X		X	X	X	X	X
2012-07-31	X		X	X		X		X	X	X	X	X
2012-08-31	X		X	X		X		X	X	X	X	X
2012-09-30	X		X	X		X		X	X	X	X	X
2012-10-31	X		X	X		X		X	X	X	X	X
2012-11-30	X		X	X		X		X	X	X	X	X
2012-12-31	X	X		X			X	X	X	X	X	X
2013-01-31	X	X		X			X	X	X	X	X	X
2013-02-28	X	X		X			X	X	X	X	X	X
2013-03-31	X	X		X			X	X	X	X	X	X
2013-04-30	X	X		X			X	X	X	X	X	X
2013-05-31	X	X		X			X	X	X	X	X	X
2013-06-30	X	X		X			X	X	X	X	X	X
2013-07-31	X	X		X			X	X	X	X	X	X
2013-08-31	X	X		X			X	X	X	X	X	X
2013-09-30	X	X		X			X	X	X	X	X	X
2013-10-31	X	X		X			X	X	X	X	X	X
2013-11-30	X	X		X			X	X	X	X	X	X
2013-12-31	X	X		X			X	X	X	X		X
2014-01-31	X	X		X			X	X	X	X		X
2014-02-28	X	X		X			X	X	X	X		X
2014-03-31	X	X		X			X	X	X	X		X
2014-04-30	X	X		X			X	X	X	X		X
2014-05-31	X	X		X			X	X	X	X		X
2014-06-30	X	X		X			X	X	X	X		X
2014-07-31	X	X		X			X	X	X	X		X
2014-08-31	X	X		X			X	X	X	X		X
2014-09-30	X	X		X			X	X	X	X		X
2014-10-31	X	X		X			X	X	X	X		X
2014-11-30	X	X		X			X	X	X	X		X
2014-12-31	X	X		X			X	X	X	X		X
2015-01-31	X	X		X			X	X	X	X		X
2015-02-28	X	X		X			X	X	X	X		X
2015-03-31	X	X		X	X		X	X	X	X		X
2015-04-30	X	X		X	X		X	X	X	X		X
2015-05-31	X	X		X	X		X	X	X	X		X
2015-06-30	X	X		X	X		X	X	X	X		X
2015-07-31	X	X		X	X		X	X	X	X		X
2015-08-31	X	X		X	X		X	X	X	X		X
2015-09-30	X	X		X	X		X	X	X	X		X
2015-10-31	X	X		X	X		X	X	X	X		X
2015-11-30	X	X		X	X		X	X	X	X		X
2015-12-31	X	X		X	X		X	X		X		X

Appendix D

Time-zone and Date Adjustments

Quarterly Earnings Release Dates for Top 55 Nasdaq Constituents from 2011-2015

Quarter/Stock	AAPL	ADBE	ADP	ADSK	AKAM	AMAT	AMGN	AMZN	ATVI	BBBY	BIDU	BIIB	CA	CELG	CERN	CHKP	CMCSA	COST	CSCO
2011Q1	2011-01-18	2011-03-22	2011-01-26	2011-02-24	2011-02-09	2011-02-24	2011-01-24	2011-01-27	2011-02-09	2011-04-06	2011-01-31	2011-02-01	2011-01-25	2011-01-27	2011-02-08	2011-01-31	2011-02-16	2011-03-02	2011-02-09
2011Q2	2011-04-20	2011-06-21	2011-05-02	2011-05-19	2011-04-27	2011-05-24	2011-04-20	2011-04-26	2011-05-09	2011-06-22	2011-04-27	2011-04-21	2011-05-12	2011-04-28	2011-04-28	2011-04-14	2011-05-03	2011-05-25	2011-05-11
2011Q3	2011-07-19	2011-09-20	2011-07-28	2011-08-18	2011-07-27	2011-08-24	2011-07-29	2011-07-26	2011-08-03	2011-09-21	2011-07-25	2011-07-26	2011-07-20	2011-07-28	2011-07-28	2011-07-18	2011-08-03	2011-10-05	2011-08-10
2011Q4	2011-10-18	2011-12-15	2011-10-26	2011-11-15	2011-10-26	2011-11-16	2011-10-24	2011-10-25	2011-11-08	2011-12-21	2011-10-27	2011-10-28	2011-10-26	2011-10-27	2011-10-27	2011-10-18	2011-11-02	2011-12-08	2011-11-09
2012Q1	2012-01-24	2012-03-19	2012-01-25	2012-02-23	2012-02-08	2012-02-16	2012-01-28	2012-01-31	2012-02-09	2012-04-04	2012-02-16	2012-01-31	2012-01-24	2012-01-26	2012-02-07	2012-01-17	2012-02-15	2012-02-29	2012-02-08
2012Q2	2012-04-24	2012-06-19	2012-05-01	2012-05-17	2012-04-25	2012-05-17	2012-04-24	2012-04-26	2012-05-09	2012-06-20	2012-04-24	2012-05-01	2012-05-10	2012-04-26	2012-04-26	2012-04-23	2012-05-02	2012-05-24	2012-05-09
2012Q3	2012-07-24	2012-09-19	2012-08-01	2012-08-23	2012-07-25	2012-08-15	2012-07-26	2012-07-26	2012-08-02	2012-09-19	2012-07-23	2012-07-24	2012-07-26	2012-07-26	2012-07-26	2012-07-18	2012-08-01	2012-10-10	2012-08-15
2012Q4	2012-10-25	2012-12-13	2012-11-01	2012-11-15	2012-10-24	2012-11-15	2012-10-23	2012-10-25	2012-11-07	2012-12-19	2012-10-29	2012-10-25	2012-10-25	2012-10-25	2012-10-25	2012-10-17	2012-10-26	2012-12-12	2012-11-13
2013Q1	2013-01-23	2013-03-19	2013-02-05	2013-02-25	2013-02-06	2013-02-13	2013-01-23	2013-01-29	2013-02-07		2013-02-04	2013-01-28	2013-01-22	2013-01-24	2013-02-05	2013-01-23	2013-02-12	2013-03-12	2013-02-13
2013Q2	2013-04-23	2013-06-18	2013-05-03	2013-05-16	2013-04-24	2013-05-16	2013-04-23	2013-04-25	2013-05-08	2013-04-10	2013-04-25	2013-04-25	2013-05-07	2013-04-25	2013-04-25	2013-04-22	2013-05-01	2013-05-30	2013-05-15
2013Q3	2013-07-23	2013-09-17	2013-08-01	2013-08-22	2013-07-24	2013-08-15	2013-07-30	2013-07-25	2013-08-01	2013-06-26	2013-07-24	2013-07-25	2013-07-24	2013-07-25	2013-07-25	2013-07-18	2013-07-31	2013-10-09	2013-08-14
2013Q4	2013-10-28	2013-12-12	2013-10-30	2013-11-21	2013-10-23	2013-11-14	2013-10-22	2013-10-24	2013-11-06	2013-09-25	2013-10-29	2013-10-28	2013-10-24	2013-10-24	2013-10-24	2013-10-21	2013-10-30	2013-12-11	2013-11-13
2014Q1	2014-01-27	2014-03-18	2014-02-05	2014-02-26	2014-02-05	2014-02-12	2014-01-28	2014-01-30	2014-02-06	2014-01-08	2014-02-26	2014-01-29	2014-01-21	2014-01-30	2014-02-04	2014-01-28	2014-01-28	2014-03-06	2014-02-12
2014Q2	2014-04-23	2014-06-17	2014-04-30	2014-05-15	2014-05-01	2014-05-15	2014-04-22	2014-04-24	2014-05-06	2014-04-09	2014-04-24	2014-04-23	2014-05-15	2014-04-24	2014-04-24	2014-04-29	2014-04-22	2014-05-29	2014-05-14
2014Q3	2014-07-22	2014-09-16	2014-07-31	2014-08-14	2014-07-30	2014-08-14	2014-07-29	2014-07-24	2014-08-05	2014-06-25	2014-07-24	2014-07-23	2014-07-23	2014-07-24	2014-07-24	2014-07-23	2014-07-22	2014-10-08	2014-08-13
2014Q4	2014-10-20	2014-12-11	2014-10-29	2014-11-20	2014-10-29	2014-11-13	2014-10-27	2014-10-23	2014-11-04	2014-09-23	2014-10-29	2014-10-22	2014-10-22	2014-10-23	2014-10-23	2014-10-23	2014-10-23	2014-12-10	2014-11-12
2015Q1	2015-01-27	2015-03-17	2015-02-04	2015-02-26	2015-02-10	2015-02-11	2015-01-27	2015-01-29	2015-02-05	2015-01-08	2015-02-11	2015-01-29	2015-01-20	2015-01-29	2015-02-10	2015-01-29	2015-02-24	2015-03-05	2015-02-11
2015Q2	2015-04-27	2015-06-16	2015-04-30	2015-05-19	2015-04-28	2015-05-14	2015-04-21	2015-04-23	2015-05-06	2015-04-08	2015-04-29	2015-04-24	2015-05-07	2015-04-30	2015-05-07	2015-04-20	2015-05-04	2015-05-27	2015-05-13
2015Q3	2015-07-21	2015-09-17	2015-07-30	2015-08-27	2015-07-28	2015-08-13	2015-07-30	2015-07-23	2015-08-04	2015-06-24	2015-07-27	2015-07-24	2015-07-23	2015-07-23	2015-08-04	2015-07-22	2015-07-23	2015-09-29	2015-08-12
2015Q4	2015-10-27	2015-12-10	2015-10-28	2015-11-19	2015-10-27	2015-11-12	2015-10-28	2015-10-22	2015-11-02	2015-09-24	2015-10-29	2015-10-21	2015-10-21	2015-11-05	2015-11-03	2015-10-26	2015-10-27	2015-12-08	2015-11-12

Quarter/Stock	CTSH	CTXS	DLTR	EBAY	ESRX	FAST	FISV	GILD	GOOGL	HSIC	INTC	INTU	ISRG	KLAC	LLTC	MAT	MSFT	MYL
2011Q1	2011-02-07	2011-01-26	2011-02-23	2011-01-19	2011-02-16	2011-01-18	2011-02-03	2011-01-25	2011-01-20	2011-02-22	2011-01-13	2011-02-17	2011-01-20	2011-01-27	2011-01-18	2011-02-02	2011-01-27	2011-02-24
2011Q2	2011-05-03	2011-04-27	2011-05-19	2011-04-27	2011-04-25	2011-04-12	2011-04-27	2011-04-20	2011-04-14	2011-05-03	2011-04-19	2011-05-19	2011-04-19	2011-04-28	2011-04-19	2011-04-15	2011-04-28	2011-05-03
2011Q3	2011-08-02	2011-07-27	2011-08-18	2011-07-20	2011-07-21	2011-07-12	2011-07-26	2011-07-26	2011-07-14	2011-08-02	2011-07-20	2011-08-18	2011-07-19	2011-07-28	2011-07-26	2011-07-15	2011-07-21	2011-07-27
2011Q4	2011-11-02	2011-10-27	2011-11-17	2011-10-19	2011-10-25	2011-10-13	2011-11-01	2011-10-27	2011-10-13	2011-11-01	2011-10-18	2011-11-17	2011-10-18	2011-10-27	2011-10-18	2011-10-14	2011-10-20	2011-10-26
2012Q1	2012-02-08	2012-01-25	2012-02-22	2012-01-18	2012-02-22	2012-01-18	2012-02-02	2012-02-04	2012-01-19	2012-02-15	2012-01-19	2012-02-21	2012-01-19	2012-01-26	2012-01-17	2012-01-31	2012-01-19	2012-02-21
2012Q2	2012-05-07	2012-04-25	2012-05-17	2012-04-18	2012-05-10	2012-04-12	2012-05-01	2012-04-26	2012-04-12	2012-05-08	2012-04-17	2012-05-17	2012-04-17	2012-04-26	2012-04-17	2012-04-16	2012-04-19	2012-04-26
2012Q3	2012-08-06	2012-07-25	2012-08-16	2012-07-18	2012-08-07	2012-07-12	2012-07-30	2012-07-26	2012-07-19	2012-08-02	2012-07-17	2012-08-21	2012-07-19	2012-07-26	2012-07-24	2012-07-17	2012-07-19	2012-07-26
2012Q4	2012-11-07	2012-10-24	2012-11-15	2012-10-17	2012-11-05	2012-10-11	2012-10-30	2012-10-23	2012-10-18	2012-11-07	2012-10-16	2012-11-15	2012-10-16	2012-10-25	2012-10-16	2012-10-16	2012-10-18	2012-10-25
2013Q1	2013-02-07	2013-01-30	2013-02-27	2013-01-16	2013-02-18	2013-01-17	2013-02-05	2013-02-04	2013-01-22	2013-02-13	2013-01-17	2013-02-21	2013-01-22	2013-01-24	2013-01-15	2013-02-01	2013-01-24	2013-02-27
2013Q2	2013-05-08	2013-04-24	2013-05-23	2013-04-17	2013-04-29	2013-04-10	2013-04-30	2013-05-02	2013-04-18	2013-05-07	2013-04-16	2013-05-21	2013-04-18	2013-04-25	2013-04-16	2013-04-17	2013-04-18	2013-05-02
2013Q3	2013-08-06	2013-07-24	2013-08-22	2013-07-17	2013-07-29	2013-07-10	2013-07-30	2013-07-25	2013-07-18	2013-08-06	2013-07-17	2013-08-20	2013-07-18	2013-07-25	2013-07-23	2013-07-17	2013-07-18	2013-08-01
2013Q4	2013-11-05	2013-10-23	2013-11-21	2013-10-16	2013-10-24	2013-10-09	2013-10-29	2013-10-29	2013-10-17	2013-11-05	2013-10-15	2013-11-21	2013-10-17	2013-10-24	2013-10-15	2013-10-16	2013-10-24	2013-10-31
2014Q1	2014-02-05	2014-01-29	2014-02-26	2014-01-22	2014-02-20	2014-01-15	2014-02-05	2014-02-04	2014-01-30	2014-02-11	2014-01-16	2014-02-20	2014-01-23	2014-01-23	2014-01-14	2014-01-31	2014-01-23	2014-02-27
2014Q2	2014-05-07	2014-04-23	2014-05-22	2014-04-29	2014-04-29	2014-04-11	2014-04-29	2014-04-22	2014-04-16	2014-05-06	2014-04-15	2014-05-20	2014-04-22	2014-04-24	2014-04-15	2014-04-17	2014-04-24	2014-05-01
2014Q3	2014-08-06	2014-07-23	2014-08-21	2014-07-16	2014-07-29	2014-07-11	2014-07-29	2014-07-23	2014-07-17	2014-08-04	2014-07-15	2014-08-21	2014-07-22	2014-07-24	2014-07-22	2014-07-17	2014-07-22	2014-08-07
2014Q4	2014-11-05	2014-10-22	2014-11-20	2014-10-15	2014-10-28	2014-10-10	2014-10-28	2014-10-28	2014-10-16	2014-11-06	2014-10-14	2014-11-20	2014-10-21	2014-10-23	2014-10-14	2014-10-16	2014-10-23	2014-10-30
2015Q1	2015-02-04	2015-01-28	2015-02-25	2015-01-21	2015-02-23	2015-01-15	2015-02-03	2015-02-03	2015-01-29	2015-02-11	2015-01-15	2015-02-19	2015-01-22	2015-01-22	2015-01-13	2015-01-30	2015-01-26	2015-03-02
2015Q2	2015-05-04	2015-04-22	2015-05-21	2015-04-22	2015-04-28	2015-04-14	2015-05-05	2015-04-30	2015-04-23	2015-05-04	2015-04-14	2015-05-21	2015-04-21	2015-04-23	2015-04-14	2015-04-16	2015-04-23	2015-05-05
2015Q3	2015-08-05	2015-07-28	2015-09-01	2015-07-16	2015-07-28	2015-07-14	2015-07-29	2015-07-28	2015-07-16	2015-07-29	2015-07-15	2015-08-20	2015-07-21	2015-07-30	2015-07-21	2015-07-16	2015-07-21	2015-08-06
2015Q4	2015-11-04	2015-10-21	2015-11-24	2015-10-21	2015-10-27	2015-10-13	2015-10-27	2015-10-27	2015-10-22	2015-11-04	2015-10-13	2015-11-19	2015-10-20	2015-10-21	2015-10-13	2015-10-15	2015-10-22	2015-10-30

Quarter/Stock	MU	NTAP	NVDA	ORLY	PAYX	PCAR	PCLN	QCOM	ROST	SBUX	SNDK	SRCL	STX	SYMC	VOD	VRTX	XLNX	YHOO
2011Q1	2011-03-23	2011-02-16	2011-02-16	2011-02-16	2011-03-23	2011-02-01	2011-02-23	2011-01-26	2011-03-17	2011-01-26	2011-01-27	2011-02-03	2011-01-19	2011-01-26	2011-02-03	2011-02-03	2011-01-19	2011-01-25
2011Q2	2011-06-23	2011-05-25	2011-05-12	2011-04-27	2011-06-22	2011-04-19	2011-05-05	2011-04-20	2011-05-19	2011-04-27	2011-04-21	2011-04-27	2011-04-19	2011-05-11	2011-05-17	2011-05-03	2011-04-27	2011-04-19
2011Q3	2011-09-29	2011-08-17	2011-08-11	2011-07-27	2011-09-27	2011-07-26	2011-08-04	2011-07-20	2011-08-18	2011-07-28	2011-07-21	2011-07-27	2011-07-20	2011-07-27	2011-07-22	2011-07-28	2011-07-20	2011-07-19
2011Q4	2011-12-21	2011-11-16	2011-11-10	2011-10-26	2011-12-20	2011-10-25	2011-11-07	2011-11-02	2011-11-17	2011-11-03	2011-10-20	2011-10-26	2011-10-20	2011-10-26	2011-11-08	2011-10-27	2011-10-19	2011-10-18
2012Q1	2012-03-22	2012-02-15	2012-02-15	2012-02-08	2012-03-28	2012-01-31	2012-02-27	2012-02-01	2012-03-15	2012-01-26	2012-01-25	2012-02-02	2012-01-31	2012-01-25	2012-02-09	2012-02-02	2012-01-18	2012-01-24
2012Q2	2012-06-20	2012-05-23	2012-05-11	2012-04-25	2012-06-27	2012-04-24	2012-05-09	2012-04-18	2012-05-17	2012-04-26	2012-04-19	2012-04-25	2012-04-17	2012-05-02	2012-05-22	2012-04-26	2012-04-25	2012-04-17
2012Q3	2012-09-27	2012-08-15	2012-08-09	2012-07-25	2012-09-24	2012-07-24	2012-08-07	2012-07-18	2012-08-16	2012-07-26	2012-07-19	2012-07-25	2012-07-30	2012-07-25	2012-07-20	2012-07-30	2012-07-18	2012-07-17
2012Q4	2012-12-20	2012-11-14	2012-11-08	2012-10-24	2012-12-19	2012-10-25	2012-11-01	2012-11-07	2012-11-15	2012-11-01	2012-10-18	2012-10-24	2012-10-31	2012-10-24	2012-11-13	2012-11-01	2012-10-17	2012-10-22
2013Q1	2013-03-21	2013-02-13	2013-02-13	2013-02-06	2013-03-27	2013-01-31	2013-02-26	2013-01-30	2013-03-21	2013-01-24	2013-01-23	2013-02-06	2013-01-28	2013-01-23	2013-02-07	2013-01-29	2013-01-17	2013-01-28
2013Q2	2013-06-19	2013-05-21	2013-05-09	2013-04-24	2013-06-26	2013-04-23	2013-05-09	2013-04-24	2013-05-23	2013-04-25	2013-04-17	2013-04-24	2013-05-01	2013-05-07	2013-05-21	2013-04-30	2013-04-24	2013-04-16
2013Q3	2013-10-10	2013-08-14	2013-08-08	2013-07-24	2013-09-30	2013-07-23	2013-08-08	2013-07-24	2013-08-22	2013-07-25	2013-07-17	2013-07-24	2013-07-24	2013-07-30	2013-07-19	2013-07-29	2013-07-17	2013-07-16
2013Q4	2014-01-07	2013-11-13	2013-11-07	2013-10-23	2013-12-18	2013-10-29	2013-11-07	2013-11-06	2013-11-21	2013-10-30	2013-10-16	2013-10-23	2013-10-28	2013-10-23	2013-11-12	2013-10-29	2013-10-16	2013-10-15
2014Q1	2014-04-03	2014-02-12	2014-02-12	2014-02-05	2014-03-26	2014-01-31	2014-02-20	2014-01-29	2014-02-27	2014-01-23	2014-01-22	2014-02-05	2014-01-27	2014-01-29	2014-02-06	2014-01-29	2014-01-21	2014-01-28
2014Q2	2014-06-23	2014-05-21	2014-05-06	2014-04-23	2014-07-01	2014-04-29	2014-05-08	2014-04-23	2014-05-22	2014-04-24	2014-04-16	2014-04-24	2014-04-29	2014-05-08	2014-05-20	2014-05-01	2014-04-23	2014-04-15
2014Q3	2014-09-25	2014-08-13	2014-08-07	2014-07-23	2014-09-24	2014-07-29	2014-08-11	2014-07-23	2014-08-21	2014-07-24	2014-07-16	2014-07-24	2014-07-17	2014-08-06	2014-07-25	2014-07-29	2014-07-22	2014-07-15
2014Q4	2015-01-06	2014-11-12	2014-11-06	2014-10-22	2014-12-19	2014-10-28	2014-11-04	2014-11-05	2014-11-20	2014-10-30	2014-10-16	2014-10-23	2014-10-27	2014-11-05	2014-11-11	2014-10-28	2014-10-16	2014-10-21
2015Q1	2015-04-01	2015-02-11	2015-02-11	2015-02-04	2015-03-25	2015-01-30	2015-02-19	2015-01-28	2015-02-26	2015-01-22	2015-01-21	2015-02-05	2015-01-26	2015-02-05	2015-02-05	2015-01-28	2015-01-21	2015-01-27
2015Q2	2015-06-25	2015-05-20	2015-05-07	2015-04-22	2015-07-01	2015-04-21	2015-05-07	2015-04-22	2015-05-21	2015-04-23	2015-04-15	2015-04-23	2015-04-17	2015-05-14	2015-05-19	2015-04-29	2015-04-22	2015-04-21
2015Q3	2015-10-01	2015-08-19	2015-08-06	2015-07-29	2015-09-30	2015-07-28	2015-08-05	2015-07-22	2015-08-20	2015-07-23	2015-07-22	2015-07-23	2015-07-31	2015-08-11	2015-07-24	2015-07-29	2015-07-22	2015-07-21
2015Q4	2015-12-22	2015-11-18	2015-11-05	2015-10-28	2015-12-22	2015-10-27	2015-11-09	2015-11-04	2015-11-19	2015-10-29	2015-10-21	2015-10-22	2015-10-30	2015-11-05	2015-11-10	2015-10-28	2015-10-14	2015-10-20

Quarterly earnings release dates are systematically flagged and excluded from the news sample of each stock in order to focus on unscheduled news announcements. Quarterly earnings release dates are harvested for each company using a combination of Bloomberg's Earnings History (ERN) function and historical press releases.

Appendix D — continued

Time-zone and Date Adjustments

Nasdaq Trading Periods in ET and UTC Equivalents, Adjusted for Daylight Savings (2011-2015)

			0:00		4:00		9:30		16:00		20:00		0:00
				US Overnight Closed		**US Pre-Market** Thin Trading		**US Market Open** Liquid Trading		**US After-Market** Thin Trading		**US Overnight** Closed	
				00:00:00	04:00:00	04:00:00	09:30:00	09:30:00	16:00:00	16:00:00	20:00:00	20:00:00	00:00:00
Year	**Offset**	**Start Date**	**End Date**	**Start Time**	**End Time**	**Start Time**	**End Time**	**Start Time**	**End Time**	**Start Time**	**End Time**	**Start Time**	**End Time**
2011	UTC - 5	01-Jan-11	12-Mar-11	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00
2011	UTC - 4	13-Mar-11	05-Nov-11	04:00:00	08:00:00	08:00:00	13:30:00	13:30:00	20:00:00	20:00:00	00:00:00	00:00:00	04:00:00
2011	UTC - 5	06-Nov-11	10-Mar-12	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00
2012	UTC - 4	11-Mar-12	03-Nov-12	04:00:00	08:00:00	08:00:00	13:30:00	13:30:00	20:00:00	20:00:00	00:00:00	00:00:00	04:00:00
2012	UTC - 5	04-Nov-12	09-Mar-13	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00
2013	UTC - 4	10-Mar-13	02-Nov-13	04:00:00	08:00:00	08:00:00	13:30:00	13:30:00	20:00:00	20:00:00	00:00:00	00:00:00	04:00:00
2013	UTC - 5	03-Nov-13	08-Mar-14	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00
2014	UTC - 4	09-Mar-14	01-Nov-14	04:00:00	08:00:00	08:00:00	13:30:00	13:30:00	20:00:00	20:00:00	00:00:00	00:00:00	04:00:00
2014	UTC - 5	02-Nov-14	07-Mar-15	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00
2015	UTC - 4	08-Mar-15	31-Oct-15	04:00:00	08:00:00	08:00:00	13:30:00	13:30:00	20:00:00	20:00:00	00:00:00	00:00:00	04:00:00
2015	UTC - 5	01-Nov-15	31-Dec-15	05:00:00	09:00:00	09:00:00	14:30:00	14:30:00	21:00:00	21:00:00	01:00:00	01:00:00	05:00:00

Dummy variables are generated to identify which of the four trading periods news items fall into: the US Pre-Market, Market Open, After-Market, or Overnight session. Given that the news and price data sets are in UTC, time zone adjustments are carried out in order to correctly categorize the sample.

Appendix D — continued

Time-zone and Date Adjustments

UTC Equivalent US Holiday Start and End Dates, Adjusted for Daylight Savings (2011-2015)

Date	Holiday	NASDAQ	UTC Start Date	Start Time	UTC End Date	End Time
Saturday, January 1, 2011	New Year's Day	Closed	NA	NA	NA	NA
Monday, January 17, 2011	Martin Luther King, Jr. Day	Closed	Monday, January 17, 2011	05:00:00	Tuesday, January 18, 2011	05:00:00
Monday, February 21, 2011	President's Day - U.S.	Closed	Monday, February 21, 2011	05:00:00	Tuesday, February 22, 2011	05:00:00
Friday, April 22, 2011	Good Friday	Closed	Friday, April 22, 2011	04:00:00	Saturday, April 23, 2011	04:00:00
Monday, May 30, 2011	Memorial Day - U.S.	Closed	Monday, May 30, 2011	04:00:00	Tuesday, May 31, 2011	04:00:00
Monday, July 4, 2011	Independence Day - U.S.	Closed	Monday, July 4, 2011	04:00:00	Tuesday, July 5, 2011	04:00:00
Monday, September 5, 2011	Labor Day - U.S.	Closed	Monday, September 5, 2011	04:00:00	Tuesday, September 6, 2011	04:00:00
Thursday, November 24, 2011	Thanksgiving Day - U.S.	Closed	Thursday, November 24, 2011	05:00:00	Friday, November 25, 2011	05:00:00
Friday, November 25, 2011	Early Close - U.S.	Close Early	Friday, November 25, 2011	18:00:00	Saturday, November 26, 2011	05:00:00
Saturday, December 24, 2011	Christmas Eve	Close Early	NA	NA	NA	NA
Sunday, December 25, 2011	Christmas Day	Closed	NA	NA	NA	NA
Monday, December 26, 2011	Market Holidays (St. Stephen's Day / Boxing Day	Closed	Monday, December 26, 2011	05:00:00	Tuesday, December 27, 2011	05:00:00
Sunday, January 1, 2012	New Year's Day	Closed	NA	NA	NA	NA
Monday, January 2, 2012	New Year's Day (Observed)	Closed	Monday, January 2, 2012	05:00:00	Tuesday, January 3, 2012	05:00:00
Monday, January 16, 2012	Martin Luther King, Jr. Day	Closed	Monday, January 16, 2012	05:00:00	Tuesday, January 17, 2012	05:00:00
Monday, February 20, 2012	President's Day - U.S.	Closed	Monday, February 20, 2012	05:00:00	Tuesday, February 21, 2012	05:00:00
Friday, April 6, 2012	Good Friday	Closed	Friday, April 6, 2012	04:00:00	Saturday, April 7, 2012	04:00:00
Monday, May 28, 2012	Memorial Day - U.S.	Closed	Monday, May 28, 2012	04:00:00	Tuesday, May 29, 2012	04:00:00
Tuesday, July 3, 2012	Early Close - U.S.	Close Early	Tuesday, July 3, 2012	17:00:00	Wednesday, July 4, 2012	04:00:00
Wednesday, July 4, 2012	Independence Day - U.S.	Closed	Wednesday, July 4, 2012	04:00:00	Thursday, July 5, 2012	04:00:00
Monday, September 3, 2012	Labor Day - U.S.	Closed	Monday, September 3, 2012	04:00:00	Tuesday, September 4, 2012	04:00:00
Thursday, November 22, 2012	Thanksgiving Day - U.S.	Closed	Thursday, November 22, 2012	05:00:00	Friday, November 23, 2012	05:00:00
Friday, November 23, 2012	Early Close - U.S.	Close Early	Friday, November 23, 2012	18:00:00	Saturday, November 24, 2012	05:00:00
Monday, December 24, 2012	Christmas Eve	Close Early	Monday, December 24, 2012	18:00:00	Tuesday, December 25, 2012	05:00:00
Tuesday, December 25, 2012	Christmas Day	Closed	Tuesday, December 25, 2012	05:00:00	Wednesday, December 26, 2012	05:00:00
Tuesday, January 1, 2013	New Year's Day	Closed	Tuesday, January 1, 2013	05:00:00	Wednesday, January 2, 2013	05:00:00
Tuesday, January 1, 2013	New Year's Day (Observed)	Closed	Tuesday, January 1, 2013	05:00:00	Wednesday, January 2, 2013	05:00:00
Monday, January 21, 2013	Martin Luther King, Jr. Day	Closed	Monday, January 21, 2013	05:00:00	Tuesday, January 22, 2013	05:00:00
Monday, February 18, 2013	President's Day - U.S.	Closed	Monday, February 18, 2013	05:00:00	Tuesday, February 19, 2013	05:00:00
Friday, March 29, 2013	Good Friday	Closed	Friday, March 29, 2013	04:00:00	Saturday, March 30, 2013	04:00:00
Monday, May 27, 2013	Memorial Day	Closed	Monday, May 27, 2013	04:00:00	Tuesday, May 28, 2013	04:00:00
Wednesday, July 3, 2013	Early Close-U.S.	Close 3 Hours Early	Wednesday, July 3, 2013	17:00:00	Thursday, July 4, 2013	04:00:00
Thursday, July 4, 2013	Independence Day - U.S.	Closed	Thursday, July 4, 2013	04:00:00	Friday, July 5, 2013	04:00:00
Monday, September 2, 2013	Labor Day - U.S.	Closed	Monday, September 2, 2013	04:00:00	Tuesday, September 3, 2013	04:00:00
Thursday, November 28, 2013	Thanksgiving Day - U.S.	Closed	Thursday, November 28, 2013	05:00:00	Friday, November 29, 2013	05:00:00
Friday, November 29, 2013	Early Close-U.S.	Close 3 Hours Early	Friday, November 29, 2013	18:00:00	Saturday, November 30, 2013	05:00:00
Tuesday, December 24, 2013	Christmas Eve	Close 3 Hours Early	Tuesday, December 24, 2013	18:00:00	Wednesday, December 25, 2013	05:00:00
Wednesday, December 25, 2013	Christmas Day	Closed	Wednesday, December 25, 2013	05:00:00	Thursday, December 26, 2013	05:00:00
Sunday, January 1, 2012	New Year's Day	Closed	NA	NA	NA	NA
Wednesday, January 1, 2014	New Year's Day (Observed)	Closed	Wednesday, January 1, 2014	05:00:00	Thursday, January 2, 2014	05:00:00
Monday, January 20, 2014	Martin Luther King, Jr. Day	Closed	Monday, January 20, 2014	05:00:00	Tuesday, January 21, 2014	05:00:00
Monday, February 17, 2014	President's Day - U.S.	Closed	Monday, February 17, 2014	05:00:00	Tuesday, February 18, 2014	05:00:00
Friday, April 18, 2014	Good Friday	Closed	Friday, April 18, 2014	04:00:00	Saturday, April 19, 2014	04:00:00
Friday, May 23, 2014	Early Close-U.S.	Close at 1:00PM	Friday, May 23, 2014	17:00:00	Saturday, May 24, 2014	04:00:00
Monday, May 26, 2014	Memorial Day -- U.S./ Spring Bank Holiday -- U.K.	Closed	Monday, May 26, 2014	04:00:00	Tuesday, May 27, 2014	04:00:00
Thursday, July 3, 2014	Early Close-U.S.	Close at 1:00PM	Thursday, July 3, 2014	17:00:00	Friday, July 4, 2014	04:00:00
Friday, July 4, 2014	Independence Day - U.S.	Closed	Friday, July 4, 2014	04:00:00	Saturday, July 5, 2014	04:00:00
Monday, September 1, 2014	Labor Day - U.S.	Closed	Monday, September 1, 2014	04:00:00	Tuesday, September 2, 2014	04:00:00
Thursday, November 27, 2014	Thanksgiving Day - U.S.	Closed	Thursday, November 27, 2014	05:00:00	Friday, November 28, 2014	05:00:00
Friday, November 28, 2014	Early Close-U.S.	Close at 1:00PM	Friday, November 28, 2014	18:00:00	Saturday, November 29, 2014	05:00:00
Wednesday, December 24, 2014	Christmas Eve	Close at 1:00PM	Wednesday, December 24, 2014	18:00:00	Thursday, December 25, 2014	05:00:00
Thursday, December 25, 2014	Christmas Day	Closed	Thursday, December 25, 2014	05:00:00	Friday, December 26, 2014	05:00:00
Thursday, January 1, 2015	New Year's Day	Closed	Thursday, January 1, 2015	05:00:00	Friday, January 2, 2015	05:00:00
Monday, January 19, 2015	Martin Luther King, Jr. Day	Closed	Monday, January 19, 2015	05:00:00	Tuesday, January 20, 2015	05:00:00
Monday, February 16, 2015	President's Day - U.S.	Closed	Monday, February 16, 2015	05:00:00	Tuesday, February 17, 2015	05:00:00
Friday, April 3, 2015	Good Friday	Closed	Friday, April 3, 2015	04:00:00	Saturday, April 4, 2015	04:00:00
Friday, May 22, 2015	Early Close-U.S.	Close Early	Friday, May 22, 2015	17:00:00	Saturday, May 23, 2015	04:00:00
Monday, May 25, 2015	Memorial Day -- U.S./ Spring Bank Holiday -- U.K.	Closed	Monday, May 25, 2015	04:00:00	Tuesday, May 26, 2015	04:00:00
Friday, July 3, 2015	Independence Day - U.S.(Observed)	Closed	Friday, July 3, 2015	04:00:00	Saturday, July 4, 2015	04:00:00
Saturday, July 4, 2015	Independence Day - U.S.	Closed	NA	NA	NA	NA
Monday, September 7, 2015	Labor Day - U.S.	Closed	Monday, September 7, 2015	04:00:00	Tuesday, September 8, 2015	04:00:00
Thursday, November 26, 2015	Thanksgiving Day - U.S.	Closed	Thursday, November 26, 2015	05:00:00	Friday, November 27, 2015	05:00:00
Friday, November 27, 2015	Early Close-U.S.	Close at 1:00PM	Friday, November 27, 2015	18:00:00	Saturday, November 28, 2015	05:00:00
Thursday, December 24, 2015	Christmas Eve	Close at 1:00PM	Thursday, December 24, 2015	18:00:00	Friday, December 25, 2015	05:00:00
Friday, December 25, 2015	Christmas Day	Closed	Friday, December 25, 2015	05:00:00	Saturday, December 26, 2015	05:00:00

Appendix E

Summary of Stock Split Data

AAPL Stock Split Stats:		Factor
Size of Split	7 for 1	7
Date of Split	2014-06-09	
Average 15-min Volume Before Split	297,743	
Average 15-min Volume After Split	853,234	2.87
Average 15-min Adjusted Volume After Split	121,891	0.41
Average 15-min Number of Trades Before Split	1,790	
Average 15-min Number of Trades After Split	4,322	2.41

ADP Stock Split Stats:		Factor
Size of Split	1139 for 1000	1.139
Date of Split	2014-10-01	
Average 15-min Volume Before Split	77,016	
Average 15-min Volume After Split	63,267	0.82
Average 15-min Adjusted Volume After Split	55,561	0.72
Average 15-min Number of Trades Before Split	498	
Average 15-min Number of Trades After Split	519	1.04

CELG Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2014-06-26	
Average 15-min Volume Before Split	99,583	
Average 15-min Volume After Split	122,903	1.23
Average 15-min Adjusted Volume After Split	61,461	0.62
Average 15-min Number of Trades Before Split	636	
Average 15-min Number of Trades After Split	931	1.46

CERN Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2011-06-27	
Average 15-min Volume Before Split	24,037	
Average 15-min Volume After Split	39,510	1.64
Average 15-min Adjusted Volume After Split	19,755	0.82
Average 15-min Number of Trades Before Split	165	
Average 15-min Number of Trades After Split	272	1.65

Appendix E — continued

Summary of Stock Split Data

CERN Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2013-07-01	
Average 15-min Volume Before Split	39,510	
Average 15-min Volume After Split	61,277	1.55
Average 15-min Adjusted Volume After Split	15,323	0.39
Average 15-min Number of Trades Before Split	272	
Average 15-min Number of Trades After Split	447	1.64

CTSH Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2014-03-10	
Average 15-min Volume Before Split	89,151	
Average 15-min Volume After Split	128,730	1.44
Average 15-min Adjusted Volume After Split	64,389	0.72
Average 15-min Number of Trades Before Split	569	
Average 15-min Number of Trades After Split	882	1.55

DLTR Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2012-06-27	
Average 15-min Volume Before Split	44,530	
Average 15-min Volume After Split	86,197	1.94
Average 15-min Adjusted Volume After Split	43,098	0.97
Average 15-min Number of Trades Before Split	288	
Average 15-min Number of Trades After Split	598	2.08

EBAY Stock Split Stats:		Factor
Size of Split	2376 for 1000	2.376
Date of Split	2015-07-20	
Average 15-min Volume Before Split	315,121	
Average 15-min Volume After Split	400,719	1.27
Average 15-min Adjusted Volume After Split	168,764	0.54
Average 15-min Number of Trades Before Split	1,560	
Average 15-min Number of Trades After Split	1,993	1.28

Appendix E — continued

Summary of Stock Split Data

FAST Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2011-05-23	
Average 15-min Volume Before Split	36,948	
Average 15-min Volume After Split	74,544	2.02
Average 15-min Adjusted Volume After Split	37,276	1.01
Average 15-min Number of Trades Before Split	246	
Average 15-min Number of Trades After Split	508	2.06

FISV Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2013-12-17	
Average 15-min Volume Before Split	31,252	
Average 15-min Volume After Split	38,537	1.23
Average 15-min Adjusted Volume After Split	19,271	0.62
Average 15-min Number of Trades Before Split	211	
Average 15-min Number of Trades After Split	314	1.49

GILD Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2013-01-28	
Average 15-min Volume Before Split	272,342	
Average 15-min Volume After Split	254,103	0.93
Average 15-min Adjusted Volume After Split	127,052	0.47
Average 15-min Number of Trades Before Split	1,300	
Average 15-min Number of Trades After Split	1,533	1.18

GOOGL Stock Split Stats:		Factor
Size of Split	1998 for 1000	1.998
Date of Split	2014-04-03	
Average 15-min Volume Before Split	61,577	
Average 15-min Volume After Split	45,116	0.73
Average 15-min Adjusted Volume After Split	22,583	0.37
Average 15-min Number of Trades Before Split	442	
Average 15-min Number of Trades After Split	563	1.27

Appendix E — continued

Summary of Stock Split Data

ROST Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2011-12-16	
Average 15-min Volume Before Split	52,942	
Average 15-min Volume After Split	58,918	1.11
Average 15-min Adjusted Volume After Split	29,461	0.56
Average 15-min Number of Trades Before Split	359	
Average 15-min Number of Trades After Split	432	1.20

ROST Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2015-06-12	
Average 15-min Volume Before Split	58,918	
Average 15-min Volume After Split	92,081	1.56
Average 15-min Adjusted Volume After Split	23,024	0.39
Average 15-min Number of Trades Before Split	432	
Average 15-min Number of Trades After Split	681	1.57

SBUX Stock Split Stats:		Factor
Size of Split	2 for 1	2
Date of Split	2015-04-09	
Average 15-min Volume Before Split	172,398	
Average 15-min Volume After Split	201,334	1.17
Average 15-min Adjusted Volume After Split	100,667	0.58
Average 15-min Number of Trades Before Split	974	
Average 15-min Number of Trades After Split	1,270	1.30

VOD Stock Split Stats:		Factor
Size of Split	4905 for 5000	0.981
Date of Split	2014-02-24	
Average 15-min Volume Before Split	230,550	
Average 15-min Volume After Split	127,297	0.55
Average 15-min Adjusted Volume After Split	129,801	0.56
Average 15-min Number of Trades Before Split	857	
Average 15-min Number of Trades After Split	615	0.72

Appendix F — *Additional Cumulative Abnormal Returns Results*

Mean and Median CARs vs SPX Index for Relevant Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Returns (CARs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively. Note that Long CARs are reported for (Net) Positive News and Short CARs for (Net) Negative News so as to respect the implied direction of the market reaction.

For this series of tests, Relevance is set to 1 (most relevant news), no threshold is set for Novelty (all novelty scores are included), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CARs are measured in basis points (bps) relative to the SPX index benchmark. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

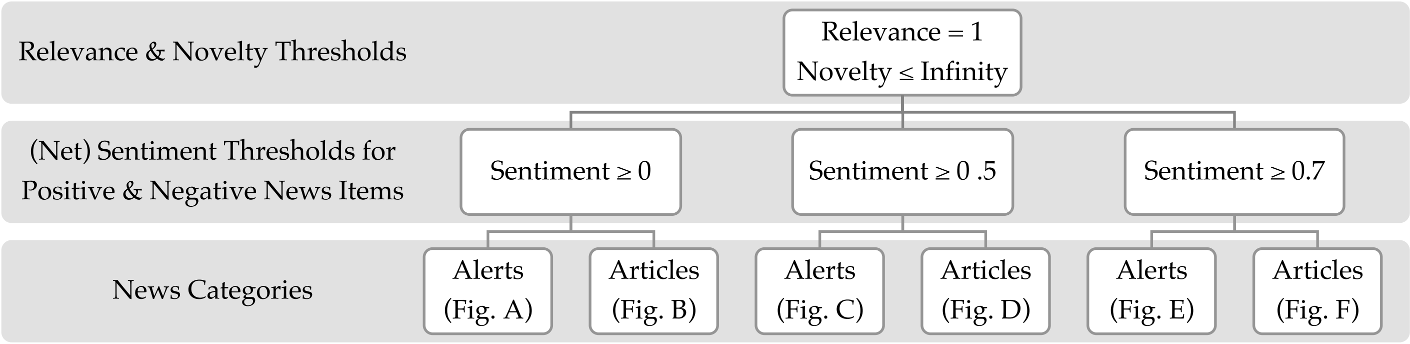
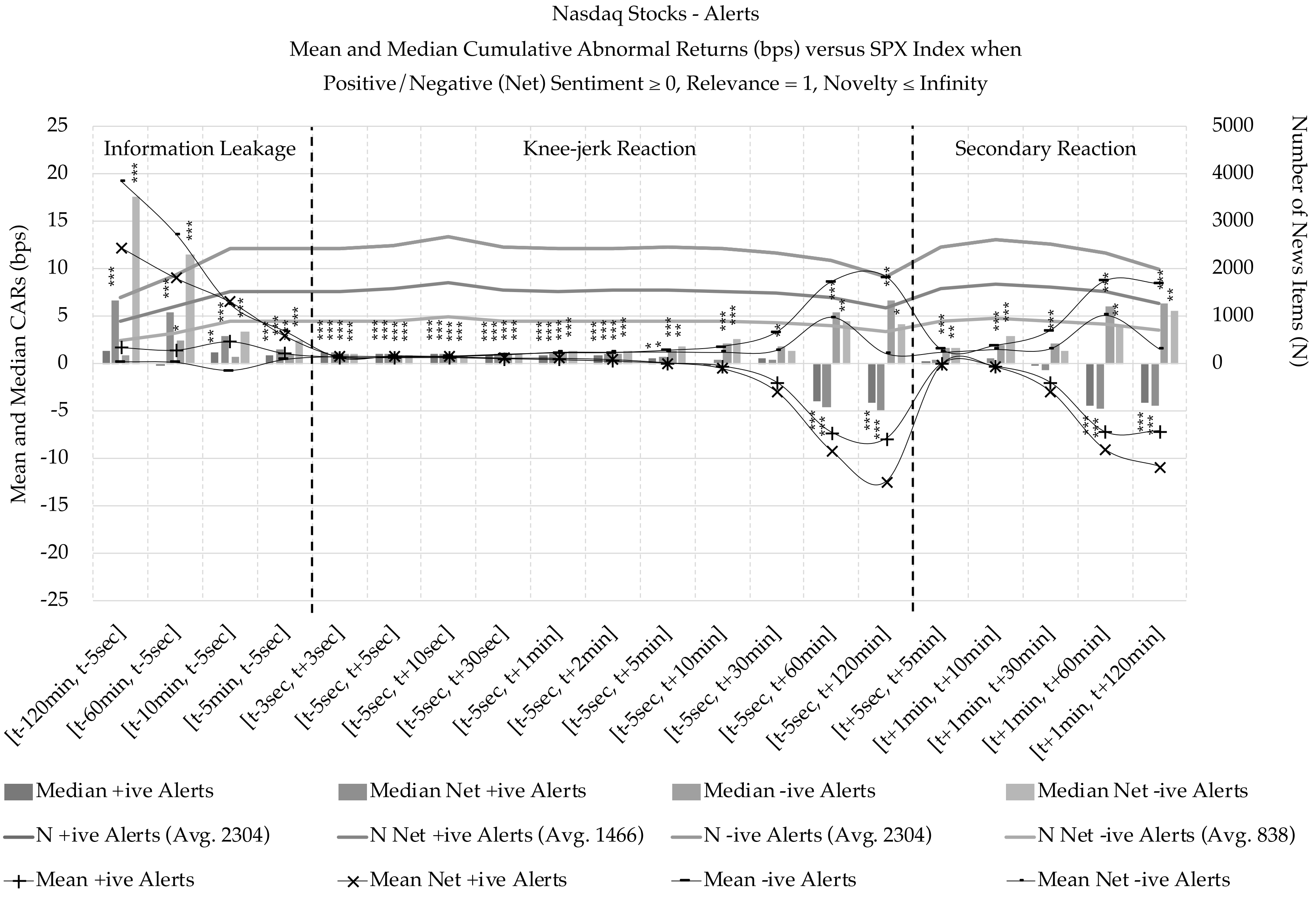


Figure A

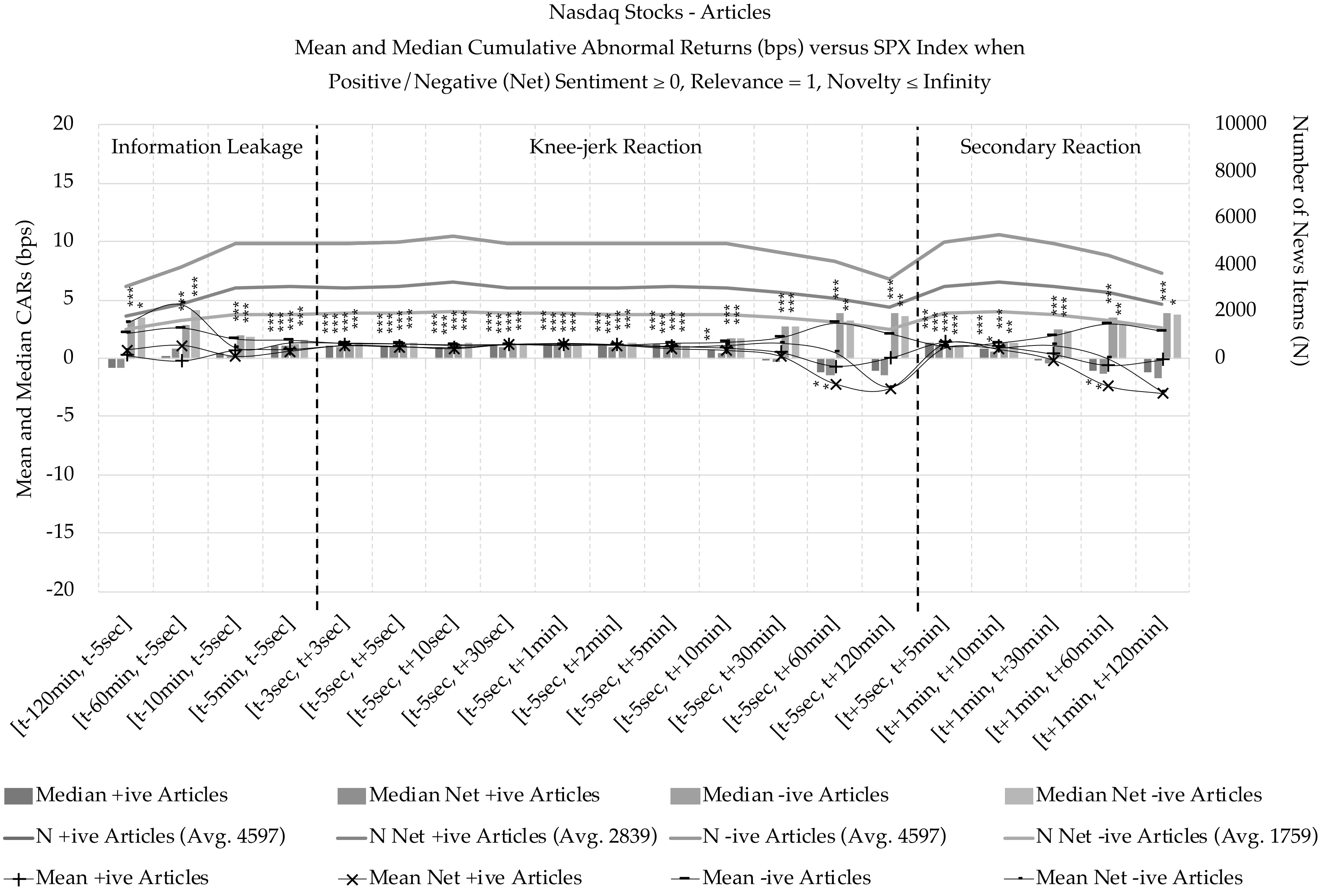
Mean and Median CARs vs SPX Index for Relevant Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

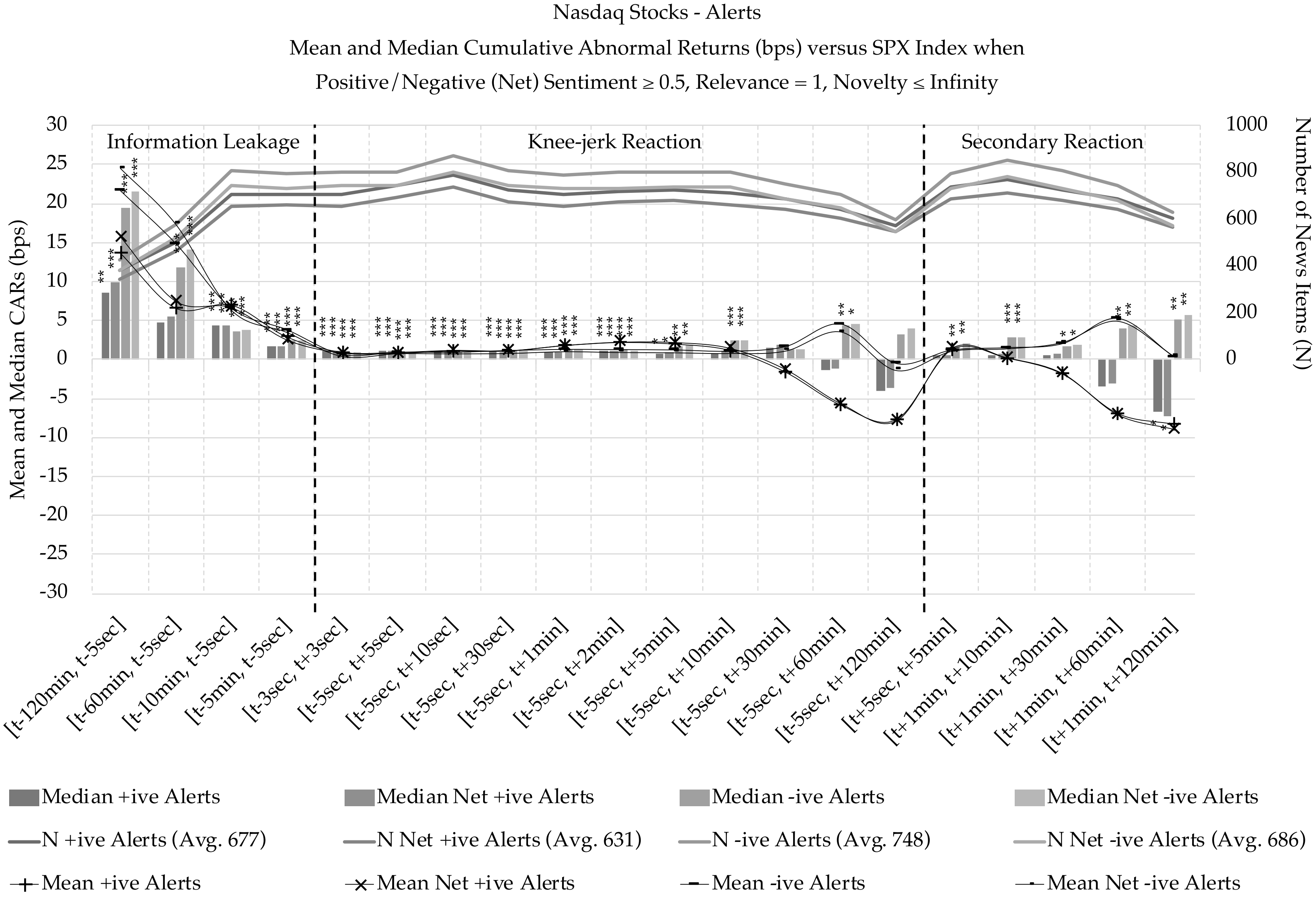
Mean and Median CARs vs SPX Index for Relevant Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

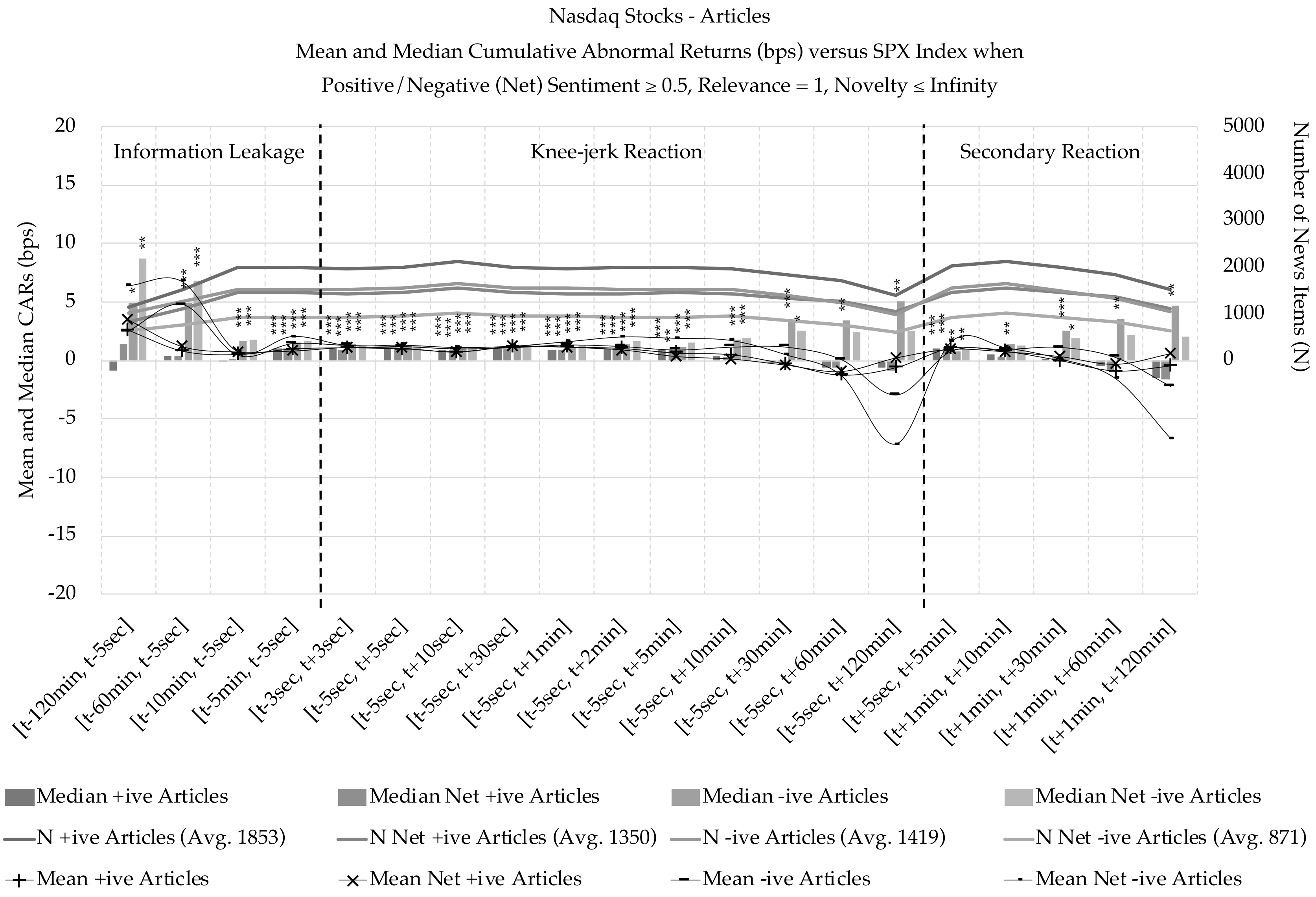
Mean and Median CARs vs SPX Index for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

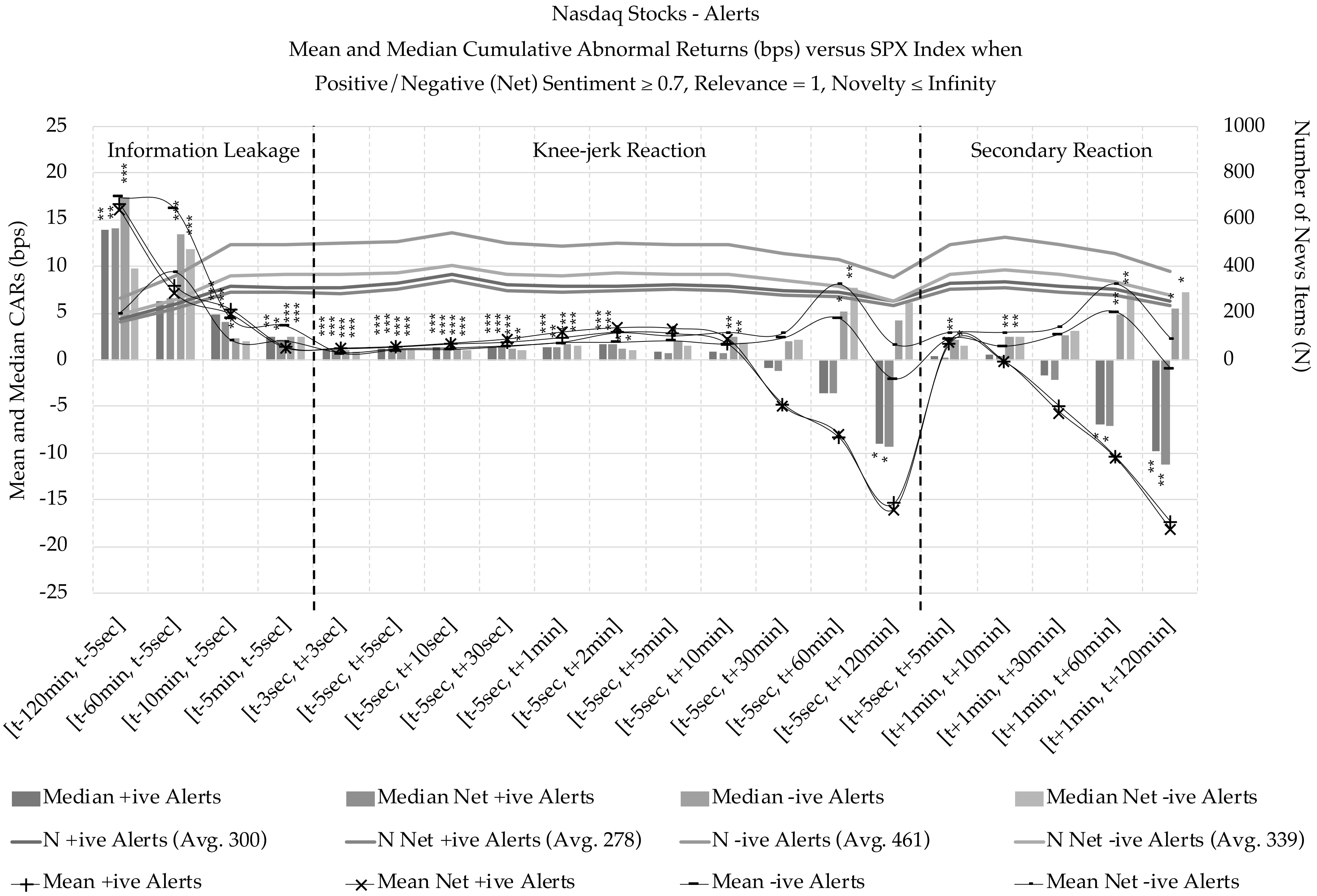
Mean and Median CARs vs SPX Index for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

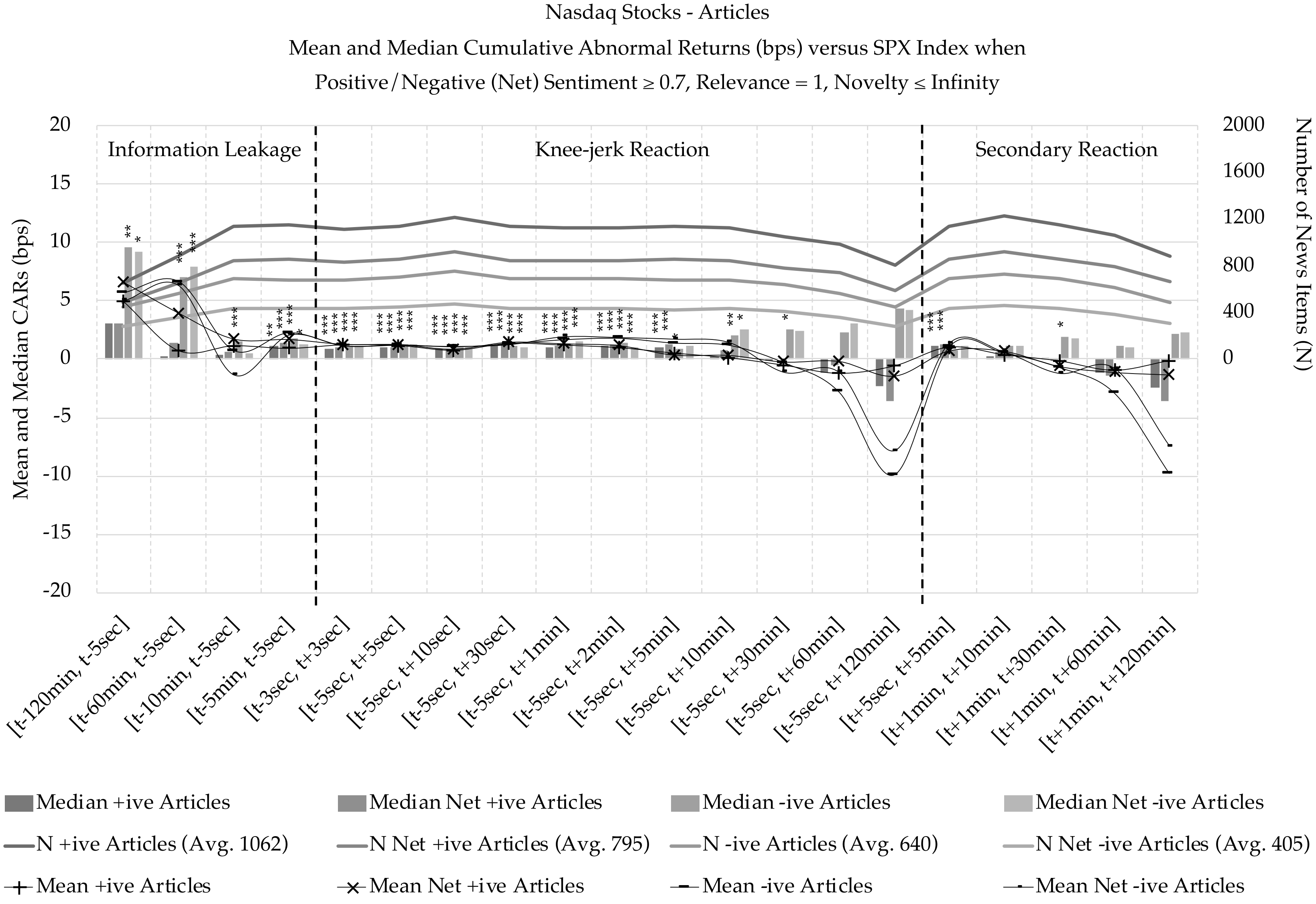
Mean and Median CARs vs SPX Index for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CARs vs SPX Index for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix F – continued — *Additional Cumulative Abnormal Returns Results*

Mean and Median CARs vs SPX Index for Novel Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Returns (CARs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively. Note that Long CARs are reported for (Net) Positive News and Short CARs for (Net) Negative News so as to respect the implied direction of the market reaction.

For this series of tests, no threshold is set for Relevance (all relevance scores are included), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CARs are measured in basis points (bps) relative to the SPX index benchmark. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

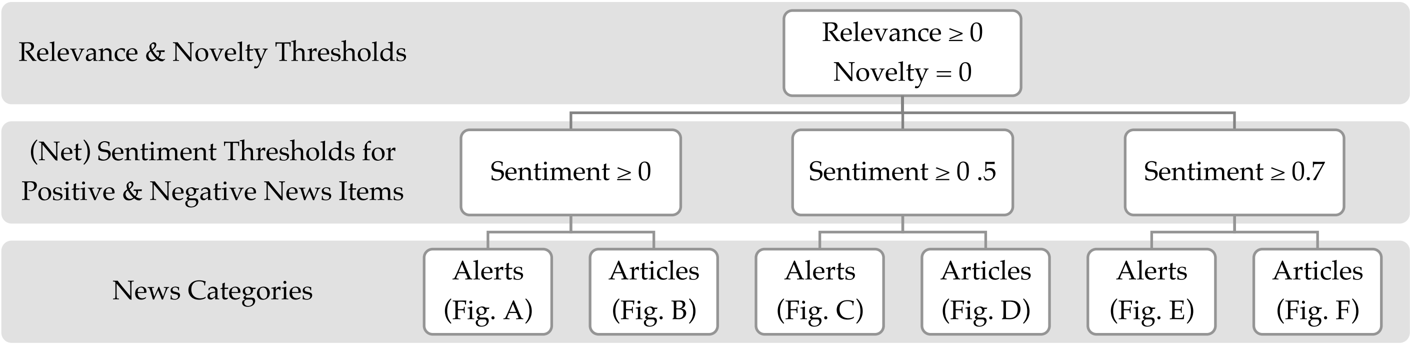
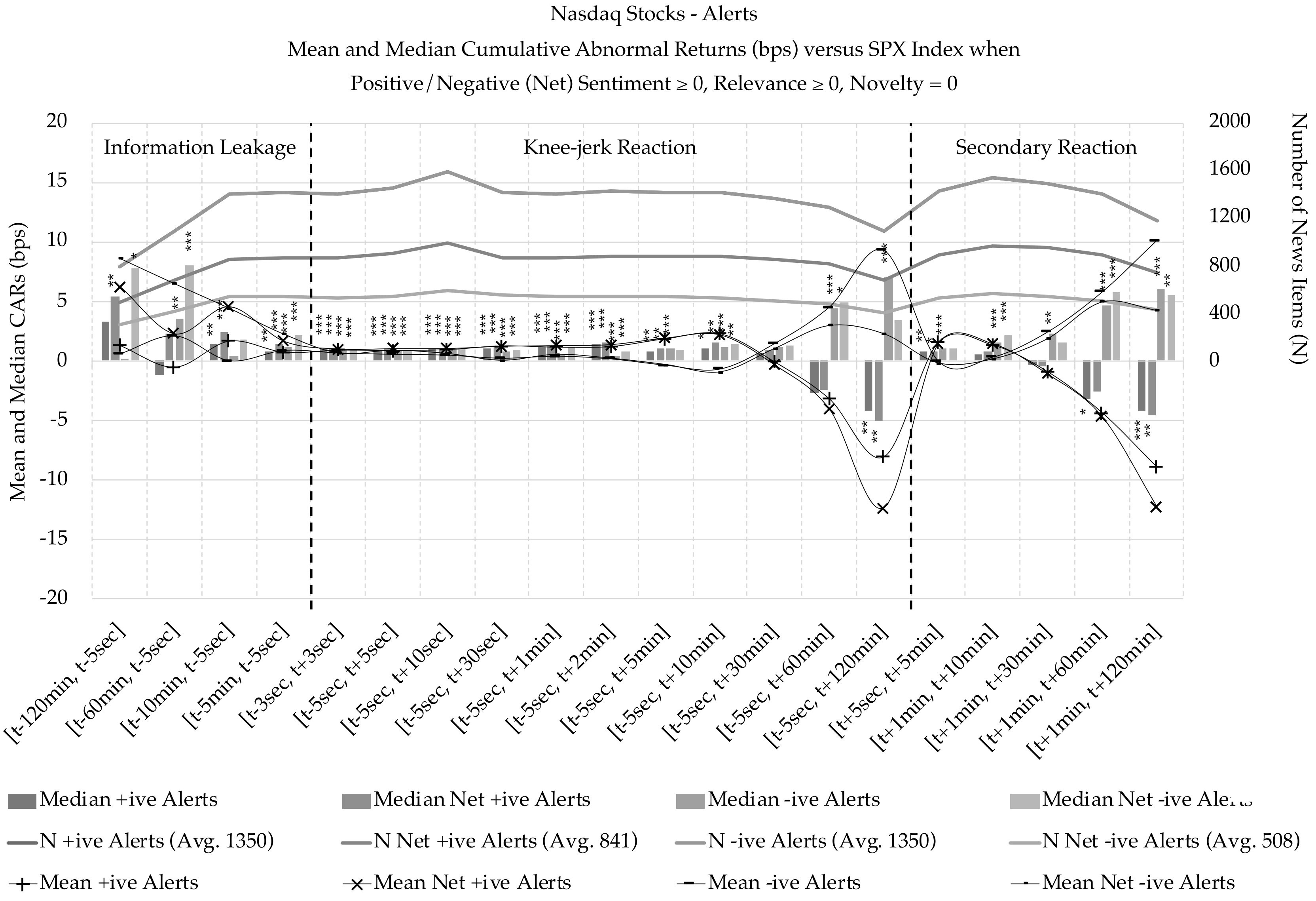


Figure A

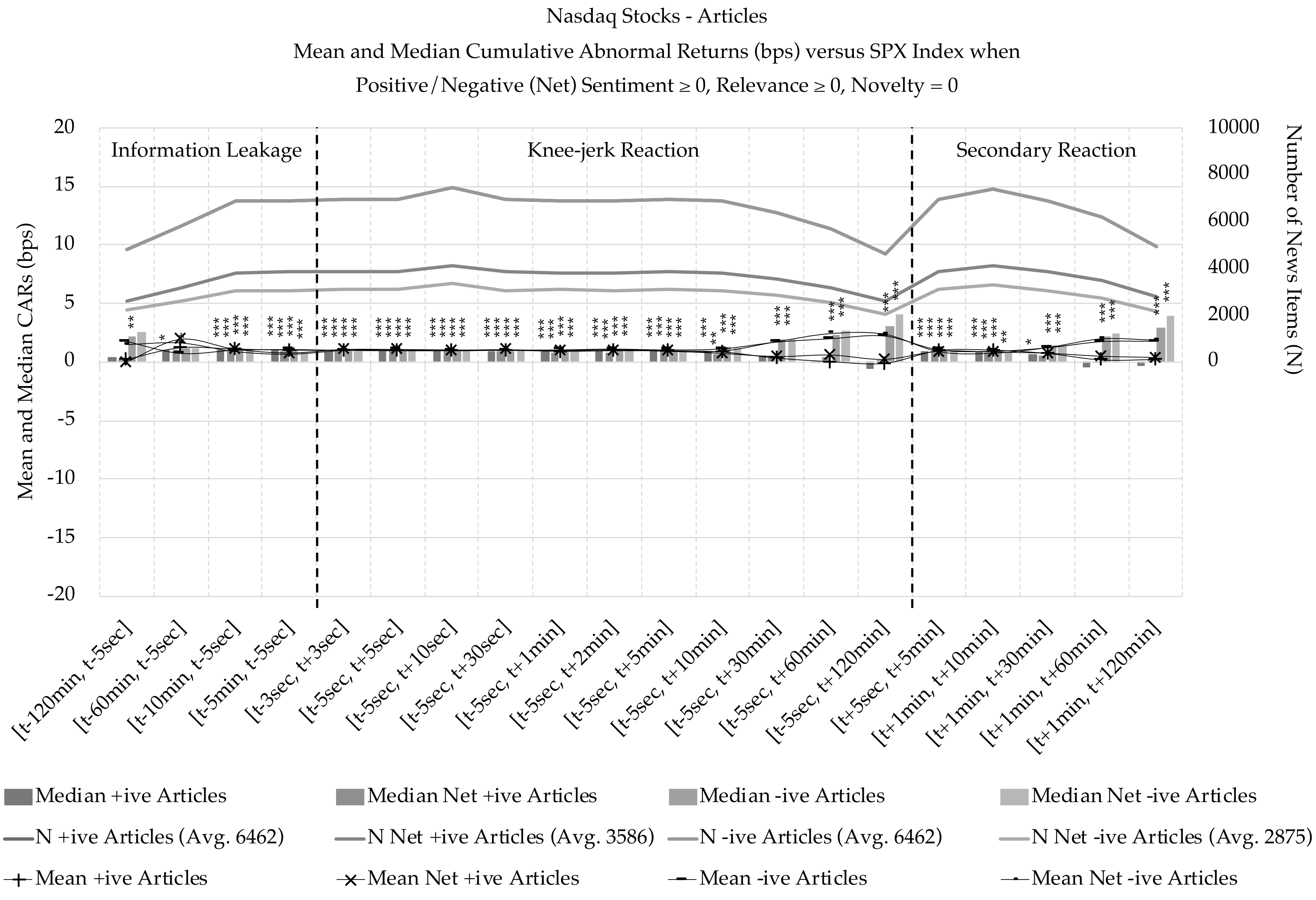
Mean and Median CARs vs SPX Index for Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

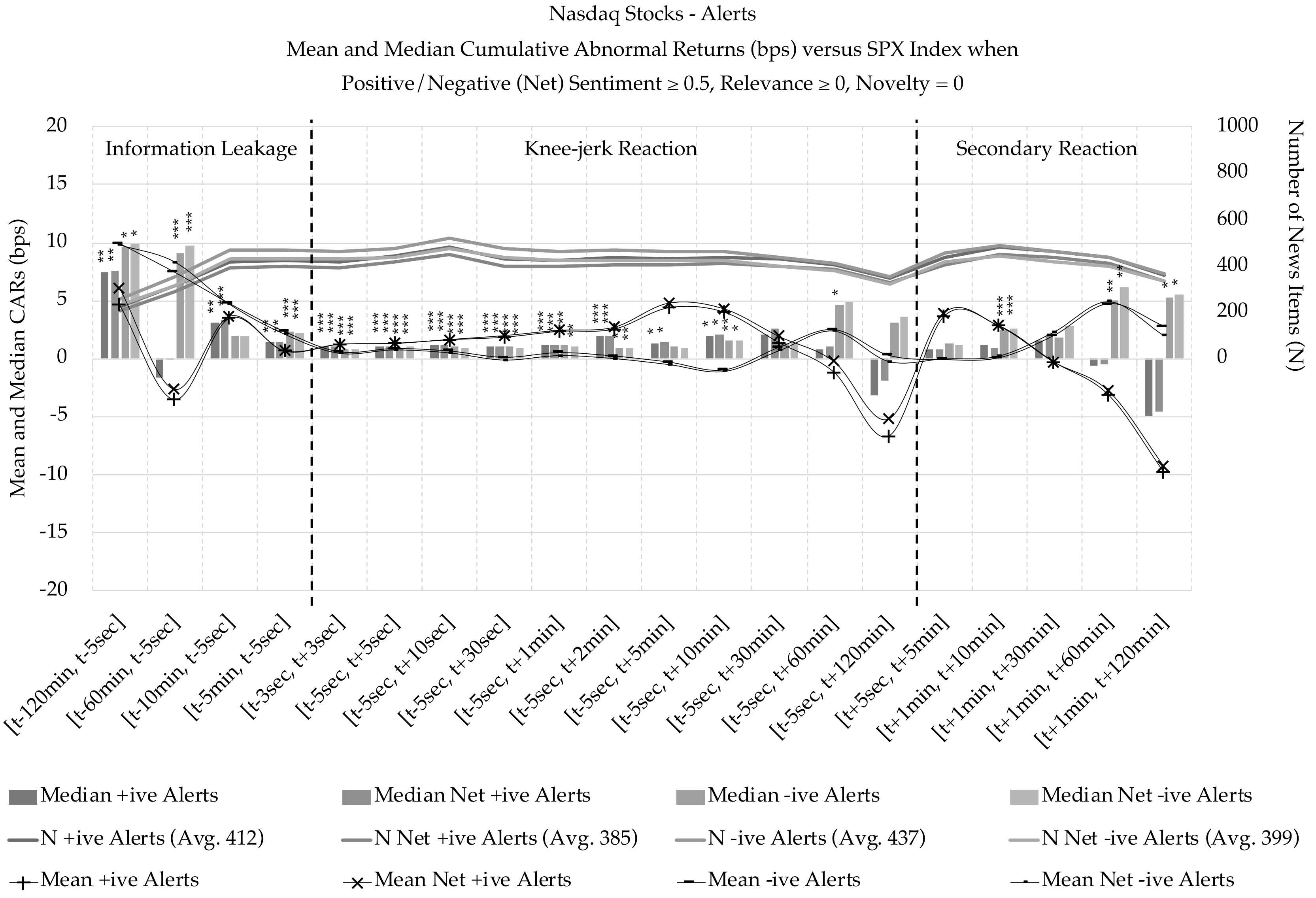
Mean and Median CARs vs SPX Index for Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

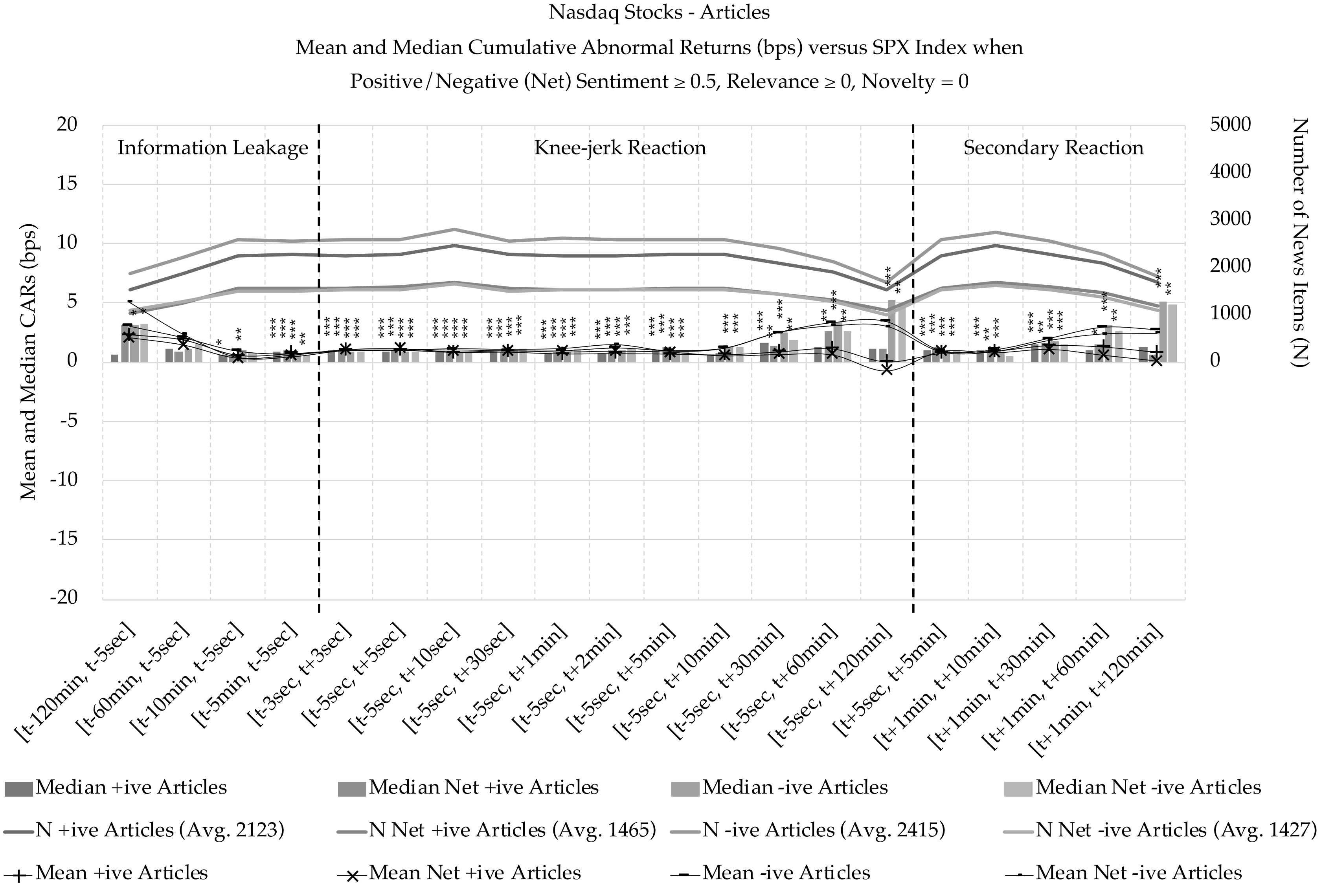
Mean and Median CARs vs SPX Index for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

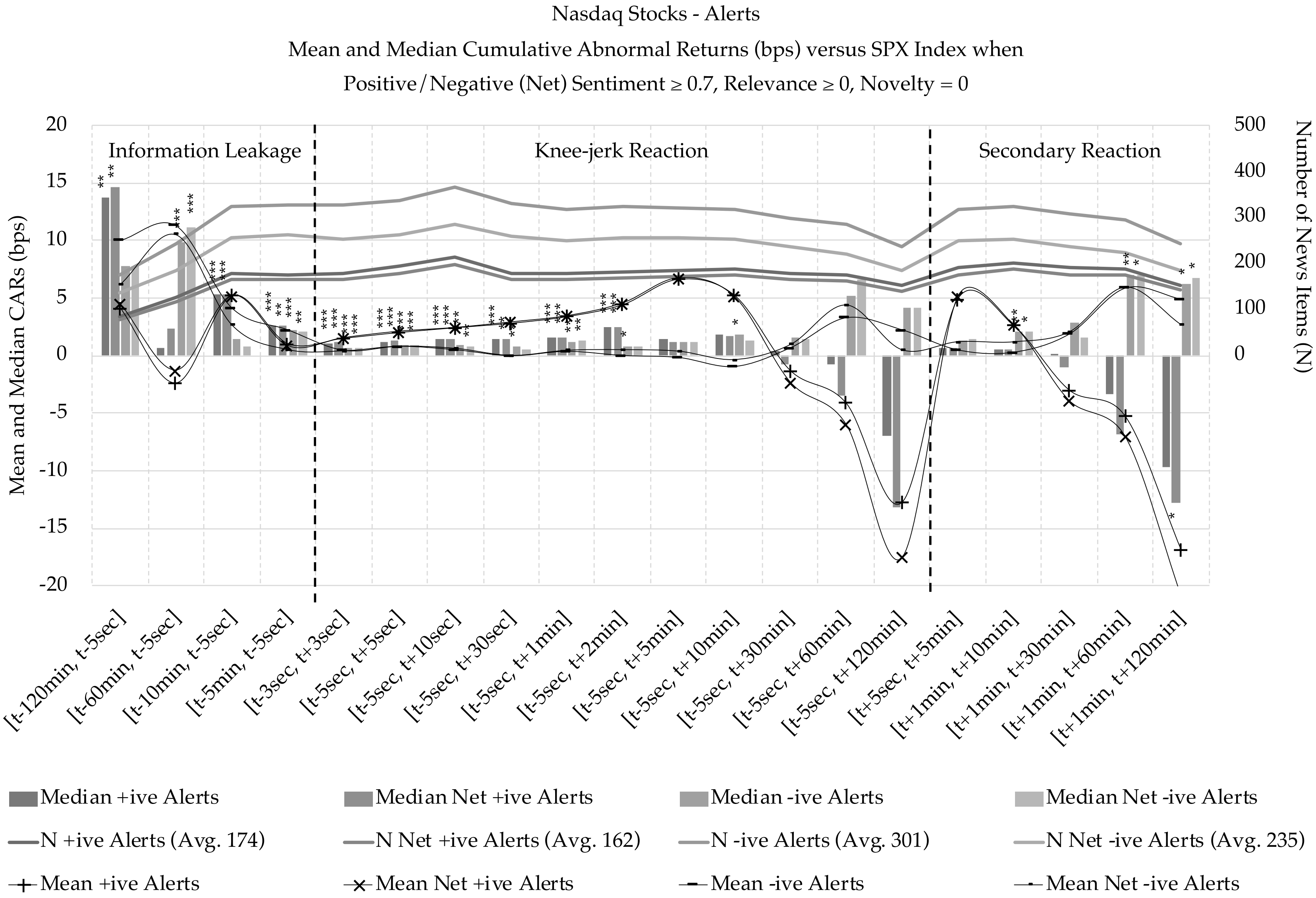
Mean and Median CARs vs SPX Index for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

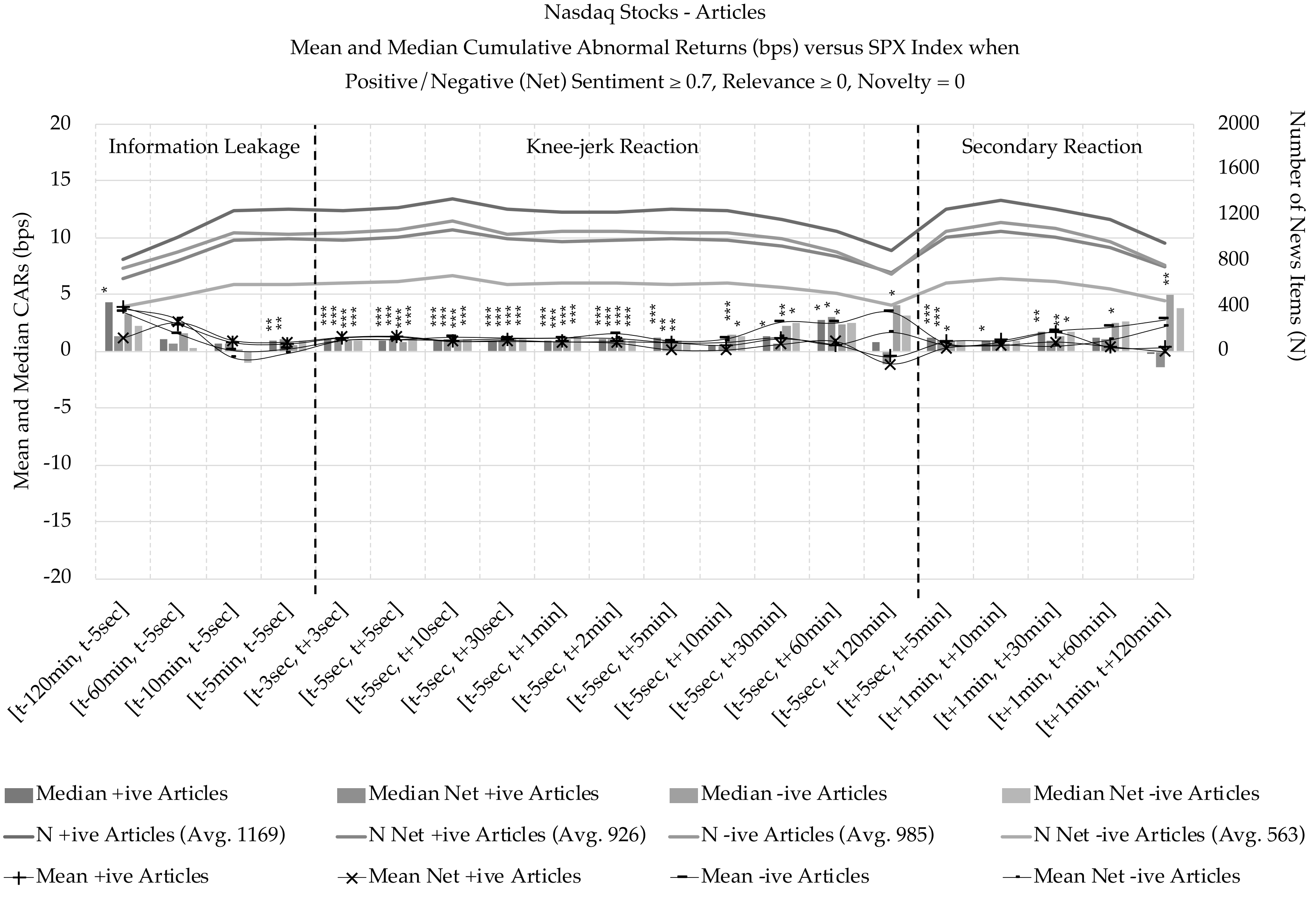
Mean and Median CARs vs SPX Index for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CARs vs SPX Index for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix F – continued — *Summary of CAR Results for Individual Stock Alerts*

Stock	Total N	% Sig. Windows	N Alerts	Alerts									
				Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CARs (bps)					
								[t0-120min, t0-5sec]		[t0-5sec, t0+1min]		[t0+1min, t0+120min]	
								Median	Average	Median	Average	Median	Average
AAPL	6498	76.98%	948	0.99	3.40	0.31	-0.60	8.81	14.38	2.14	1.28	-12.16	-21.42
CMCSA	1224	44.55%	181	0.95	0.56	0.33	-0.54	-0.61	-2.21	1.94	2.63	-3.78	-0.90
MSFT	2914	37.47%	351	0.95	1.19	0.32	-0.54	4.51	6.63	1.22	1.66	1.98	4.55
SBUX	610	34.41%	73	0.98	1.85	0.44	-0.63	-6.43	3.53	1.99	3.73	7.49	17.65
QCOM	523	33.51%	52	0.95	0.54	0.36	-0.65	37.46	27.37	0.91	0.65	-6.64	-8.26
EBAY	1020	33.23%	139	0.98	1.24	0.35	-0.51	12.28	27.42	2.45	-0.41	1.05	-11.93
MU	560	29.03%	72	0.97	4.07	0.39	-0.63	30.82	-5.57	-1.64	-1.03	77.28	31.01
GILD	537	28.96%	109	0.98	1.28	0.35	-0.60	18.15	19.26	2.03	1.86	6.78	15.14
ADP	492	27.78%	28	0.95	0.29	0.49	-0.46	18.99	18.99	4.32	3.36	61.42	61.42
SNDK	271	26.04%	63	0.97	2.46	0.48	-0.64	1.37	7.49	1.57	2.59	17.98	18.50
PCLN	290	25.73%	48	0.97	3.63	0.35	-0.66	9.12	0.32	-2.63	-1.24	-31.06	-27.12
PAYX	158	25.17%	29	0.96	0.59	0.41	-0.48	27.03	20.17	2.60	2.16	47.66	33.31
AMAT	270	24.76%	37	0.99	0.95	0.52	-0.64	54.71	41.68	-3.18	-2.01	-35.15	-32.06
COST	482	23.09%	32	0.95	0.78	0.47	-0.60	25.49	25.82	1.05	1.03	11.00	2.39
NVDA	348	22.99%	53	0.96	2.55	0.39	-0.64	-30.03	-18.13	-2.30	-1.12	0.33	5.15
INTC	1630	22.78%	239	0.96	2.38	0.38	-0.62	6.92	2.35	0.30	0.91	0.80	-2.40
VOD	1364	21.35%	171	0.91	0.47	0.42	-0.55	4.07	-5.52	0.31	0.29	-13.01	-7.62
AMZN	2381	20.31%	291	0.96	1.32	0.34	-0.60	-0.51	-8.65	-0.81	-1.07	6.94	10.32
GOOGL	289	19.06%	82	0.99	0.71	0.30	-0.45	-2.81	3.49	-3.34	-2.69	-0.18	3.44
BIIB	426	18.89%	75	0.97	0.96	0.47	-0.61	23.26	65.49	-3.13	-5.20	17.33	8.83
CHKP	68	18.54%	14	1.00	0.07	0.58	N/A	N/A	N/A	9.79	9.44	14.49	10.52
CSCO	812	17.60%	105	0.99	1.59	0.37	-0.61	12.84	10.57	0.67	0.93	5.91	8.96
ADBE	332	16.56%	48	0.99	2.65	0.26	-0.62	32.57	26.41	1.98	0.81	-1.77	13.08
SYMC	213	15.14%	32	0.89	0.19	0.55	-0.62	8.41	21.78	-0.38	-1.18	-1.54	1.99
ESRX	210	14.72%	17	0.95	0.76	0.45	-0.60	34.57	34.57	0.61	3.60	6.51	-22.62
YHOO	1278	14.69%	175	0.97	1.21	0.33	-0.62	29.43	60.25	-0.05	0.13	-14.01	-66.90
BBBY	242	14.55%	26	0.99	6.92	0.51	-0.70	-31.27	-71.34	0.74	-0.22	11.95	29.86
DLTR	240	13.58%	47	0.97	2.36	0.46	-0.57	-21.36	-25.00	0.23	-0.67	27.69	35.94

This table reports a summary of the median and average CARs for individual stocks for three event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAR Results for Individual Stocks — continued

Stock	Total N	% Sig. Windows	Alerts					CARs (bps)					
			N Alerts	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	[t0-120min, t0-5sec]		[t0-5sec, t0+1min]		[t0+1min, t0+120min]	
								Median	Average	Median	Average	Median	Average
CTSH	156	12.40%	39	0.95	0.64	0.50	-0.44	44.19	34.13	1.12	0.06	-23.74	-26.52
ADSK	248	12.26%	38	0.98	1.84	0.59	-0.61	-21.03	-29.29	3.53	4.89	-8.52	-40.02
ROST	202	11.91%	28	0.95	4.57	0.56	-0.71	12.00	53.13	-0.09	2.85	31.77	34.82
ATVI	303	11.67%	29	0.99	0.69	0.35	-0.65	60.71	38.17	1.46	4.00	29.09	41.19
AMGN	472	11.53%	101	0.95	0.64	0.38	-0.58	-12.19	5.34	2.47	1.81	-2.51	-4.93
MYL	562	10.73%	74	0.99	1.73	0.44	-0.53	33.92	25.83	-0.03	4.37	-2.81	-28.49
NTAP	141	10.66%	14	1.00	1.57	0.36	-0.61	-28.81	-13.42	1.30	3.33	-19.87	-12.30
STX	217	10.21%	58	0.98	1.81	0.46	-0.62	-0.69	2.72	-1.45	-4.08	-7.91	-17.72
	220	10.21%	52	1.00	2.12	0.46	-0.63	156.27	111.55	3.60	-0.07	-33.60	-51.13
CTXS	194	10.14%	30	0.96	2.83	0.61	-0.64	121.63	34.26	0.16	-1.40	-126.59	-48.89
BIDU	256	9.90%	35	0.98	0.94	0.34	-0.60	31.56	32.95	-7.79	-6.17	-20.79	-20.95
CELG	507	9.83%	133	1.00	1.83	0.31	-0.50	8.68	29.71	-2.44	-1.99	24.10	13.57
ISRG	87	9.17%	6	1.00	1.67	0.75	-0.60	N/A	N/A	N/A	N/A	-51.43	-51.43
INTU	149	8.99%	20	1.00	2.70	0.53	-0.66	3.90	3.62	-7.78	-5.06	19.16	17.29
LLTC	55	8.85%	27	0.87	0.89	0.52	-0.70	18.66	22.84	2.52	3.25	-38.73	-30.61
PCAR	128	8.37%	40	0.97	1.08	0.52	-0.72	39.33	30.74	2.47	5.30	-8.37	8.06
XLNX	60	8.37%	9	1.00	0.67	0.49	-0.68	40.98	38.29	-1.04	-1.40	-17.03	-21.78
MAT	205	7.22%	34	1.00	0.91	0.41	-0.58	-11.30	-28.43	0.43	4.24	-12.18	8.85
CA	110	6.42%	7	1.00	0.00	0.24	-0.76	42.06	42.06	1.44	-4.99	8.62	-33.55
KLAC	96	6.35%	22	0.96	1.32	0.52	-0.61	30.95	19.87	3.74	5.43	2.40	-2.92
FAST	55	6.18%	14	1.00	0.36	0.21	-0.74	-10.50	-8.79	1.64	2.97	-5.49	0.53
AKAM	99	6.08%	13	1.00	0.54	0.53	-0.68	80.83	80.83	2.39	3.45	-12.72	-17.45
CERN	70	5.76%	8	0.96	0.88	0.61	-0.61	10.47	10.47	4.23	4.23	-42.95	-62.58
FISV	117	5.52%	17	0.93	0.41	0.48	-0.38	58.58	58.58	-6.25	-6.25	5.05	-0.92
ORLY	52	5.14%	8	1.00	5.00	0.52	-0.59	-15.14	6.48	1.39	1.39	-40.65	-40.65
SRCL	24	5.03%	10	0.96	0.40	0.48	-0.53	-8.83	-1.49	N/A	N/A	N/A	N/A
HSIC	28	5.00%	1	1.00	0.00	N/A	-0.51	N/A	N/A	N/A	N/A	N/A	N/A

This table reports a summary of the median and average CARs for individual stocks for three event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAR Results for Individual Stocks — continued

			Articles										
Stock	Total N	% Sig. Windows	N Articles	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CARs (bps)					
								[t0-120min, t0-5sec]		[t0-5sec, t0+1min]		[t0+1min, t0+120min]	
								Median	Average	Median	Average	Median	Average
AAPL	6498	76.98%	5331	0.53	6.72	0.40	-0.49	1.90	1.52	2.10	2.14	2.10	0.26
CMCSA	1224	44.55%	993	0.52	1.33	0.49	-0.43	1.75	3.95	1.62	2.31	1.62	4.84
MSFT	2914	37.47%	2463	0.42	2.70	0.43	-0.48	-1.14	-0.13	0.77	1.26	0.77	2.13
SBUX	610	34.41%	526	0.46	1.94	0.48	-0.54	8.78	3.04	1.82	2.05	1.82	-1.85
QCOM	523	33.51%	454	0.39	1.27	0.47	-0.48	6.19	4.56	1.96	2.25	1.96	2.83
EBAY	1020	33.23%	857	0.42	2.12	0.46	-0.47	4.04	4.56	1.82	2.55	1.82	-0.48
MU	560	29.03%	476	0.42	3.47	0.48	-0.55	-1.40	-4.97	-5.34	-4.12	-5.34	-4.81
GILD	537	28.96%	411	0.43	1.93	0.45	-0.45	5.50	3.06	2.02	3.07	2.02	4.64
ADP	492	27.78%	336	0.49	1.12	0.48	-0.51	-5.86	-1.96	2.09	2.23	2.09	3.01
SNDK	271	26.04%	205	0.46	2.31	0.50	-0.54	11.52	21.51	1.88	3.30	1.88	-0.57
PCLN	290	25.73%	236	0.44	2.60	0.48	-0.49	-20.40	-22.57	-4.15	-4.36	-4.15	-21.43
PAYX	158	25.17%	129	0.76	0.99	0.52	-0.48	8.10	-4.53	1.29	1.51	1.29	10.49
AMAT	270	24.76%	220	0.36	4.24	0.49	-0.58	6.04	4.13	-2.20	-2.75	-2.20	7.11
COST	482	23.09%	407	0.33	2.51	0.44	-0.46	10.73	6.69	1.79	1.69	1.79	0.61
NVDA	348	22.99%	274	0.40	3.04	0.46	-0.53	21.42	12.34	-2.41	-1.62	-2.41	9.98
INTC	1630	22.78%	1299	0.49	2.36	0.47	-0.53	-1.06	1.91	-0.10	0.13	-0.10	0.93
VOD	1364	21.35%	1151	0.36	2.27	0.36	-0.47	2.95	1.11	0.72	1.17	0.72	3.44
AMZN	2381	20.31%	2008	0.43	3.37	0.44	-0.46	0.25	-0.35	-0.40	-0.27	-0.40	1.99
GOOGL	289	19.06%	207	0.50	2.10	0.35	-0.49	7.96	7.79	-1.72	-2.40	-1.72	-6.69
BIIB	426	18.89%	337	0.32	2.74	0.50	-0.42	3.17	12.52	-3.09	-4.18	-3.09	3.23
CHKP	68	18.54%	53	0.44	1.40	0.52	-0.49	0.15	6.45	1.01	1.95	1.01	-6.87
CSCO	812	17.60%	667	0.44	1.76	0.47	-0.56	-1.14	0.55	-0.49	-0.10	-0.49	1.08
ADBE	332	16.56%	274	0.48	3.73	0.48	-0.51	6.29	6.56	1.45	1.26	1.45	0.11
SYMC	213	15.14%	170	0.31	1.33	0.37	-0.45	-18.53	-20.14	-0.92	0.63	-0.92	-19.68
ESRX	210	14.72%	187	0.50	2.19	0.42	-0.41	-4.10	0.03	1.40	1.68	1.40	27.15
YHOO	1278	14.69%	1070	0.44	2.70	0.40	-0.47	-1.04	-1.58	-0.19	0.38	-0.19	-2.91
BBBY	242	14.55%	210	0.33	4.03	0.47	-0.55	2.63	-6.29	2.70	2.06	2.70	-5.90
DLTR	240	13.58%	193	0.34	3.30	0.53	-0.46	-11.91	0.62	0.93	0.95	0.93	1.95

This table reports a summary of the median and average CARs for individual stocks for three event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAR Results for Individual Stocks — continued

			Articles										
								CARs (bps)					
Stock	Total N	% Sig. Windows	N Articles	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	[t0-120min, t0-5sec]		[t0-5sec, t0+1min]		[t0+1min, t0+120min]	
								Median	Average	Median	Average	Median	Average
CTSH	156	12.40%	113	0.43	2.50	0.57	-0.46	12.24	6.04	1.03	1.05	1.03	-3.18
ADSK	248	12.26%	202	0.33	4.48	0.55	-0.47	21.16	5.62	0.66	-0.09	0.66	-9.86
ROST	202	11.91%	162	0.34	3.07	0.46	-0.39	12.46	22.84	0.76	0.98	0.76	-9.49
ATVI	303	11.67%	261	0.37	1.93	0.41	-0.44	-4.32	-2.67	-1.90	-2.79	-1.90	1.57
AMGN	472	11.53%	350	0.41	2.41	0.45	-0.43	2.60	10.13	0.47	0.37	0.47	-8.19
MYL	562	10.73%	484	0.90	30.10	0.33	-0.45	-22.37	-8.33	1.38	1.27	1.38	-17.39
NTAP	141	10.66%	119	0.43	3.63	0.48	-0.41	30.88	32.52	0.82	0.73	0.82	4.20
STX	217	10.21%	153	0.48	2.01	0.54	-0.48	12.39	10.39	1.78	0.09	1.78	-15.52
VRTX	220	10.21%	164	0.54	1.74	0.43	-0.51	41.20	25.02	-5.38	-8.43	-5.38	-42.52
CTXS	194	10.14%	155	0.39	2.86	0.50	-0.44	6.01	4.33	3.17	3.12	3.17	-5.77
BIDU	256	9.90%	208	0.39	1.81	0.45	-0.45	-0.08	4.35	-3.33	-2.62	-3.33	-4.49
CELG	507	9.83%	364	0.39	2.82	0.43	-0.44	-6.43	0.58	0.49	1.18	0.49	-7.50
ISRG	87	9.17%	81	0.36	2.99	0.64	-0.48	-15.00	-0.78	-6.11	-5.06	-6.11	-22.84
INTU	149	8.99%	123	0.34	2.03	0.50	-0.56	5.64	-12.97	1.03	1.61	1.03	7.49
LLTC	55	8.85%	28	0.41	2.14	0.51	-0.47	10.00	17.75	0.33	0.51	0.33	5.60
PCAR	128	8.37%	83	0.58	1.07	0.42	-0.55	0.52	12.68	1.76	0.63	1.76	-7.05
XLNX	60	8.37%	49	0.41	2.63	0.47	-0.46	4.39	5.69	1.37	0.96	1.37	4.05
MAT	205	7.22%	164	0.55	0.88	0.41	-0.34	-5.48	-14.82	1.28	2.64	1.28	5.01
CA	110	6.42%	96	0.55	1.28	0.54	-0.49	-7.26	-4.69	-0.02	1.31	-0.02	-2.81
KLAC	96	6.35%	70	0.34	2.57	0.43	-0.41	22.20	17.70	1.59	2.75	1.59	-8.60
FAST	55	6.18%	41	0.37	2.22	0.48	-0.52	-8.80	1.32	1.02	1.06	1.02	11.24
AKAM	99	6.08%	83	0.48	2.57	0.59	-0.46	1.62	-10.72	0.65	0.87	0.65	-12.94
CERN	70	5.76%	62	0.35	2.68	0.55	-0.59	25.06	12.21	-0.48	-2.33	-0.48	11.68
FISV	117	5.52%	98	0.71	1.53	0.61	-0.50	-0.56	1.37	1.46	3.83	1.46	2.05
ORLY	52	5.14%	41	0.42	2.85	0.54	-0.55	-5.06	-6.43	-0.39	-2.34	-0.39	-28.10
SRCL	24	5.03%	14	0.41	2.64	0.55	-0.43	-43.20	-43.20	-8.84	-8.84	-8.84	10.79
HSIC	28	5.00%	27	0.68	0.93	0.71	-0.31	N/A	N/A	-4.61	-4.61	-4.61	-28.02

This table reports a summary of the median and average CARs for individual stocks for three event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Appendix G — *Additional Cumulative Abnormal Volume Results*

Mean and Median CAVs for Novel Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Volumes (CAVs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, no threshold is set for Relevance (all relevance scores are included), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

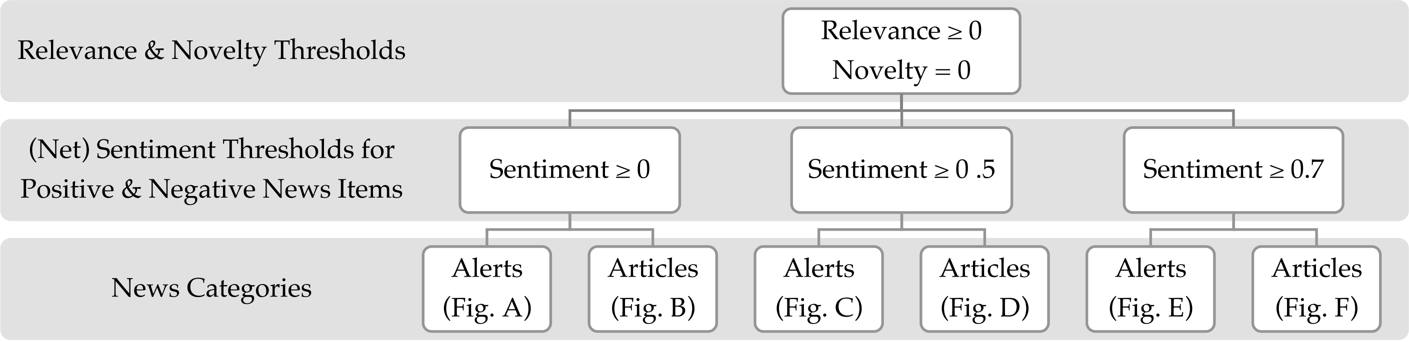
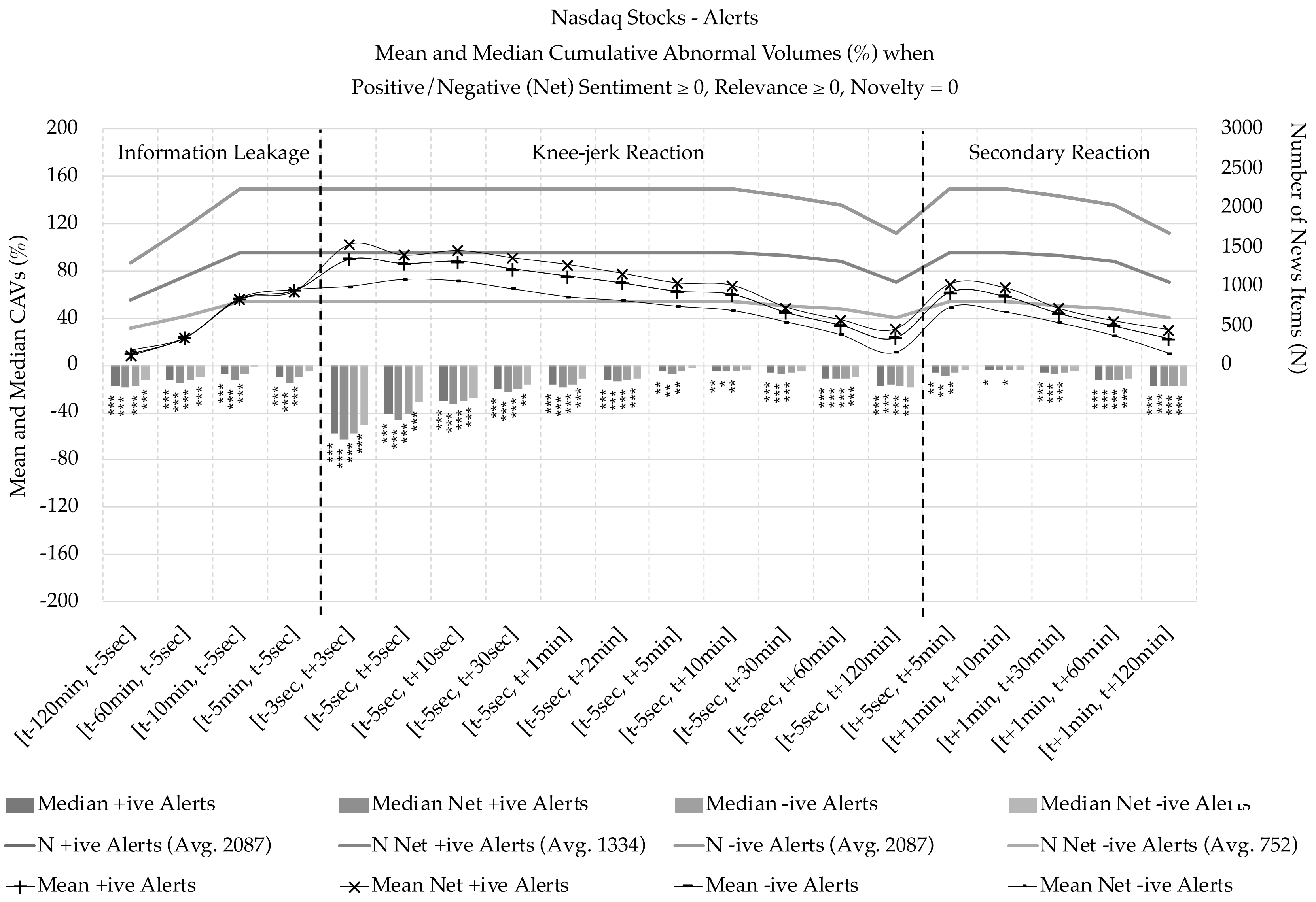


Figure A

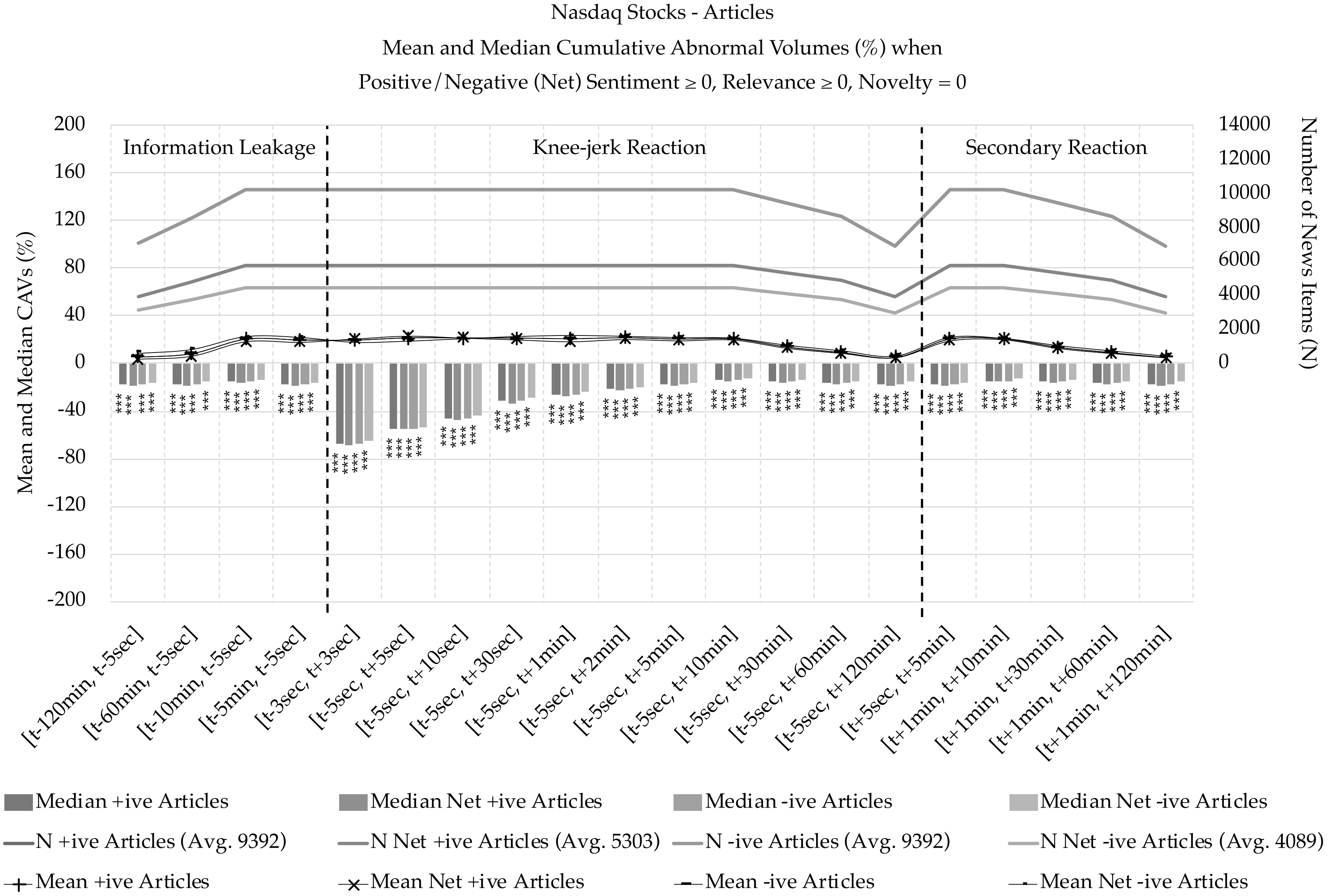
Mean and Median CAVs for Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

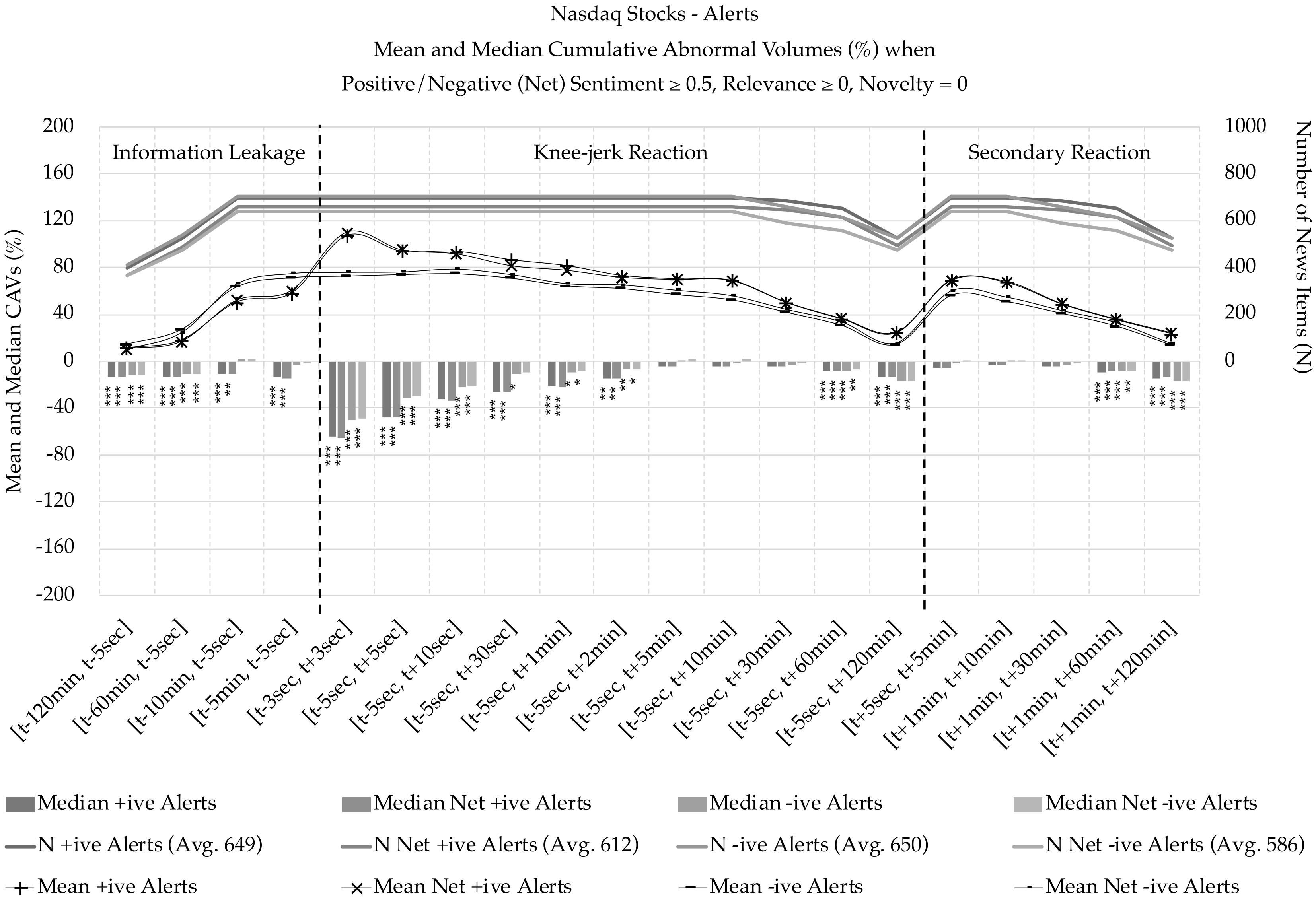
Mean and Median CAVs for Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

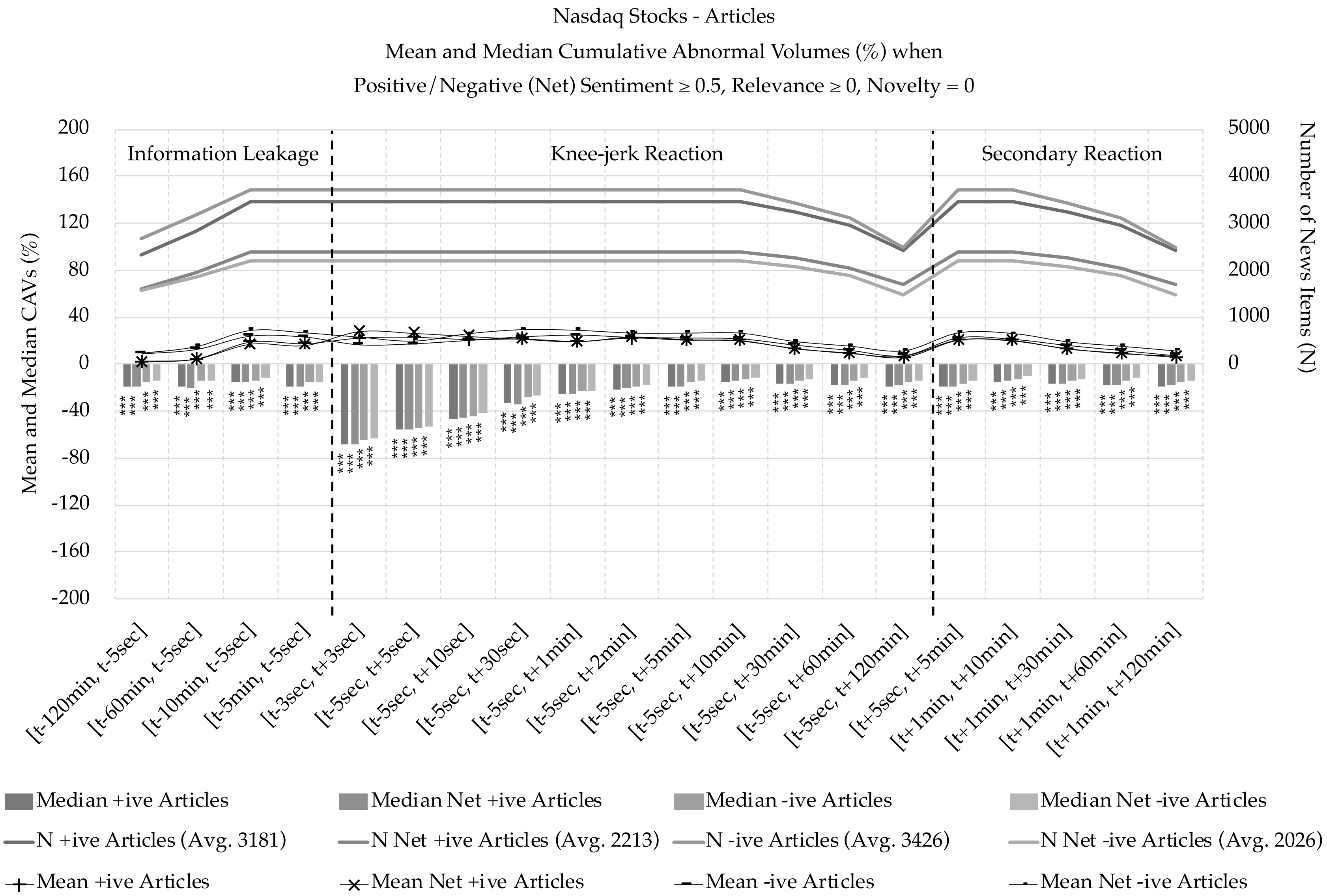
Mean and Median CAVs for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

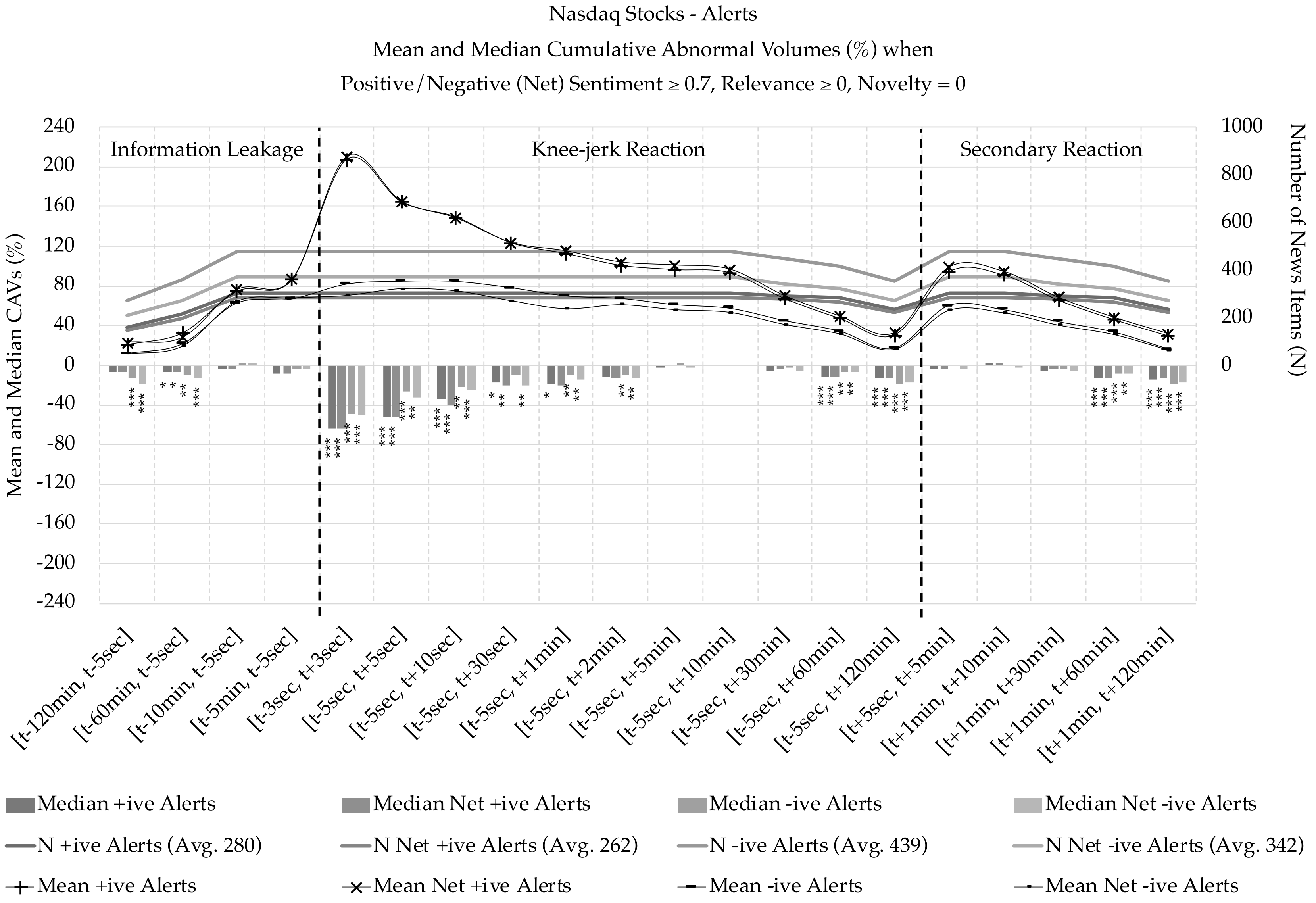
Mean and Median CAVs for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

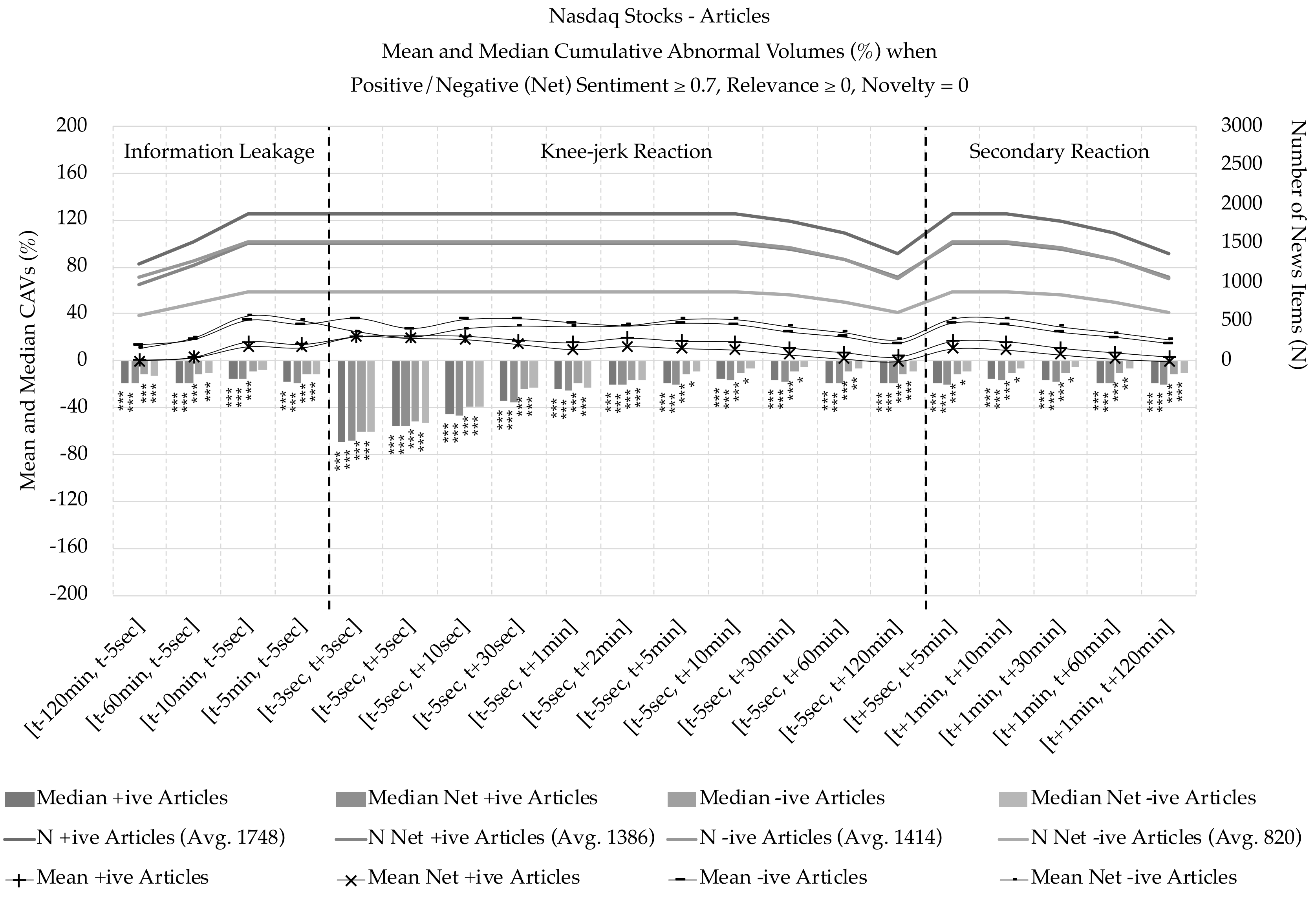
Mean and Median CAVs for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CAVs for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix G – continued — *Additional Cumulative Abnormal Volume Results*

Mean and Median CAVs for Nasdaq News across Different (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Volumes (CAVs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, no thresholds are set for Relevance and Novelty (all relevance and novelty scores are included), while absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CAVs are measured in per cent above (below) the stock's 45-day moving average volume traded during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

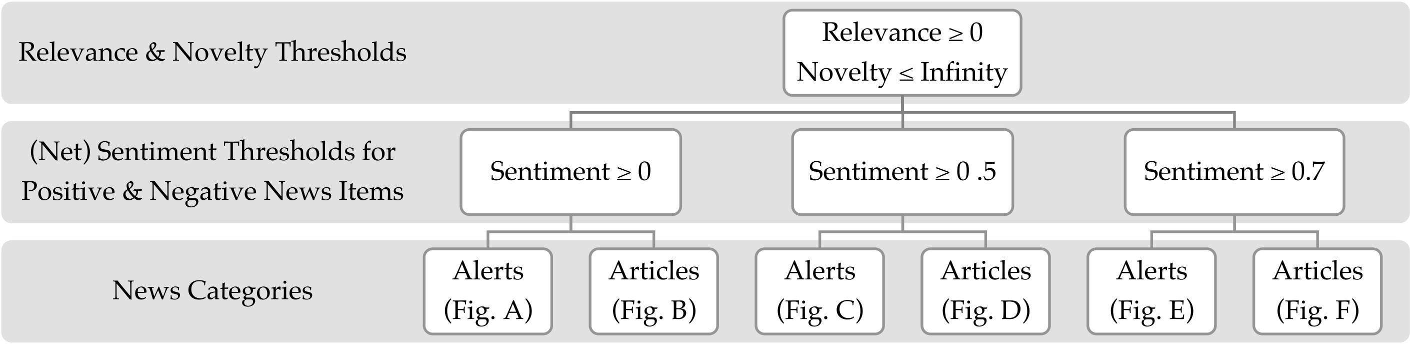
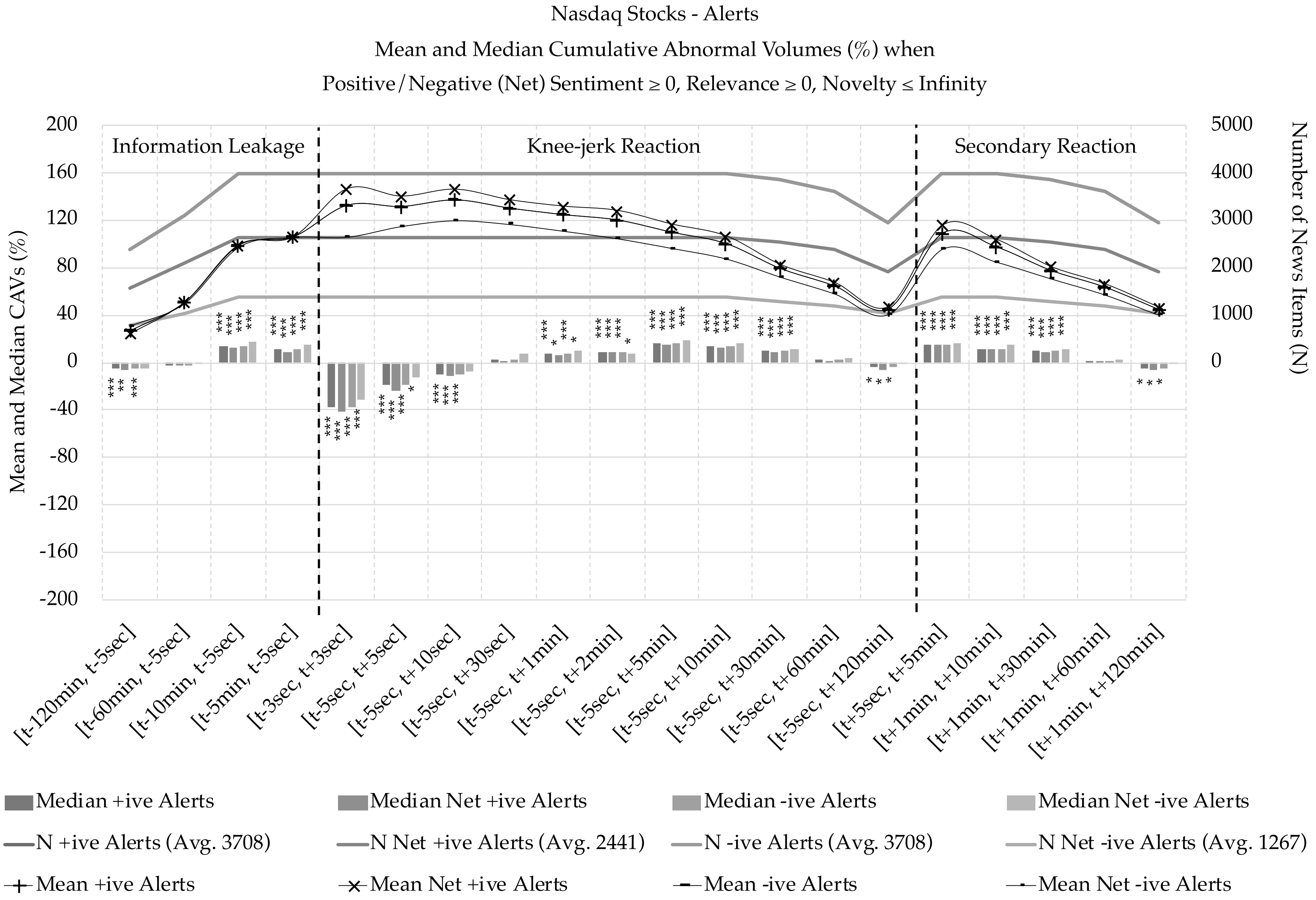


Figure A

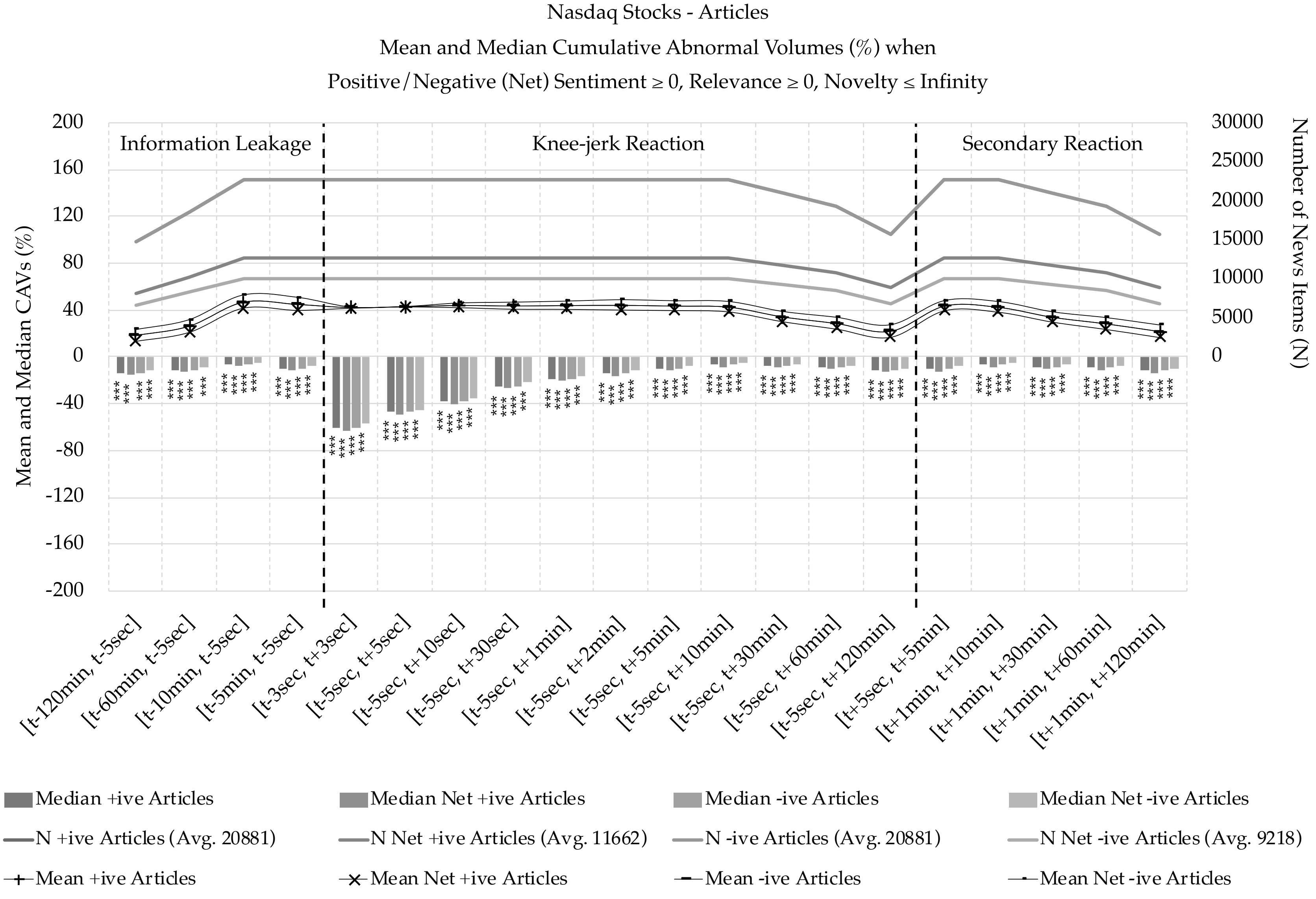
Mean and Median CAVs for Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

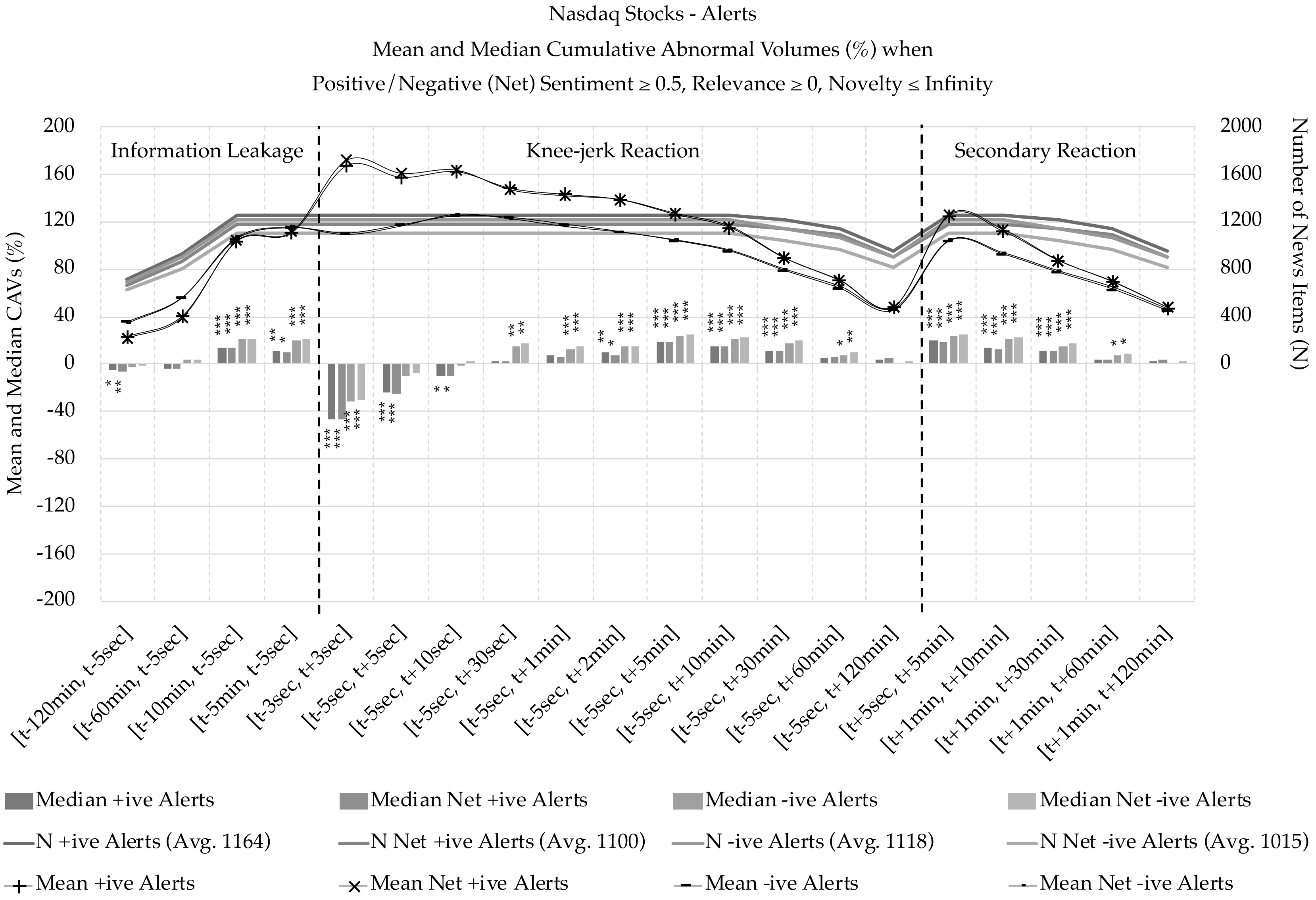
Mean and Median CAVs for Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

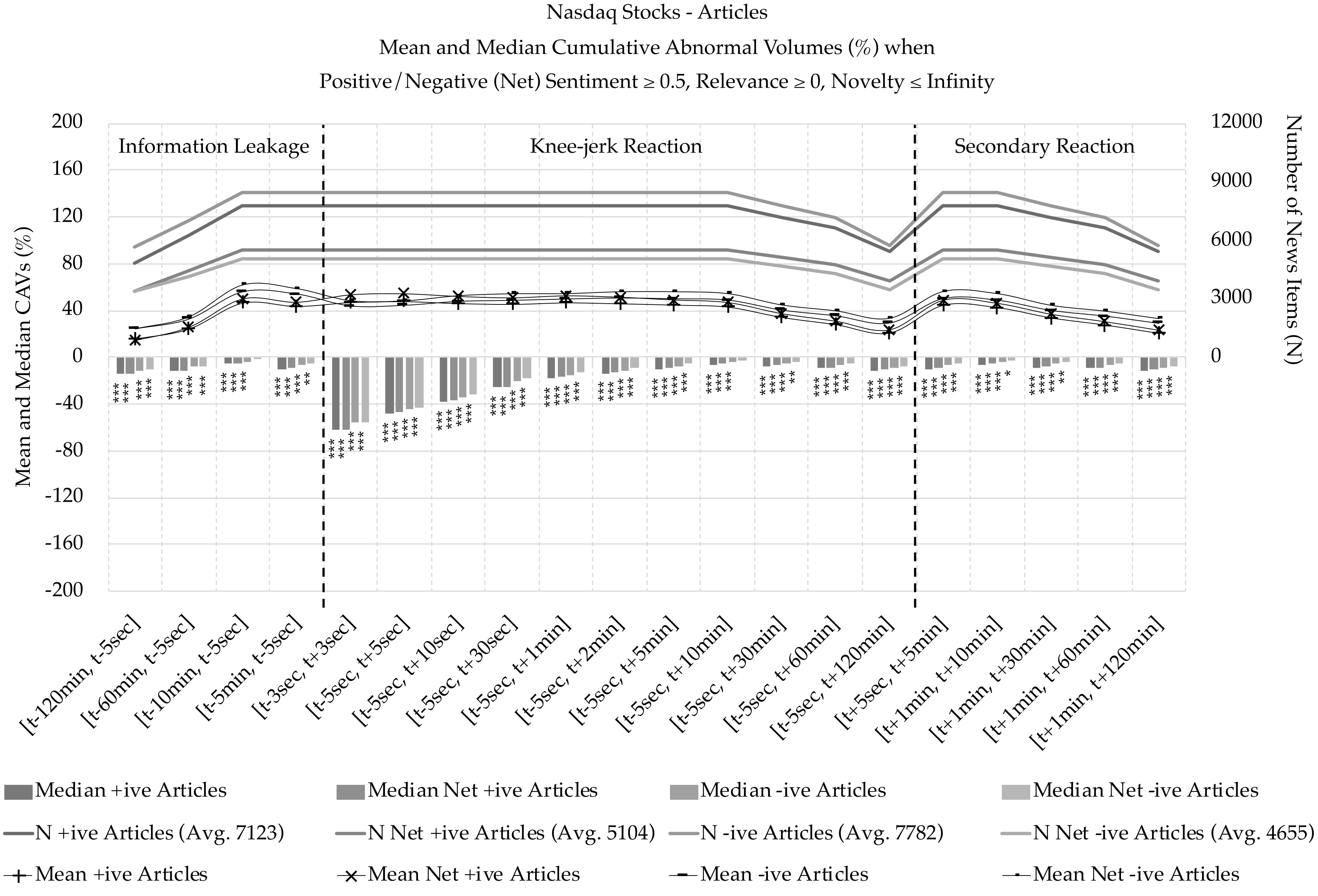
Mean and Median CAVs for Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

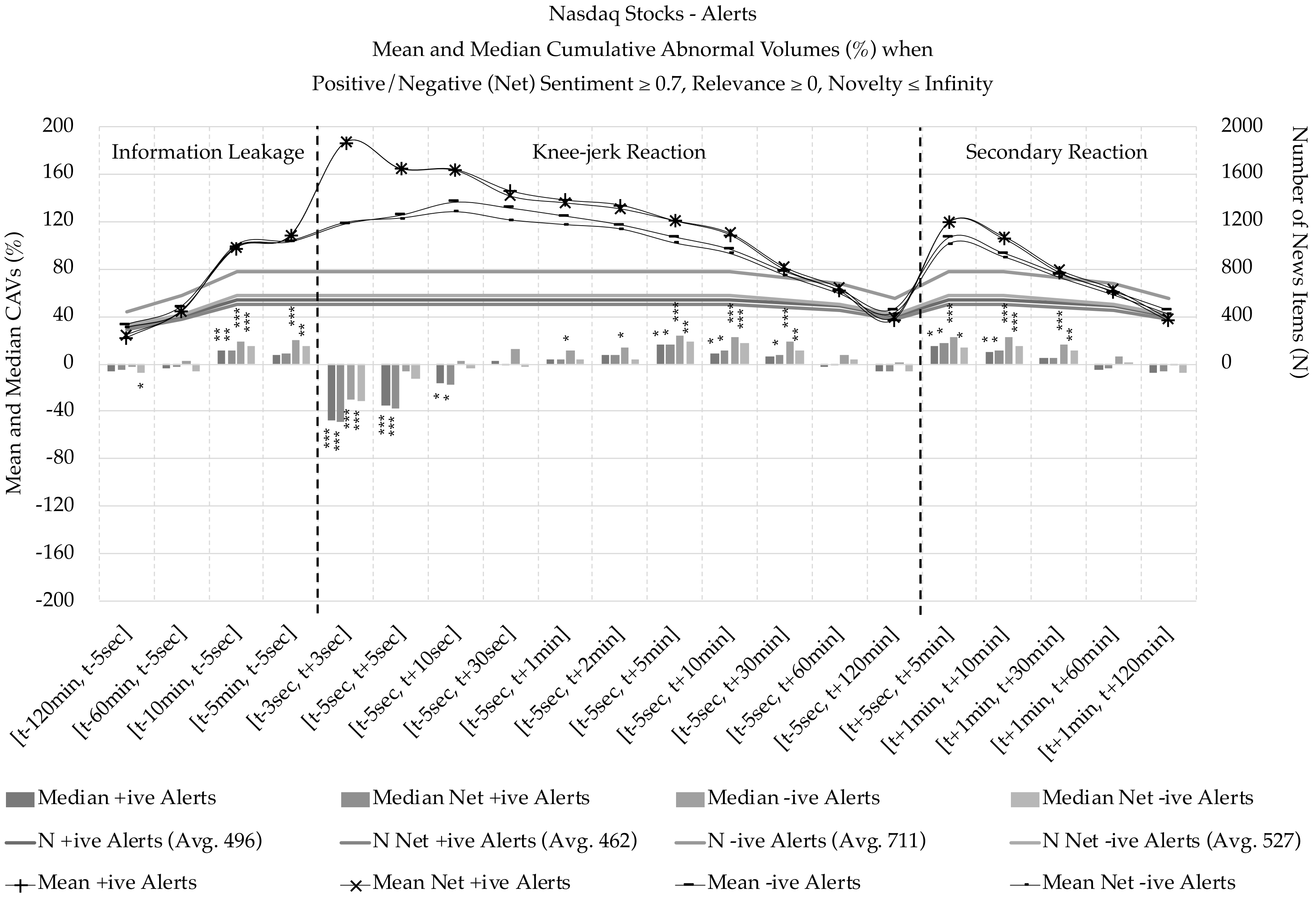
Mean and Median CAVs for Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

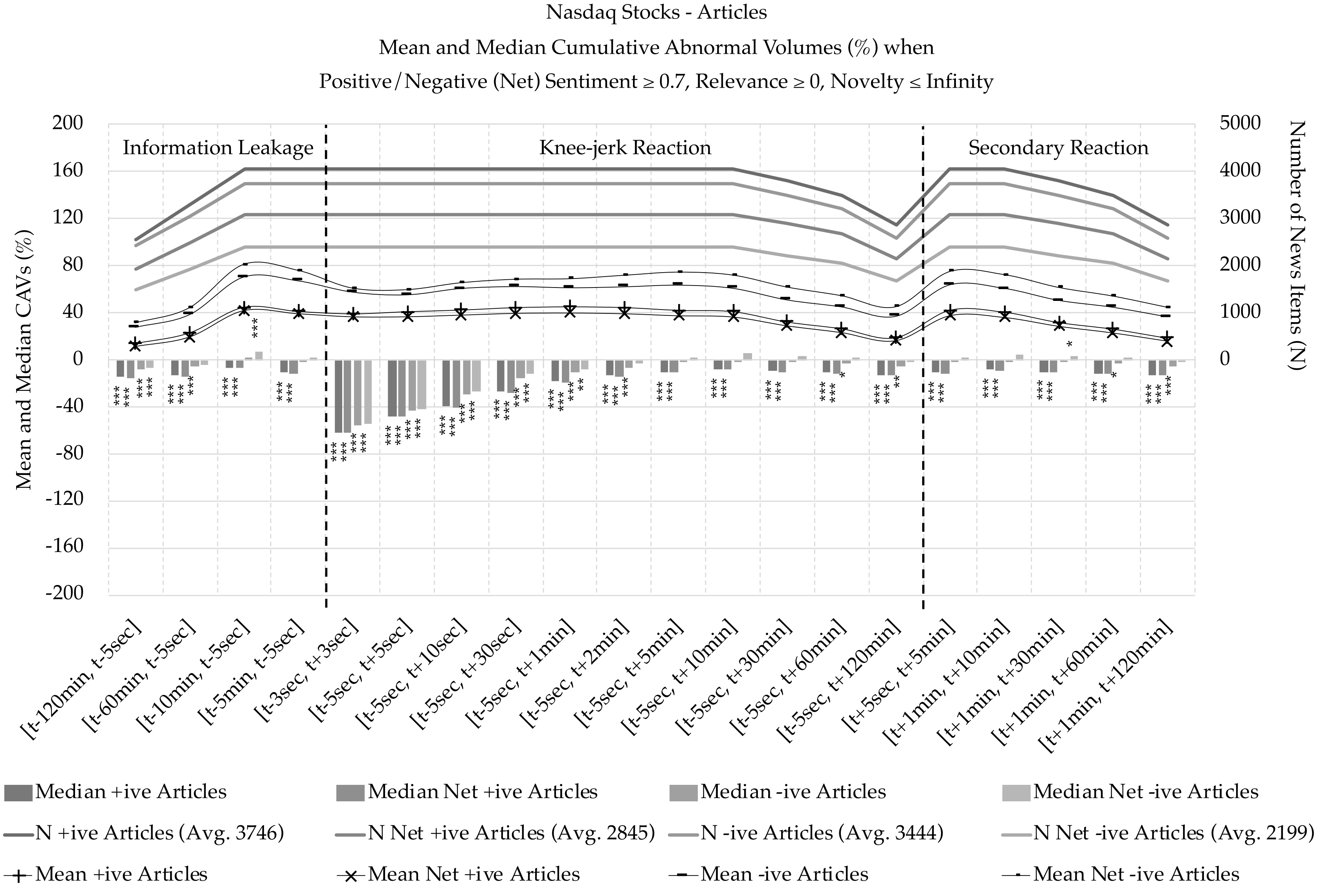
Mean and Median CAVs for Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CAVs for Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix G – continued — *Summary of CAV Results for Individual Stocks*

Stock	Total N	% Sig. Windows	Alerts														
			N Alerts	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CAVs (%)							
								[t-120min, t-5sec]		[t-3sec, t+3sec]		[t-5sec, t+1min]		[t+1min, t+120min]	
								Median	Average	Median	Average	Median	Average	Median	Average
MSFT	2914	78.96%	351	0.95	1.19	0.32	-0.54	-21.50	-7.10	-61.55	20.22	-20.84	14.18	-20.94	-11.98
CMCSA	1224	76.74%	181	0.95	0.56	0.33	-0.54	-23.49	-11.65	-73.54	40.18	-23.39	32.66	-23.87	-10.89
INTC	1630	76.08%	239	0.96	2.38	0.38	-0.62	8.27	46.40	-46.25	98.63	15.16	82.80	41.08	70.29
AAPL	6498	66.04%	948	0.99	3.40	0.31	-0.60	20.21	47.57	27.71	126.33	38.10	122.04	9.68	45.66
AMZN	2381	54.24%	291	0.96	1.32	0.34	-0.60	-11.18	15.94	-53.69	44.82	-1.02	53.08	-9.55	13.62
GOOGL	289	49.38%	82	0.99	0.71	0.30	-0.45	-22.22	-14.74	-84.08	-24.76	-38.41	-20.05	-21.85	-16.54
CTSH	156	49.38%	39	0.95	0.64	0.50	-0.44	-46.29	-35.11	-86.04	91.78	-53.68	34.50	-33.44	11.87
CSCO	812	48.26%	105	0.99	1.59	0.37	-0.61	-24.86	-12.67	-75.31	40.92	37.66	56.54	9.61	11.10
SBUX	610	45.80%	73	0.98	1.85	0.44	-0.63	-30.86	-9.20	-46.09	42.18	20.55	90.71	-17.84	29.42
VOD	1364	45.69%	171	0.91	0.47	0.42	-0.55	-3.02	0.24	-87.42	27.91	-23.10	28.56	-19.44	-11.83
CELG	507	43.19%	133	1.00	1.83	0.31	-0.50	16.12	53.11	23.51	358.84	168.47	339.96	81.63	119.44
ADP	492	42.92%	28	0.95	0.29	0.49	-0.46	-18.89	-3.09	-100.00	102.13	-82.69	15.29	17.23	17.23
FISV	117	42.57%	17	0.93	0.41	0.48	-0.38	3.98	1.22	-100.00	-76.21	-53.47	-33.54	-25.07	-0.07
GILD	537	40.17%	109	0.98	1.28	0.35	-0.60	9.30	26.12	13.95	128.90	44.09	155.44	43.17	40.27
YHOO	1278	39.83%	175	0.97	1.21	0.33	-0.62	-6.55	91.47	27.52	427.72	58.15	382.62	23.58	285.75
EBAY	1020	33.75%	139	0.98	1.24	0.35	-0.51	-11.36	9.67	-3.14	490.18	16.14	474.23	20.81	56.34
STX	217	31.15%	58	0.98	1.81	0.46	-0.62	-8.19	23.41	-63.00	311.26	22.66	160.67	29.53	39.67
ADSK	248	29.55%	38	0.98	1.84	0.59	-0.61	-13.35	51.22	-65.48	156.01	-4.01	184.85	30.52	120.77
NVDA	348	29.38%	53	0.96	2.55	0.39	-0.64	25.21	98.99	-66.92	235.99	2.27	173.62	-23.57	75.69
ATVI	303	28.92%	29	0.99	0.69	0.35	-0.65	-16.51	-11.36	-49.97	71.60	-29.75	271.62	-8.39	117.62
BBBY	242	28.23%	26	0.99	6.92	0.51	-0.70	82.81	166.75	41.93	133.96	144.07	292.75	113.28	190.03
PCAR	128	26.81%	40	0.97	1.08	0.52	-0.72	-20.95	-8.50	18.28	1235.42	56.43	435.55	-8.95	31.76
CTXS	194	26.56%	30	0.96	2.83	0.61	-0.64	4.16	13.77	-68.05	264.18	-43.97	188.58	39.57	107.71
QCOM	523	25.90%	52	0.95	0.54	0.36	-0.65	-29.89	17.66	-27.49	138.97	-5.39	132.52	-38.91	-9.64
MU	560	25.66%	72	0.97	4.07	0.39	-0.63	5.10	47.26	-48.19	70.84	-0.66	93.27	9.43	60.08
ESRX	210	24.72%	17	0.95	0.76	0.45	-0.60	-22.13	-21.67	-98.48	-25.89	-33.89	-3.22	-29.16	-7.48
ADBE	332	22.40%	48	0.99	2.65	0.26	-0.62	-11.17	7.49	-77.40	137.95	24.49	138.45	27.66	103.68
SNDK	271	22.12%	63	0.97	2.46	0.48	-0.64	7.61	115.71	-5.73	263.36	4.37	244.74	-13.20	93.64

This table reports a summary of the median and average CAVs for individual stocks for four event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAV Results for Individual Stocks — continued

Stock	Total N	% Sig. Windows	Alerts												
			N Alerts	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CAVs (%)							
								[t-120min, t-5sec]		[t-3sec, t+3sec]		[t-5sec, t+1min]		[t+1min, t+120min]	
								Median	Average	Median	Average	Median	Average	Median	Average
CHKP	68	20.80%	14	1.00	0.07	0.58	N/A	N/A	N/A	-100.00	-79.61	-6.42	40.87	44.35	34.38
MYL	562	20.63%	74	0.99	1.73	0.44	-0.53	-25.71	12.92	-26.64	148.95	-4.05	175.36	-28.83	53.80
PCLN	290	20.24%	48	0.97	3.63	0.35	-0.66	-38.15	-18.71	-99.33	-44.39	20.38	55.10	11.53	28.72
DLTR	240	19.86%	47	0.97	2.36	0.46	-0.57	-22.97	-13.17	25.72	92.49	9.13	76.21	2.33	38.75
BIIB	426	19.55%	75	0.97	0.96	0.47	-0.61	-13.39	31.29	-22.65	93.29	-0.36	156.93	-19.97	7.79
VRTX	220	19.20%	52	1.00	2.12	0.46	-0.63	46.12	83.67	-34.60	292.41	41.93	266.94	8.79	63.80
AMAT	270	19.17%	37	0.99	0.95	0.52	-0.64	17.80	42.42	-82.30	-20.73	-35.09	28.94	-4.11	3.15
MAT	205	19.10%	34	1.00	0.91	0.41	-0.58	187.32	-7.60	117.59	74.69	180.68	149.44	173.90	48.28
NTAP	141	18.54%	14	1.00	1.57	0.36	-0.61	-10.87	-11.82	-100.00	9.87	8.19	29.69	2.32	3.96
BIDU	256	17.22%	35	0.98	0.94	0.34	-0.60	-29.24	-11.45	-37.76	-10.99	-13.19	31.65	-22.41	-14.74
AMGN	472	16.39%	101	0.95	0.64	0.38	-0.58	6.87	47.20	-52.32	38.93	-22.60	65.90	-7.20	11.59
PAYX	158	16.15%	29	0.96	0.59	0.41	-0.48	2.57	-0.04	3.99	273.25	-14.98	112.84	-3.56	2.75
SYMC	213	14.27%	32	0.89	0.19	0.55	-0.62	-4.64	11.54	-89.27	235.01	-31.48	396.75	-24.31	-2.82
INTU	149	13.51%	20	1.00	2.70	0.53	-0.66	29.85	100.64	103.99	152.08	75.02	200.35	43.74	127.66
ISRG	87	13.02%	6	1.00	1.67	0.75	-0.60	N/A	N/A	-100.00	-100.00	80.33	81.67	30.44	36.43
HSIC	28	11.60%	1	1.00	0.00	N/A	-0.51	5.57	5.57	-100.00	-100.00	-22.42	-22.42	11.11	11.11
COST	482	11.28%	32	0.95	0.78	0.47	-0.60	-4.68	24.54	-36.54	45.59	4.23	51.09	-1.24	28.88
CA	110	10.90%	7	1.00	0.00	0.24	-0.76	-46.76	-46.76	-83.25	210.63	10.56	231.87	-8.39	90.14
LLTC	55	10.63%	27	0.87	0.89	0.52	-0.70	-60.92	-25.18	-31.17	107.39	-58.70	-32.18	61.92	57.32
ROST	202	10.49%	28	0.95	4.57	0.56	-0.71	-26.68	74.06	-73.99	45.48	49.32	111.04	7.12	116.14
KLAC	96	10.03%	22	0.96	1.32	0.52	-0.61	-18.56	13.37	-44.77	52.24	-10.22	37.57	-0.70	28.84
FAST	55	9.41%	14	1.00	0.36	0.21	-0.74	-9.74	20.76	-96.26	26.42	-18.50	35.60	-1.29	11.30
AKAM	99	8.89%	13	1.00	0.54	0.53	-0.68	-42.83	-38.68	-4.52	74.51	7.82	156.34	-7.08	34.88
ORLY	52	8.51%	8	1.00	5.00	0.52	-0.59	14.39	16.17	-100.00	58.67	40.77	71.09	17.49	12.46
CERN	70	7.78%	8	0.96	0.88	0.61	-0.61	-34.17	-34.17	-72.07	-58.47	125.80	117.17	11.29	44.45
SRCL	24	7.50%	10	0.96	0.40	0.48	-0.53	-12.78	-3.02	-100.00	17.47	-18.73	56.94	-38.23	-30.92
XLNX	60	5.28%	9	1.00	0.67	0.49	-0.68	8.51	4.99	93.30	1040.19	-37.24	128.90	-17.79	12.26

This table reports a summary of the median and average CAVs for individual stocks for four event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAV Results Individual Stocks — continued

			Articles												
Stock	Total N	% Sig. Windows	N Articles	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CAVs (%)							
								[t-120min, t-5sec]		[t-3sec, t+3sec]		[t-5sec, t+1min]		[t+1min, t+120min]	
								Median	Average	Median	Average	Median	Average	Median	Average
MSFT	2914	78.96%	2463	0.42	2.70	0.43	-0.48	-21.17	-3.70	-72.16	10.56	-26.14	11.49	-19.75	0.12
CMCSA	1224	76.74%	993	0.52	1.33	0.49	-0.43	-23.55	-13.10	-79.73	15.95	-31.14	3.21	-22.64	-13.46
INTC	1630	76.08%	1299	0.49	2.36	0.47	-0.53	39.95	73.98	-55.22	87.94	31.07	104.41	42.69	76.26
AAPL	6498	66.04%	5331	0.53	6.72	0.40	-0.49	-19.56	-3.82	-35.79	14.30	-21.78	13.93	-18.33	-1.68
AMZN	2381	54.24%	2008	0.43	3.37	0.44	-0.46	-24.01	-3.97	-64.81	17.73	-26.25	13.66	-21.75	-1.81
GOOGL	289	49.38%	207	0.50	2.10	0.35	-0.49	-25.09	-19.40	-79.12	-20.02	-39.11	-17.51	-31.15	-26.25
CTSH	156	49.38%	113	0.43	2.50	0.57	-0.46	-24.71	-6.79	-72.68	36.16	-40.43	9.79	-24.86	-14.82
CSCO	812	48.26%	667	0.44	1.76	0.47	-0.56	-22.08	-11.49	-78.79	6.31	-33.26	1.07	-22.90	-10.07
SBUX	610	45.80%	526	0.46	1.94	0.48	-0.54	-21.52	18.51	-52.30	32.51	-12.68	42.85	-11.50	28.78
VOD	1364	45.69%	1151	0.36	2.27	0.36	-0.47	-2.59	33.64	-79.76	33.32	-34.69	45.62	-13.33	17.61
CELG	507	43.19%	364	0.39	2.82	0.43	-0.44	11.50	48.15	-31.36	81.01	11.62	68.59	10.11	36.97
ADP	492	42.92%	336	0.49	1.12	0.48	-0.51	-14.85	1.21	-71.33	98.59	-23.29	19.59	-14.16	-3.33
FISV	117	42.57%	98	0.71	1.53	0.61	-0.50	-23.42	-19.60	-100.00	20.62	-38.21	-0.94	-25.11	-2.08
GILD	537	40.17%	411	0.43	1.93	0.45	-0.45	7.13	41.31	-30.48	56.87	5.98	70.34	8.97	40.83
YHOO	1278	39.83%	1070	0.44	2.70	0.40	-0.47	-13.17	43.06	-72.85	70.72	-24.76	63.65	-16.39	51.61
EBAY	1020	33.75%	857	0.42	2.12	0.46	-0.47	-13.13	8.36	-63.58	15.15	-21.41	27.36	-10.53	3.41
STX	217	31.15%	153	0.48	2.01	0.54	-0.48	-25.46	1.63	-82.90	56.29	-20.83	42.33	-9.49	18.58
ADSK	248	29.55%	202	0.33	4.48	0.55	-0.47	26.27	110.32	-55.94	163.87	25.86	192.30	61.25	153.97
NVDA	348	29.38%	274	0.40	3.04	0.46	-0.53	23.70	75.02	-70.13	132.20	11.61	134.92	23.52	86.44
ATVI	303	28.92%	261	0.37	1.93	0.41	-0.44	-26.04	52.22	-89.94	118.56	-44.63	94.02	-25.90	66.05
BBBY	242	28.23%	210	0.33	4.03	0.47	-0.55	29.79	135.62	-29.43	268.46	59.47	248.32	63.08	166.65
PCAR	128	26.81%	83	0.58	1.07	0.42	-0.55	-20.48	-6.82	-77.03	17.73	-23.37	43.77	-10.23	14.59
CTXS	194	26.56%	155	0.39	2.86	0.50	-0.44	5.11	31.03	-59.51	300.98	-8.08	167.95	8.57	42.58
QCOM	523	25.90%	454	0.39	1.27	0.47	-0.48	-19.45	-2.26	-62.92	19.29	-20.59	24.72	-14.46	8.37
MU	560	25.66%	476	0.42	3.47	0.48	-0.55	15.76	50.53	-68.69	52.40	-8.11	69.43	22.87	40.63
ESRX	210	24.72%	187	0.50	2.19	0.42	-0.41	-11.27	15.43	-57.05	16.74	-20.57	33.94	-11.55	6.05
ADBE	332	22.40%	274	0.48	3.73	0.48	-0.51	6.46	59.90	-54.59	140.16	25.40	134.71	25.35	83.39
SNDK	271	22.12%	205	0.46	2.31	0.50	-0.54	14.46	108.71	-54.37	119.79	25.61	138.48	12.77	113.82

This table reports a summary of the median and average CAVs for individual stocks for four event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Summary of CAV Results Individual Stocks — continued

			Articles												
Stock	Total N	% Sig. Windows	N Articles	Relevance	Novelty	Net Tone +ive News	Net Tone -ive News	CAVs (%)							
								[t-120min, t-5sec]		[t-3sec, t+3sec]		[t-5sec, t+1min]		[t+1min, t+120min]	
								Median	Average	Median	Average	Median	Average	Median	Average
CHKP	68	20.80%	53	0.44	1.40	0.52	-0.49	-5.17	-1.27	-96.76	31.35	-4.77	-7.83	-20.28	-8.50
MYL	562	20.63%	484	0.90	30.10	0.33	-0.45	-21.15	30.16	-59.49	23.39	-24.99	29.14	-25.63	2.29
PCLN	290	20.24%	236	0.44	2.60	0.48	-0.49	6.39	26.79	-99.71	75.22	-10.82	48.64	9.00	28.90
DLTR	240	19.86%	193	0.34	3.30	0.53	-0.46	-23.11	-3.98	-81.80	-9.94	-24.03	19.52	-8.04	12.75
BIIB	426	19.55%	337	0.32	2.74	0.50	-0.42	9.02	52.62	-60.21	87.65	2.78	91.56	-6.74	42.51
VRTX	220	19.20%	164	0.54	1.74	0.43	-0.51	16.14	81.66	-40.71	192.08	36.13	271.70	22.07	115.23
AMAT	270	19.17%	220	0.36	4.24	0.49	-0.58	6.34	42.38	-82.15	36.99	-20.98	78.27	3.06	48.31
MAT	205	19.10%	164	0.55	0.88	0.41	-0.34	-23.87	2.84	-89.06	6.96	-42.93	24.74	-13.42	12.10
NTAP	141	18.54%	119	0.43	3.63	0.48	-0.41	-14.58	-10.05	-88.34	-28.14	-43.23	-6.13	-27.38	-16.63
BIDU	256	17.22%	208	0.39	1.81	0.45	-0.45	-19.38	0.96	-63.70	6.68	-11.92	24.12	-24.02	-10.77
AMGN	472	16.39%	350	0.41	2.41	0.45	-0.43	3.67	40.45	-50.82	42.38	-11.57	36.88	3.50	19.68
PAYX	158	16.15%	129	0.76	0.99	0.52	-0.48	-15.42	19.80	-83.15	76.87	6.35	76.21	-0.34	31.95
SYMC	213	14.27%	170	0.31	1.33	0.37	-0.45	-6.39	39.06	-85.50	105.20	-30.04	94.07	-13.33	42.60
INTU	149	13.51%	123	0.34	2.03	0.50	-0.56	25.26	118.90	-25.87	178.66	20.80	141.12	34.68	89.98
ISRG	87	13.02%	81	0.36	2.99	0.64	-0.48	12.64	233.41	-100.00	218.46	66.99	432.72	62.18	315.04
HSIC	28	11.60%	27	0.68	0.93	0.71	-0.31	-3.89	19.72	-100.00	62.86	-53.68	5.30	-31.95	-9.96
COST	482	11.28%	407	0.33	2.51	0.44	-0.46	-1.75	28.45	-50.41	34.91	-3.29	51.83	-0.34	29.44
CA	110	10.90%	96	0.55	1.28	0.54	-0.49	-8.54	40.90	-80.76	44.19	-21.24	44.47	-16.76	25.69
LLTC	55	10.63%	28	0.41	2.14	0.51	-0.47	15.71	24.22	-80.63	140.42	15.06	48.70	44.59	50.89
ROST	202	10.49%	162	0.34	3.07	0.46	-0.39	54.77	98.87	-54.82	77.82	13.39	123.87	24.26	86.21
KLAC	96	10.03%	70	0.34	2.57	0.43	-0.41	-0.90	46.46	-53.69	65.61	21.81	90.92	35.33	59.36
FAST	55	9.41%	41	0.37	2.22	0.48	-0.52	51.94	57.85	-54.73	-28.31	25.90	54.57	38.30	34.18
AKAM	99	8.89%	83	0.48	2.57	0.59	-0.46	-21.31	-0.27	-49.21	65.22	-14.74	87.59	-3.01	31.04
ORLY	52	8.51%	41	0.42	2.85	0.54	-0.55	-7.92	-3.20	-100.00	19.75	-51.84	84.34	-2.95	-3.31
CERN	70	7.78%	62	0.35	2.68	0.55	-0.59	-25.55	1.22	-100.00	1.15	-23.61	17.97	-16.27	4.01
SRCL	24	7.50%	14	0.41	2.64	0.55	-0.43	25.78	69.67	-100.00	-42.31	-13.68	69.10	13.69	32.27
XLNX	60	5.28%	49	0.41	2.63	0.47	-0.46	-28.14	39.59	-71.01	191.95	-23.04	77.61	-12.55	33.67

This table reports a summary of the median and average CAVs for individual stocks for four event windows from the Information Leakage, Knee-jerk Reaction, and Secondary Reaction groups, respectively. The number of new items per stock and the average Relevance, Novelty, and Net Tone scores are reported alongside, and are organized in decreasing order of significance (determined as a percentage of all event windows across all tests).

Appendix H — *Additional Cumulative Abnormal Trade Results*

Mean and Median CATs for Relevant, Novel Nasdaq News across (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Trades (CATs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, Relevance is set to 1 (most relevant news), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CATs are measured in per cent above (below) the stock's 45-day moving average number of trades during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

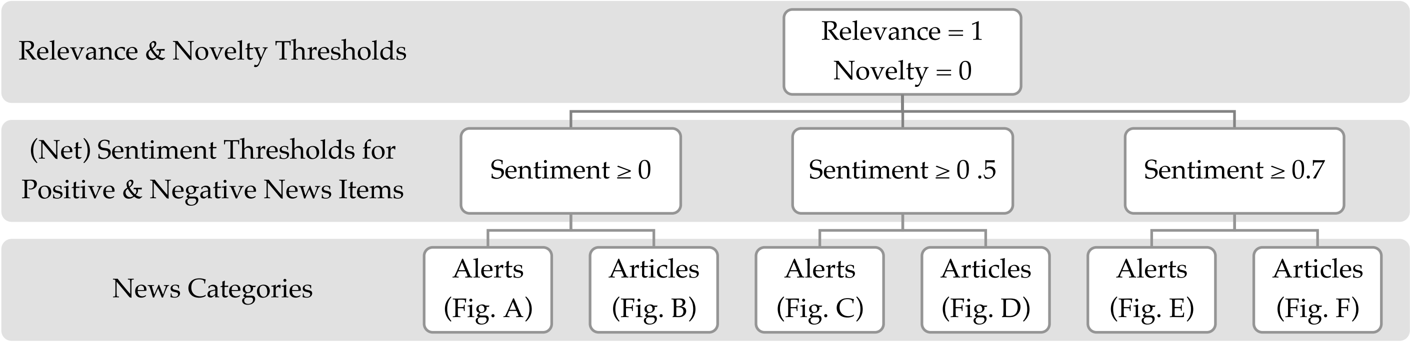
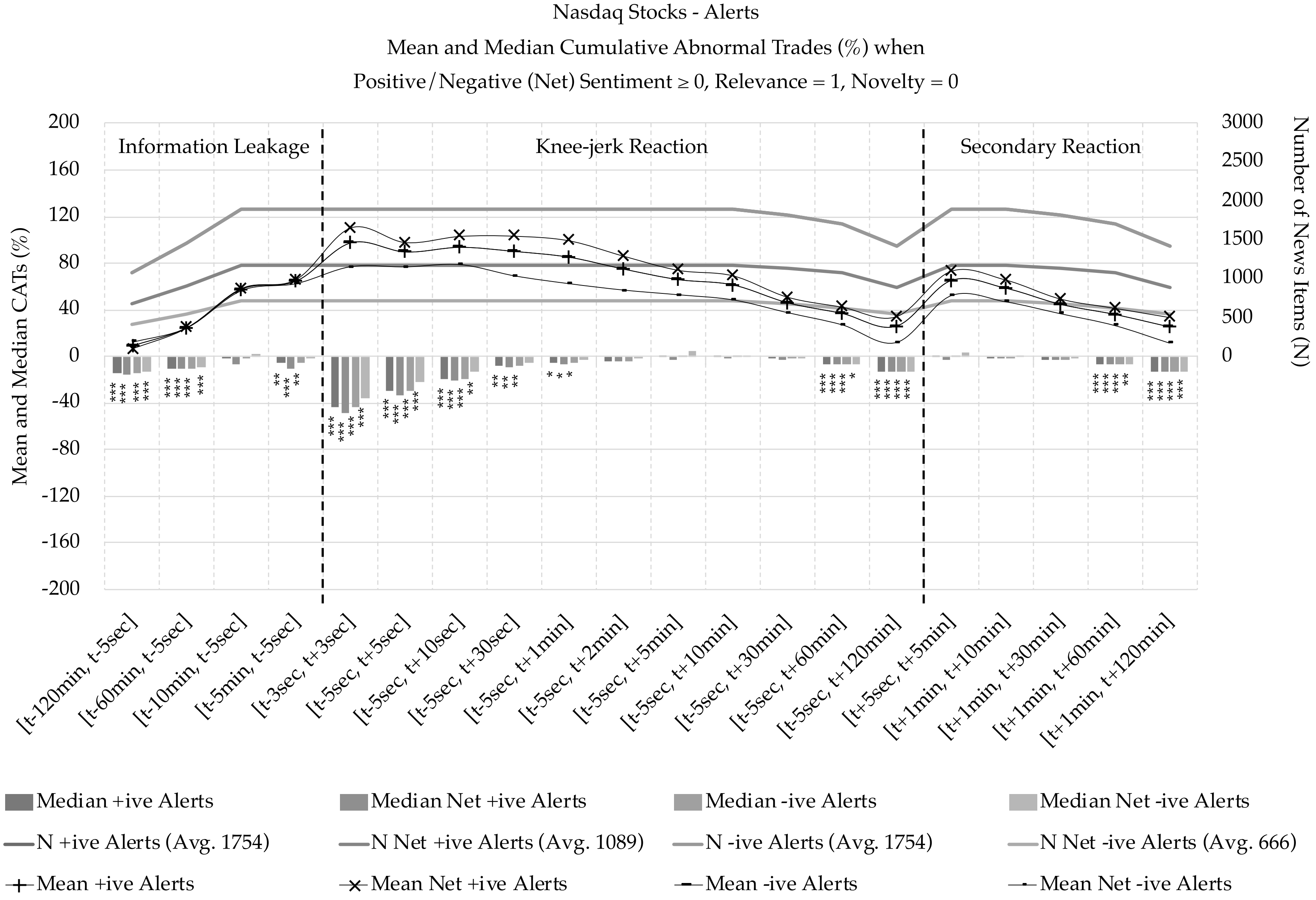


Figure A

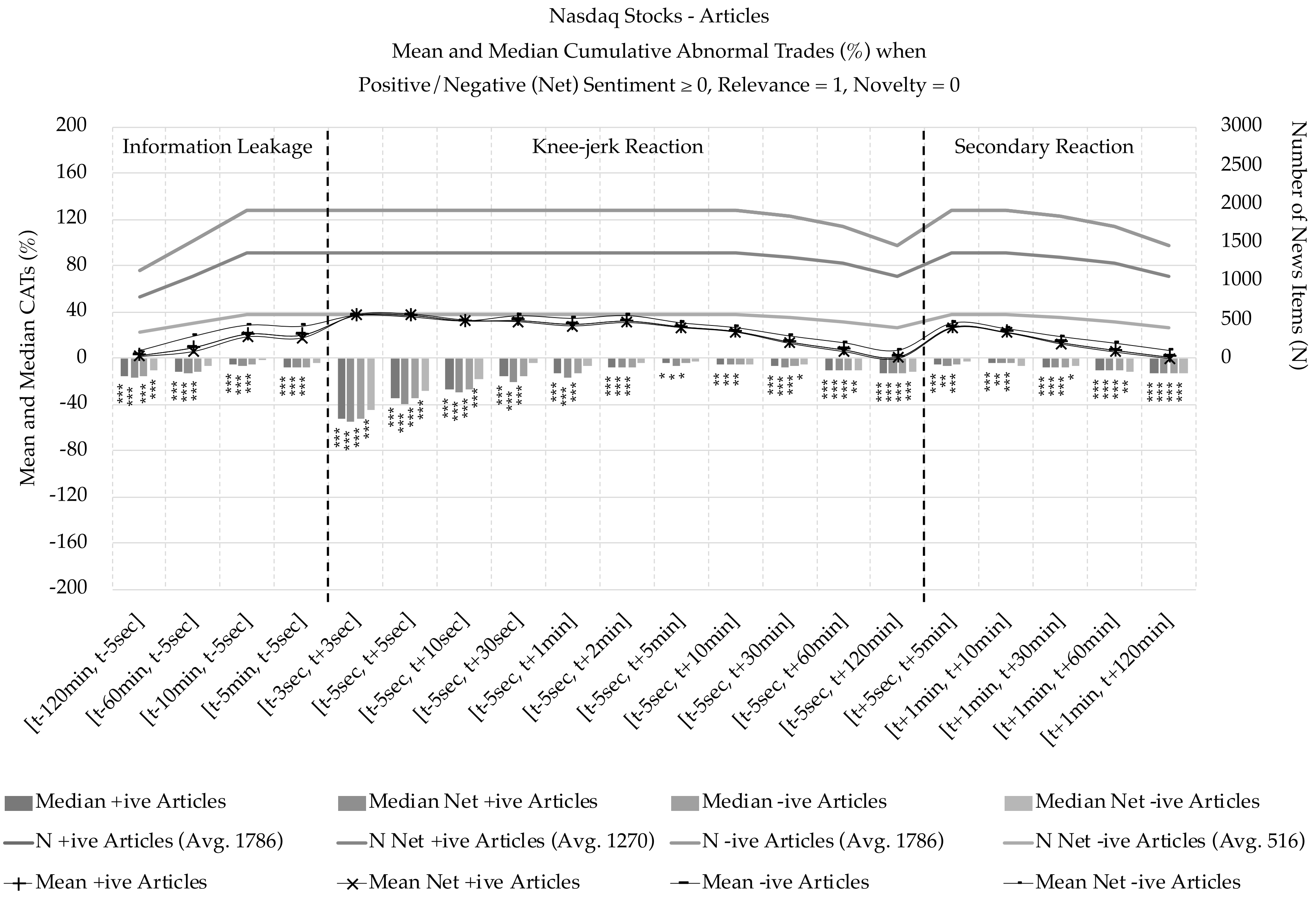
Mean and Median CATs for Relevant, Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

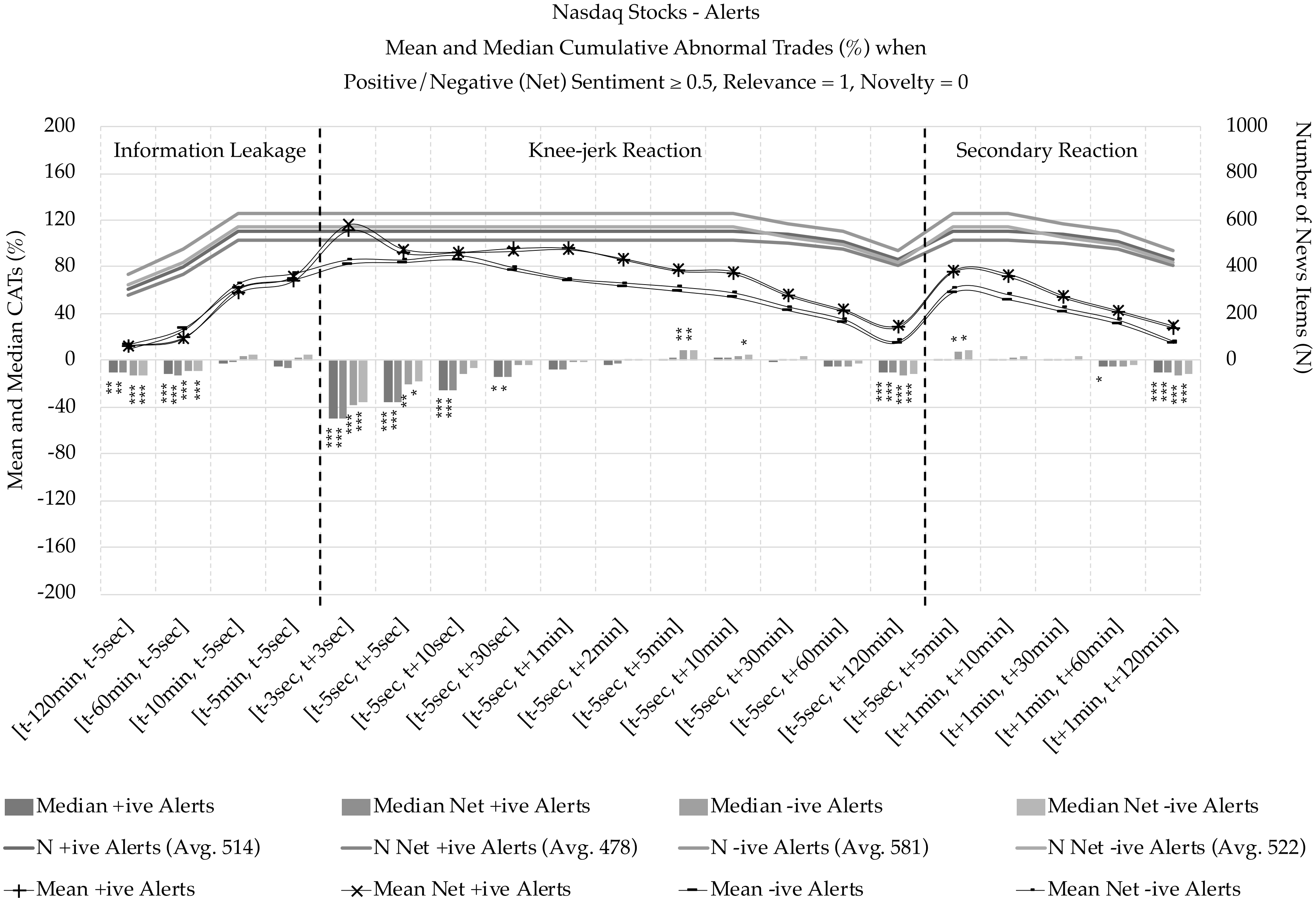
Mean and Median CATs for Relevant, Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

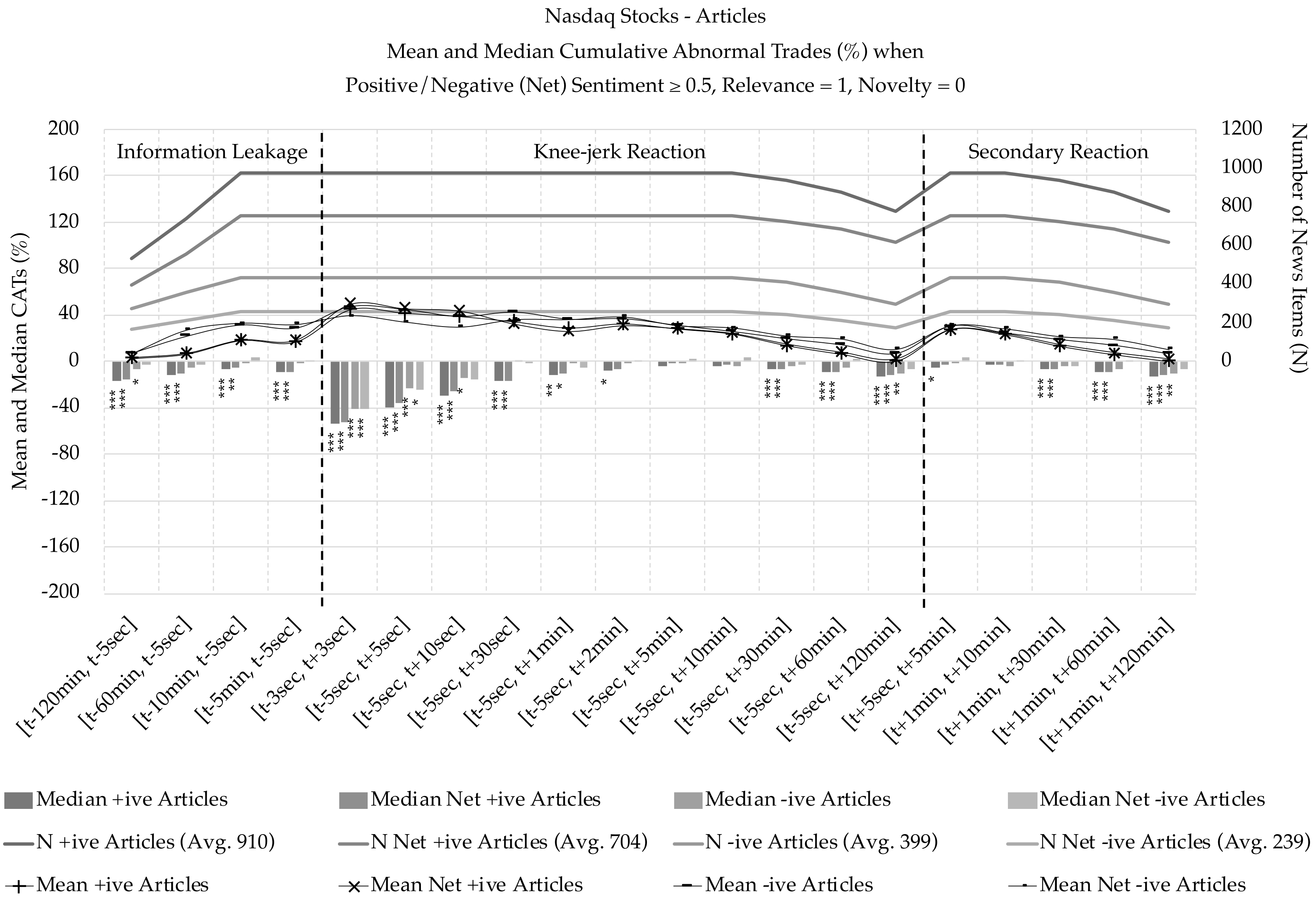
Mean and Median CATs for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

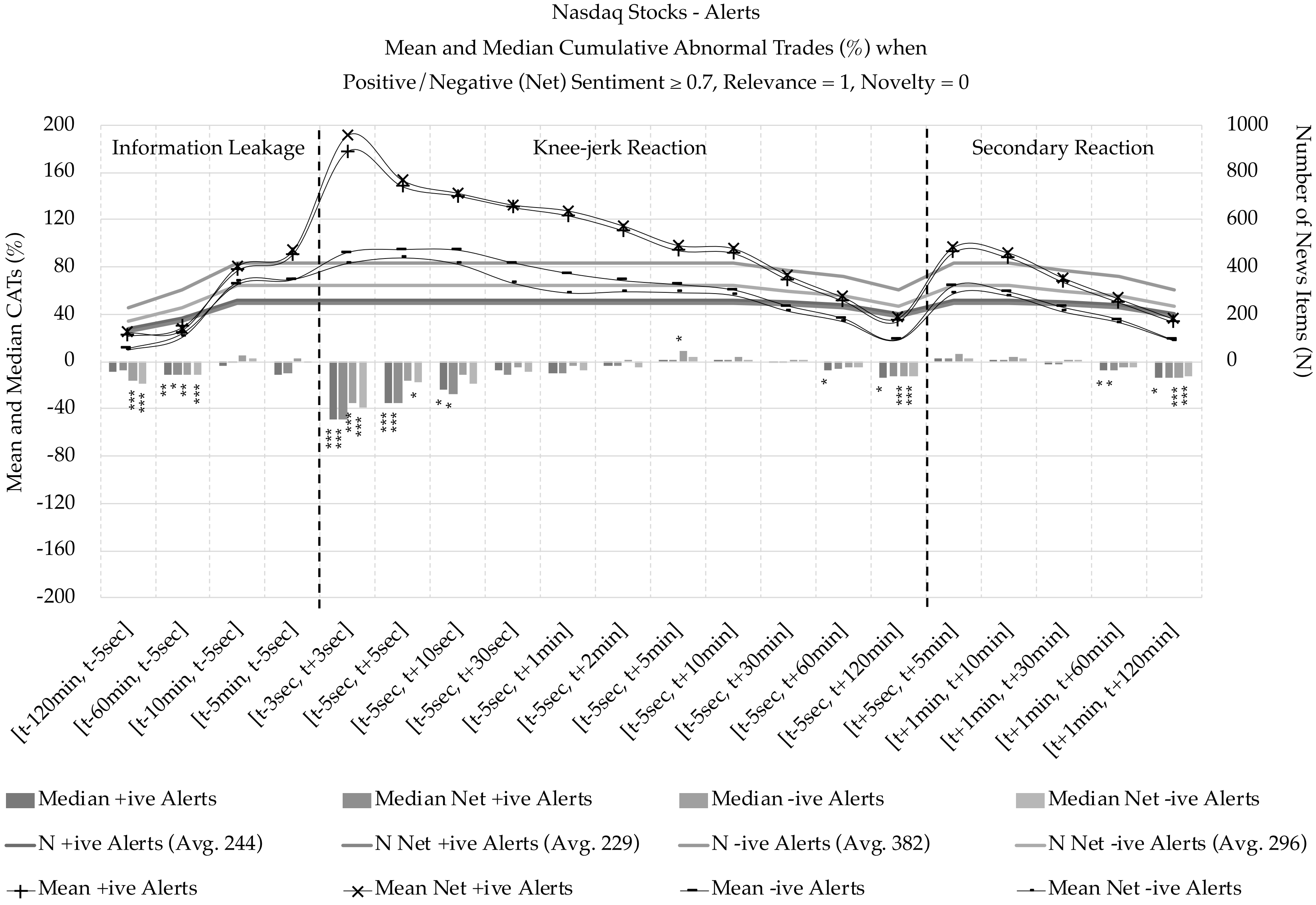
Mean and Median CATs for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

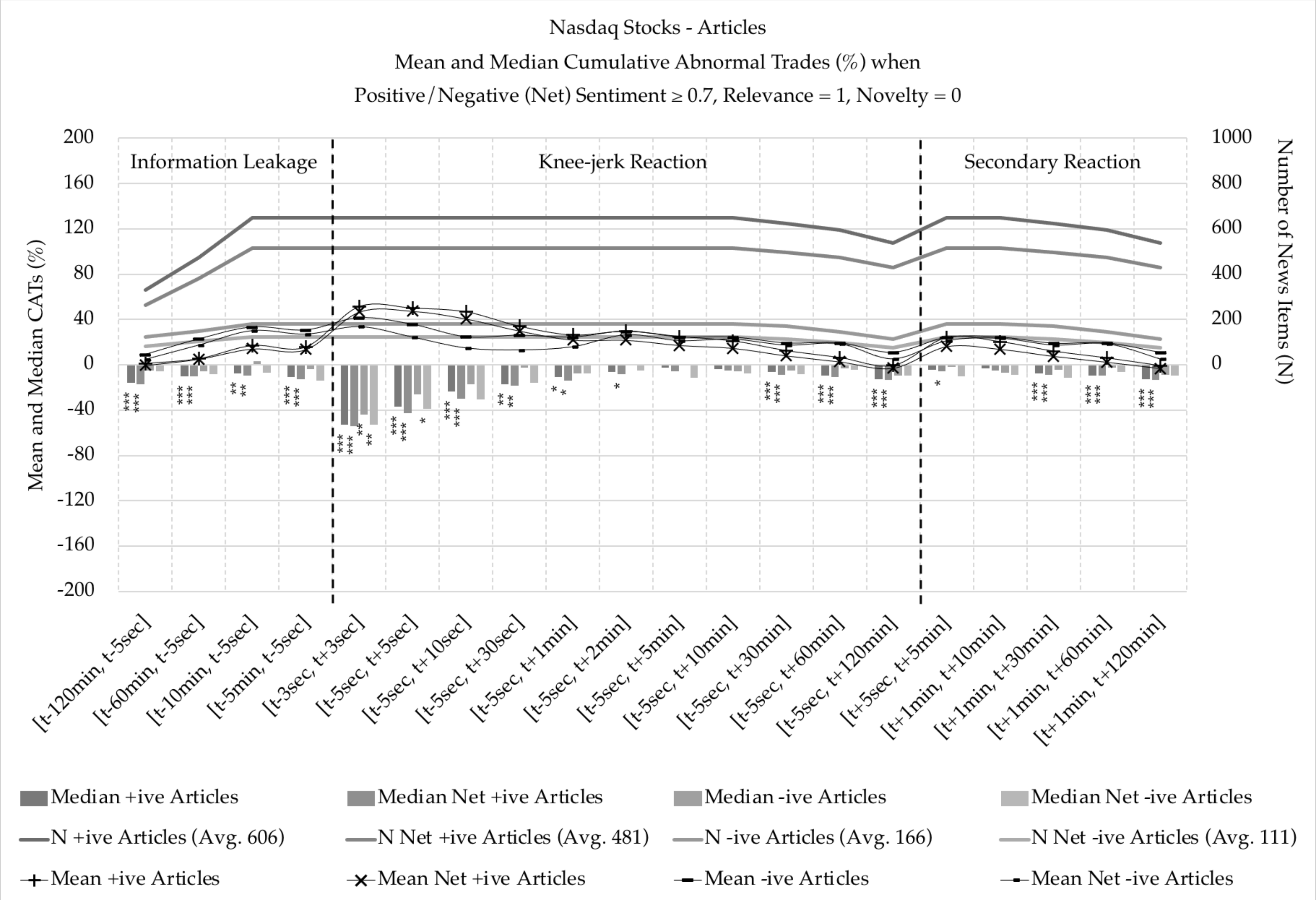
Mean and Median CATs for Relevant, Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CATs for Relevant, Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix H – continued — *Additional Cumulative Abnormal Trade Results*

Mean and Median CATs for Relevant Nasdaq News across (Net) Sentiment Thresholds

The six charts that follow report the median and mean Cumulative Abnormal Trades (CATs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, Relevance is set to 1 (most relevant news), no threshold is set for Novelty (all novelty scores are included), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CATs are measured in per cent above (below) the stock's 45-day moving average number of trades during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

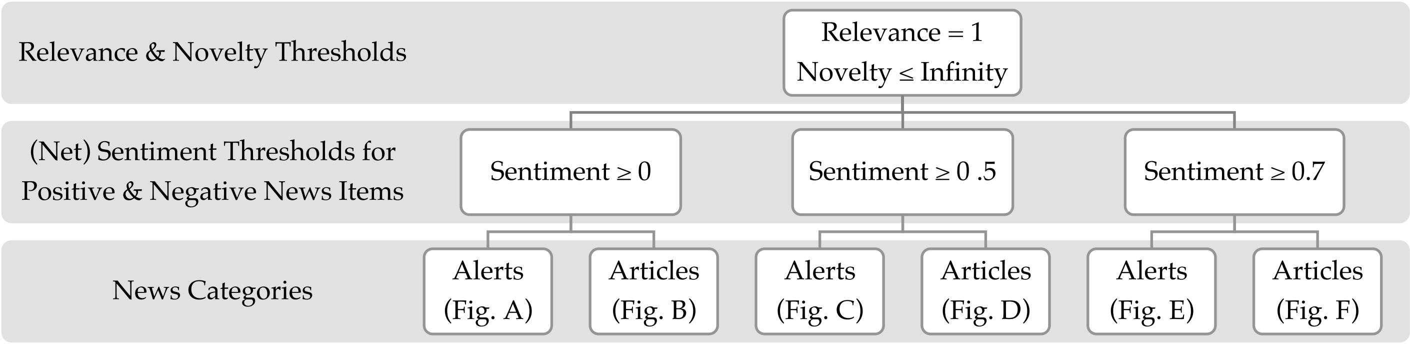
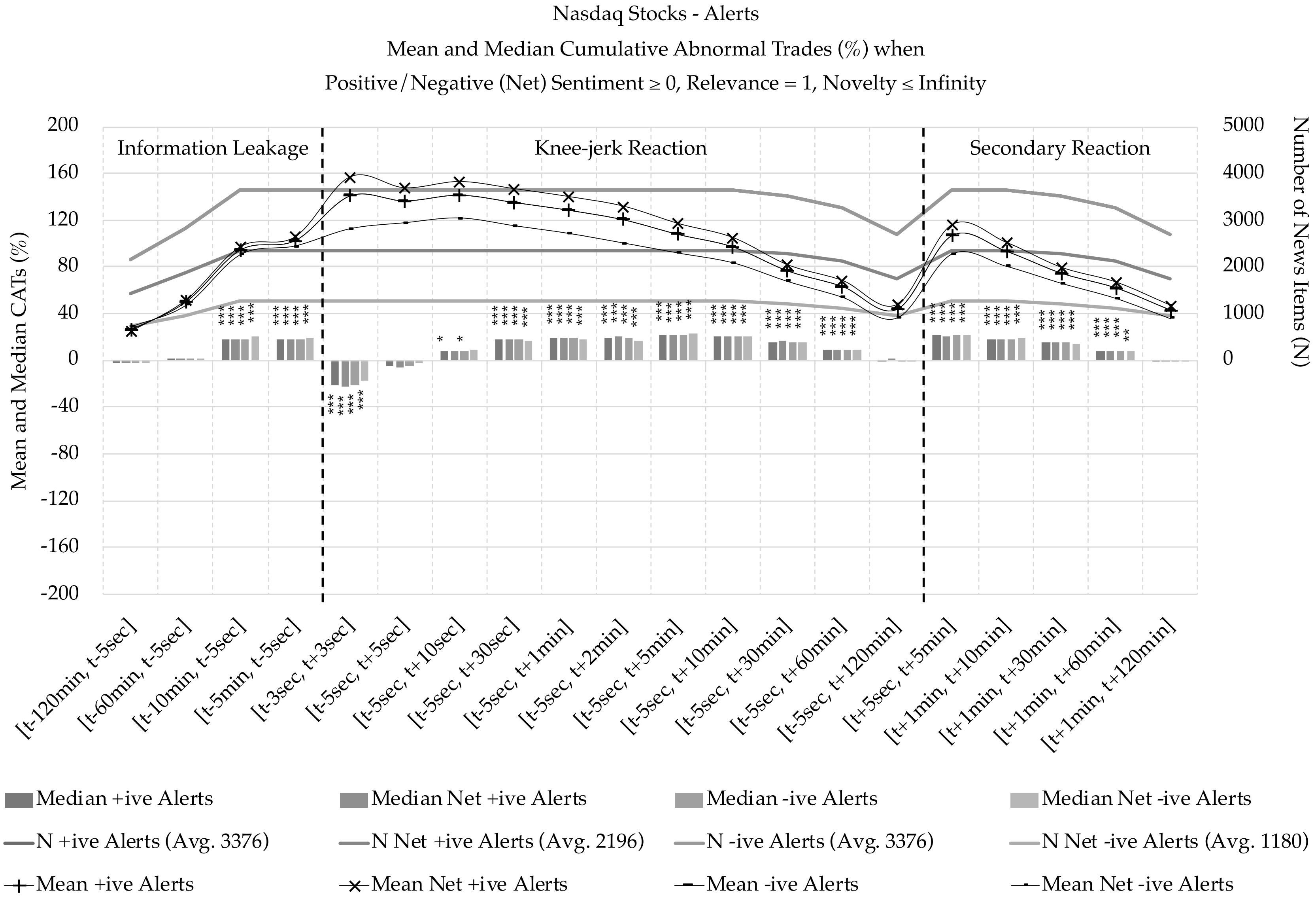


Figure A

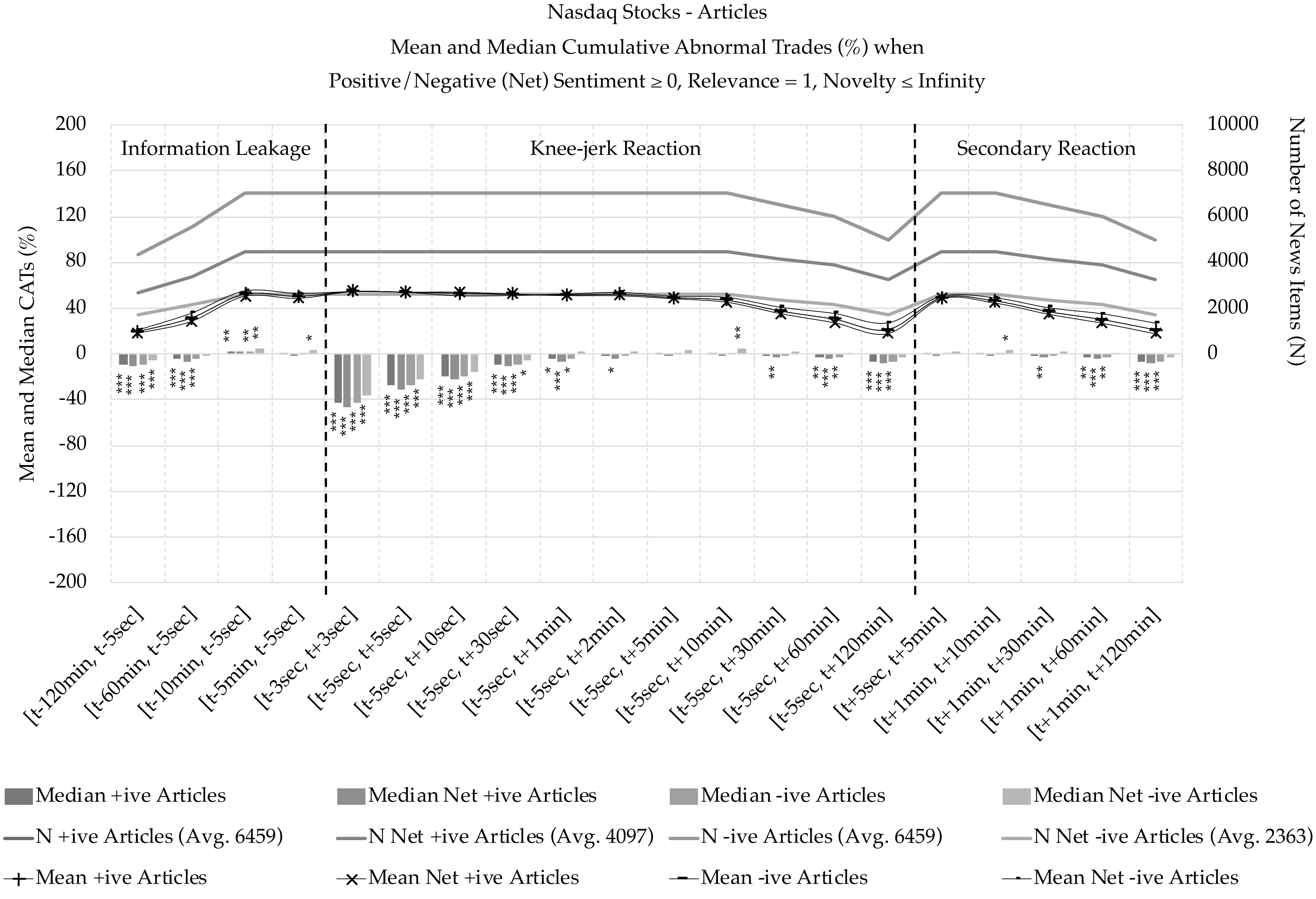
Mean and Median CATs for Relevant Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

Mean and Median CATs for Relevant Nasdaq Articles for all (Net) Sentiment Values

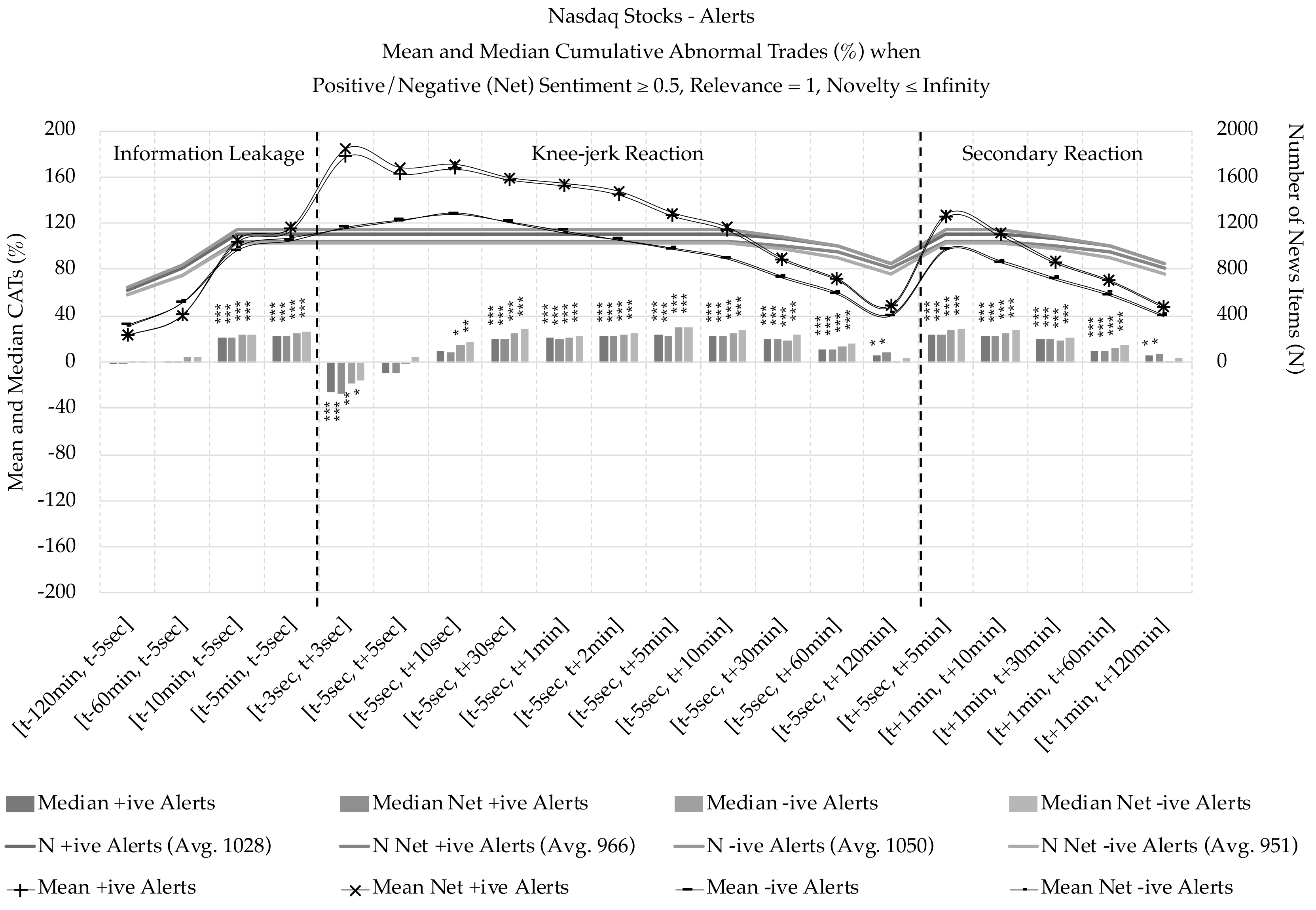


Nasdaq Stocks - Articles
Mean and Median Cumulative Abnormal Trades (%) when
Positive/Negative (Net) Sentiment ≥ 0, Relevance = 1, Novelty ≤ Infinity

Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

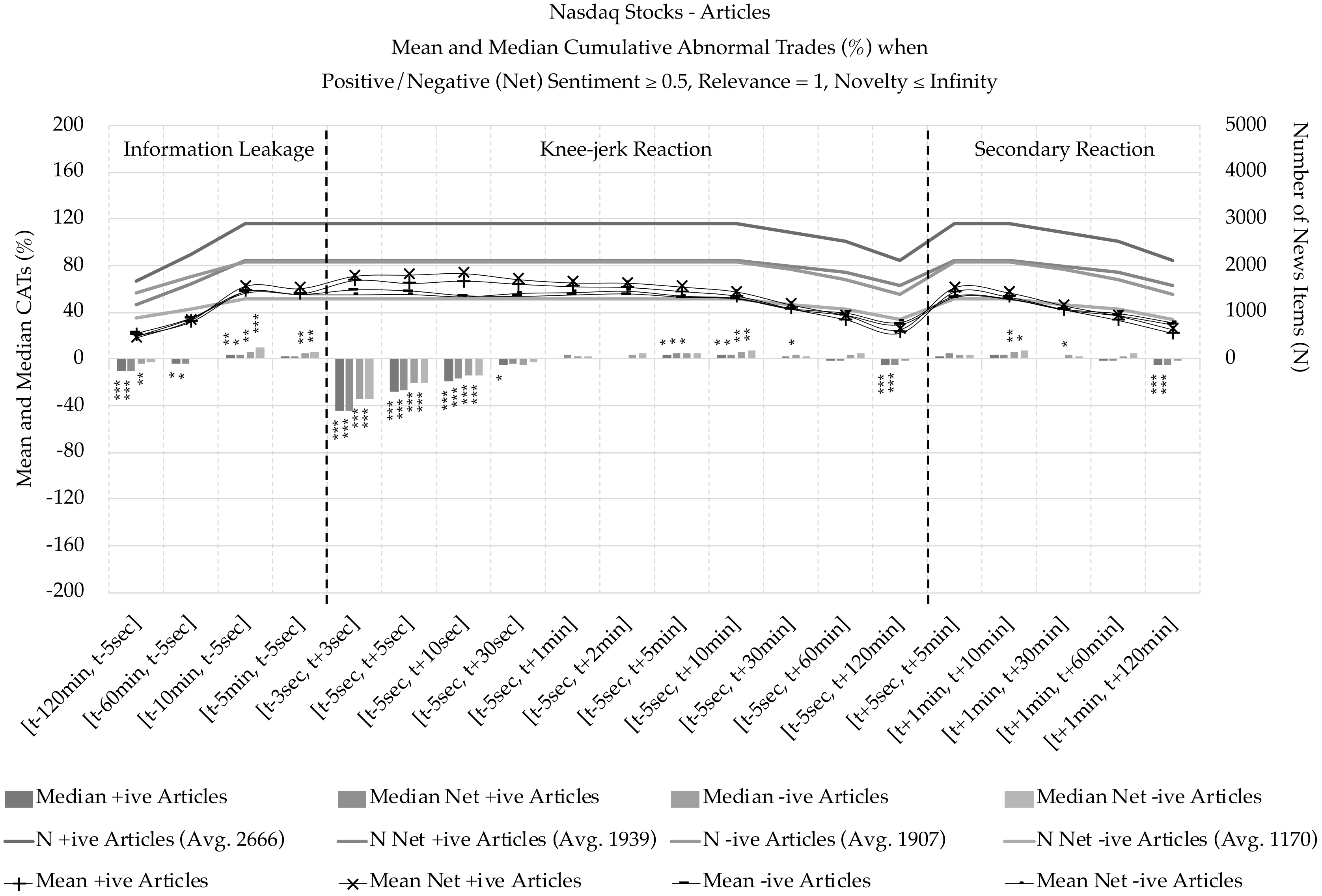
Mean and Median CATs for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

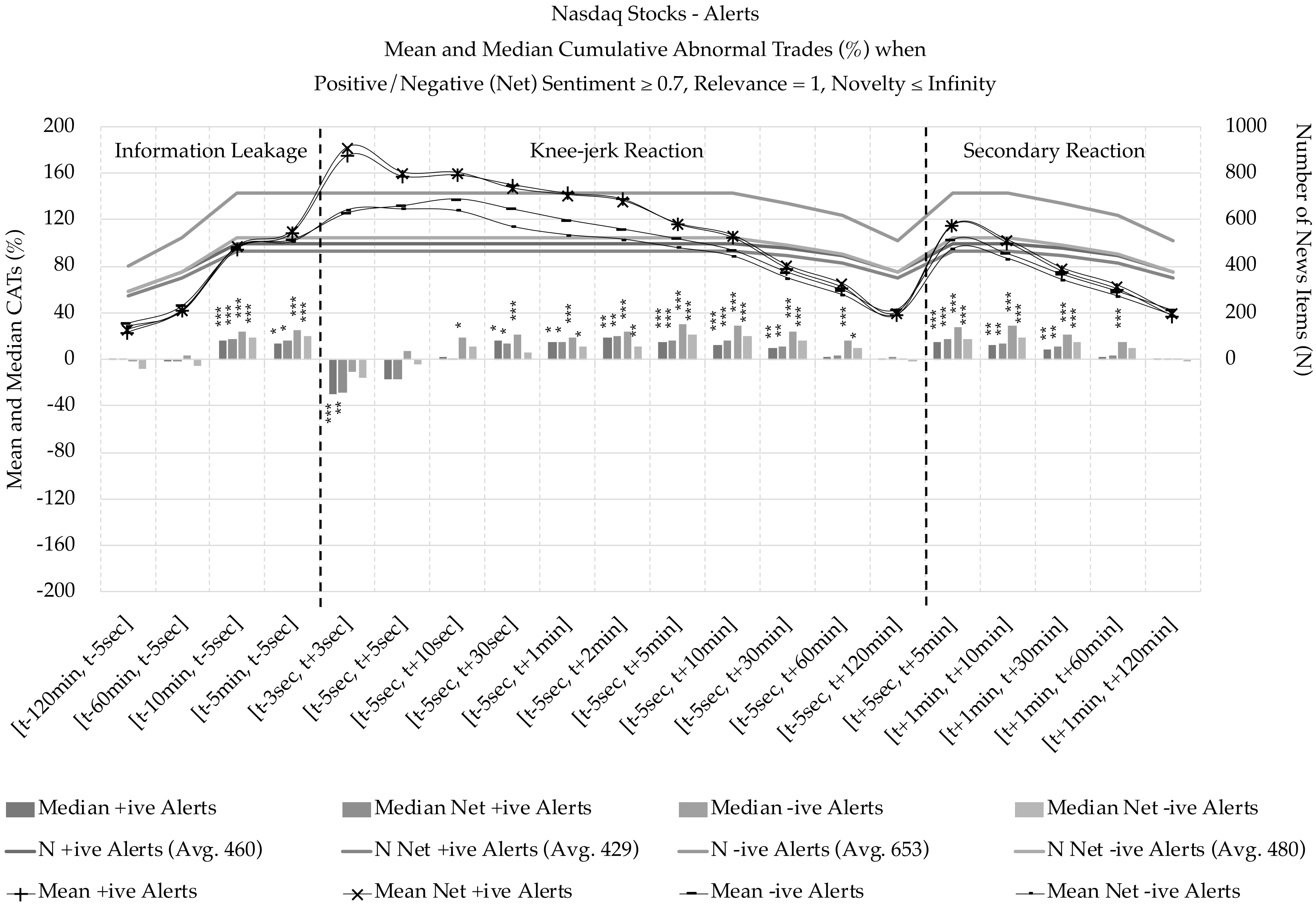
Mean and Median CATs for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

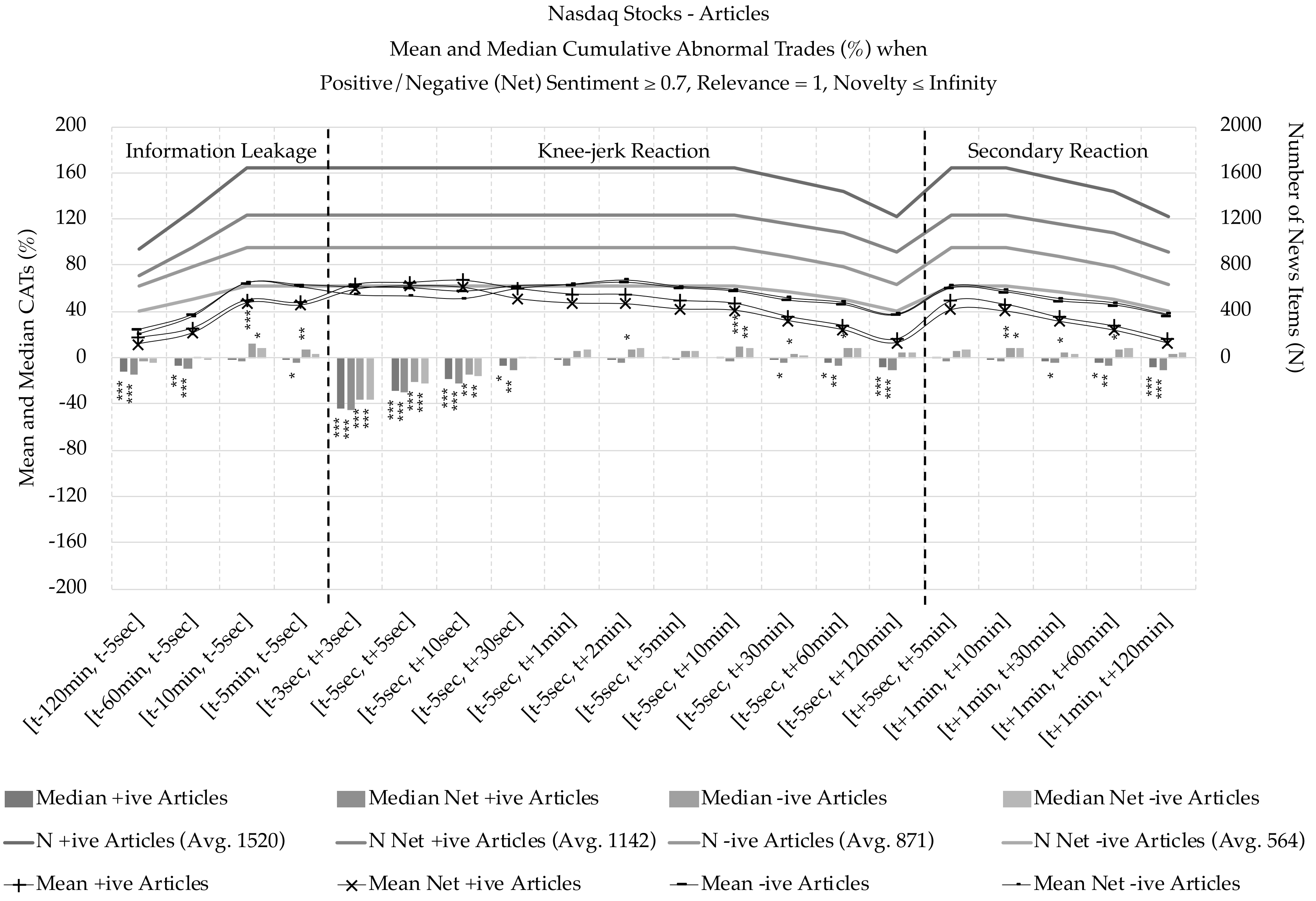
Mean and Median CATs for Relevant Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CATs for Relevant Nasdaq Articles when (Net) Sentiment ≥ 0.7



Nasdaq Stocks - Articles

Mean and Median Cumulative Abnormal Trades (%) when

Positive/Negative (Net) Sentiment ≥ 0.7, Relevance = 1, Novelty ≤ Infinity

Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix H – continued — *Additional Cumulative Abnormal Trade Results*

Mean and Median CATs for Novel Nasdaq News across (Net) Sentiment Thresholds

---

The six charts that follow report the median and mean Cumulative Abnormal Trades (CATs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, no threshold is set for Relevance (all relevance scores are included), Novelty is set to 0 (most novel news), and absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CATs are measured in per cent above (below) the stock's 45-day moving average number of trades during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

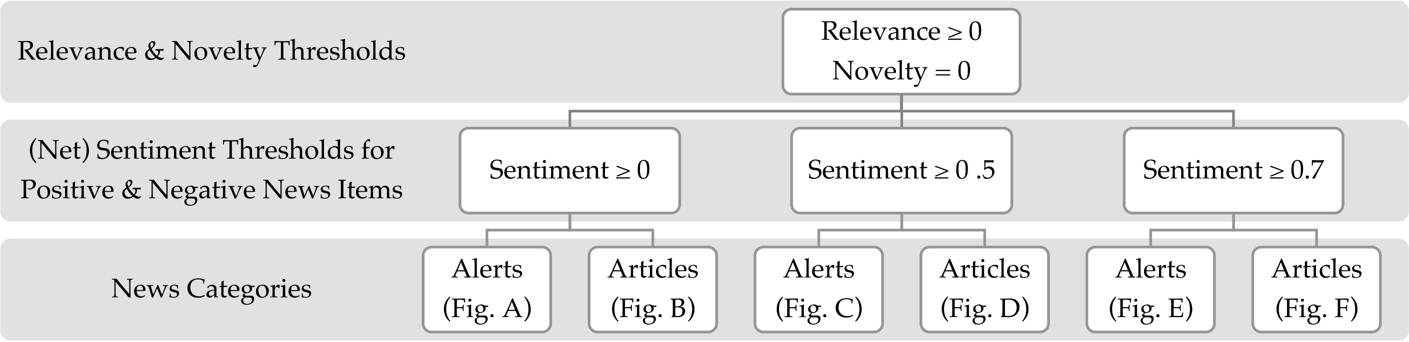
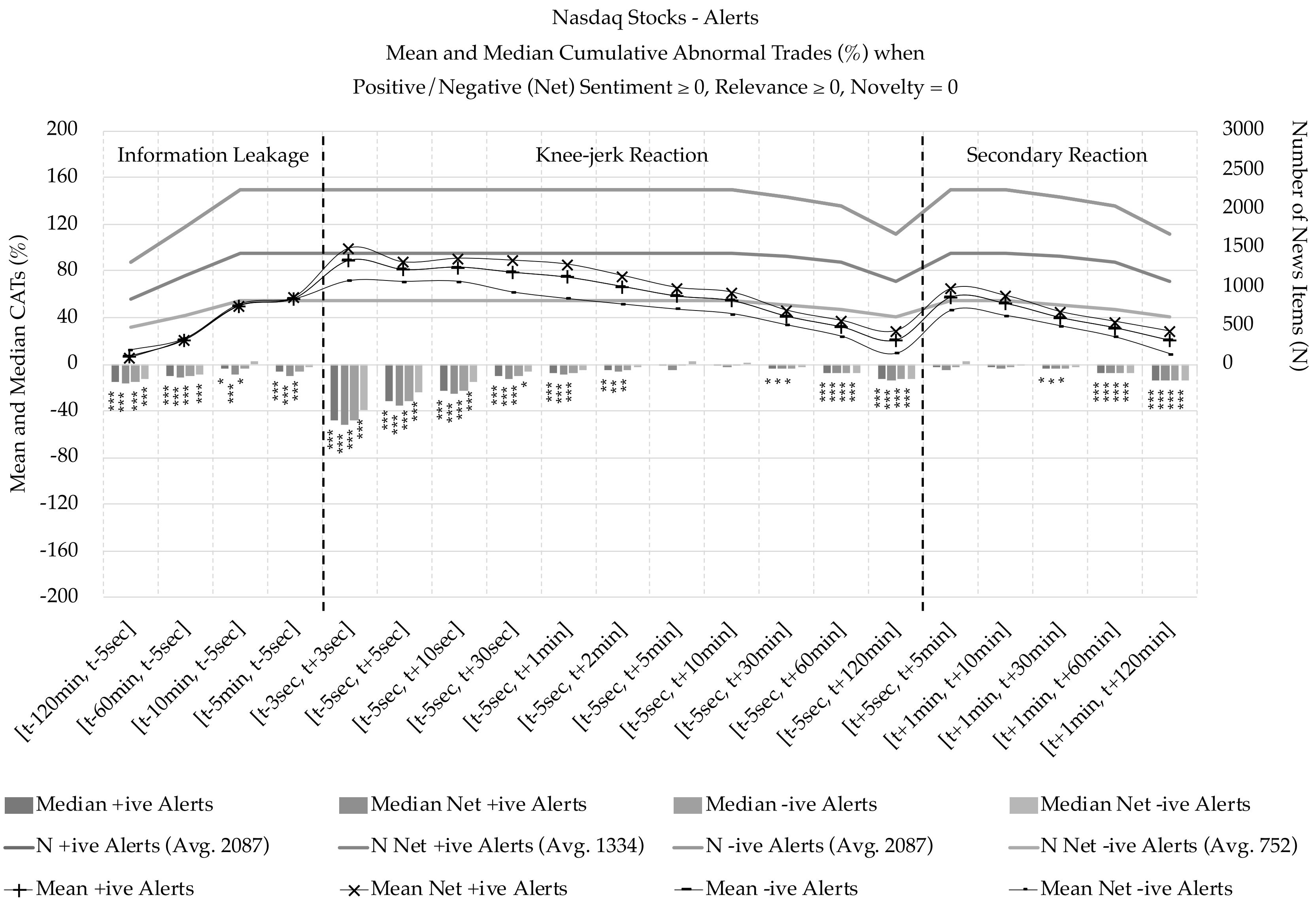


Figure A

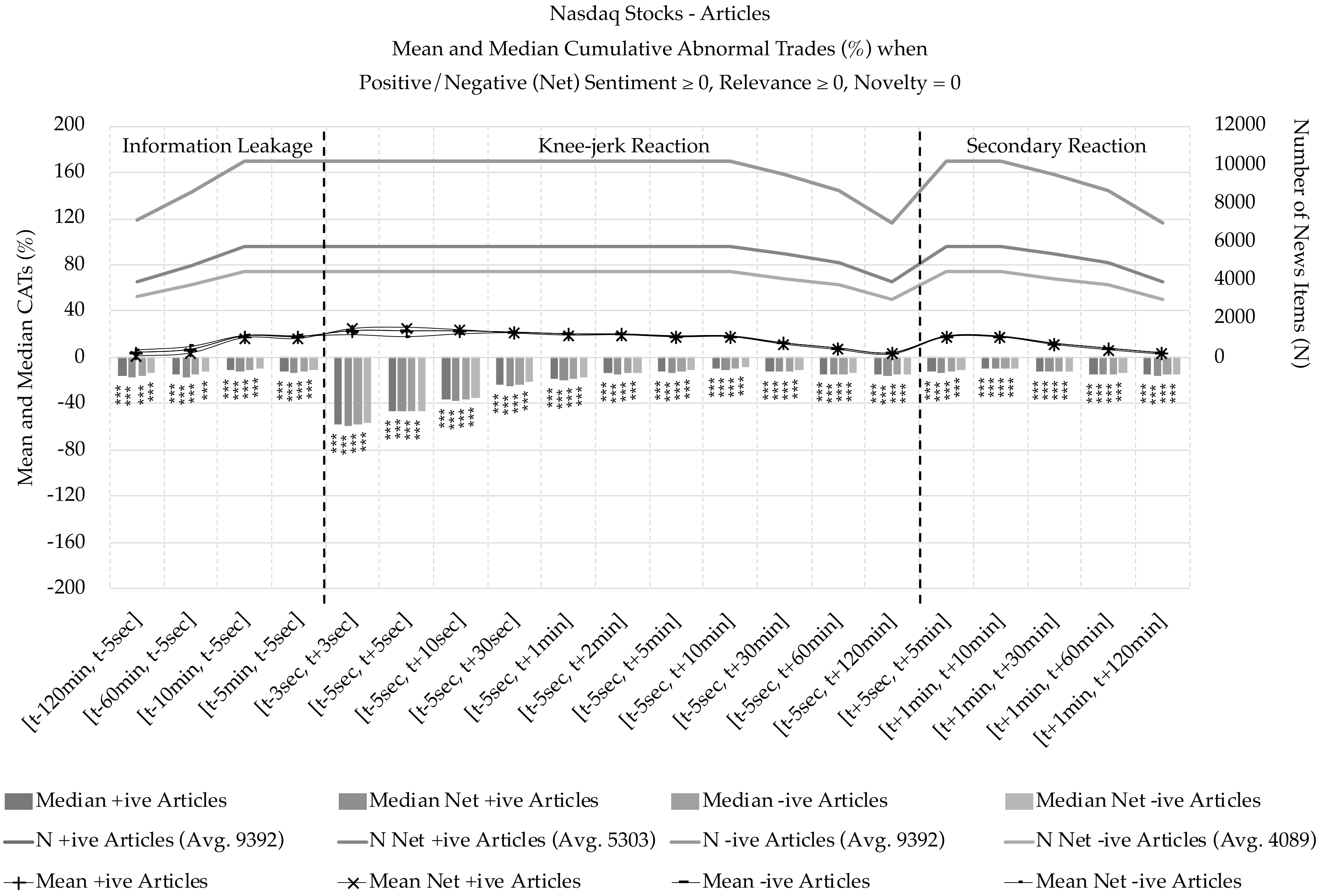
Mean and Median CATs for Novel Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

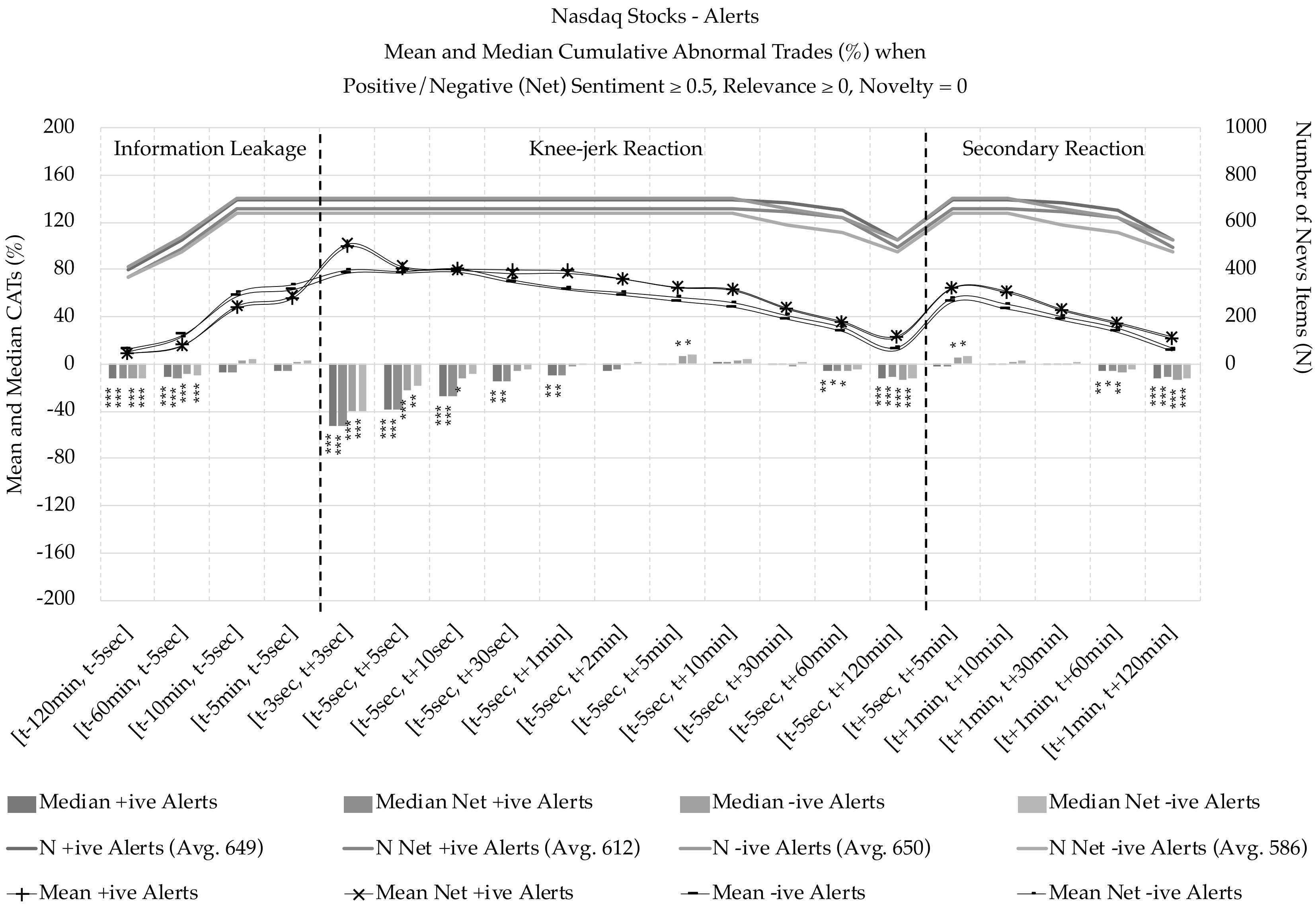
Mean and Median CATs for Novel Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

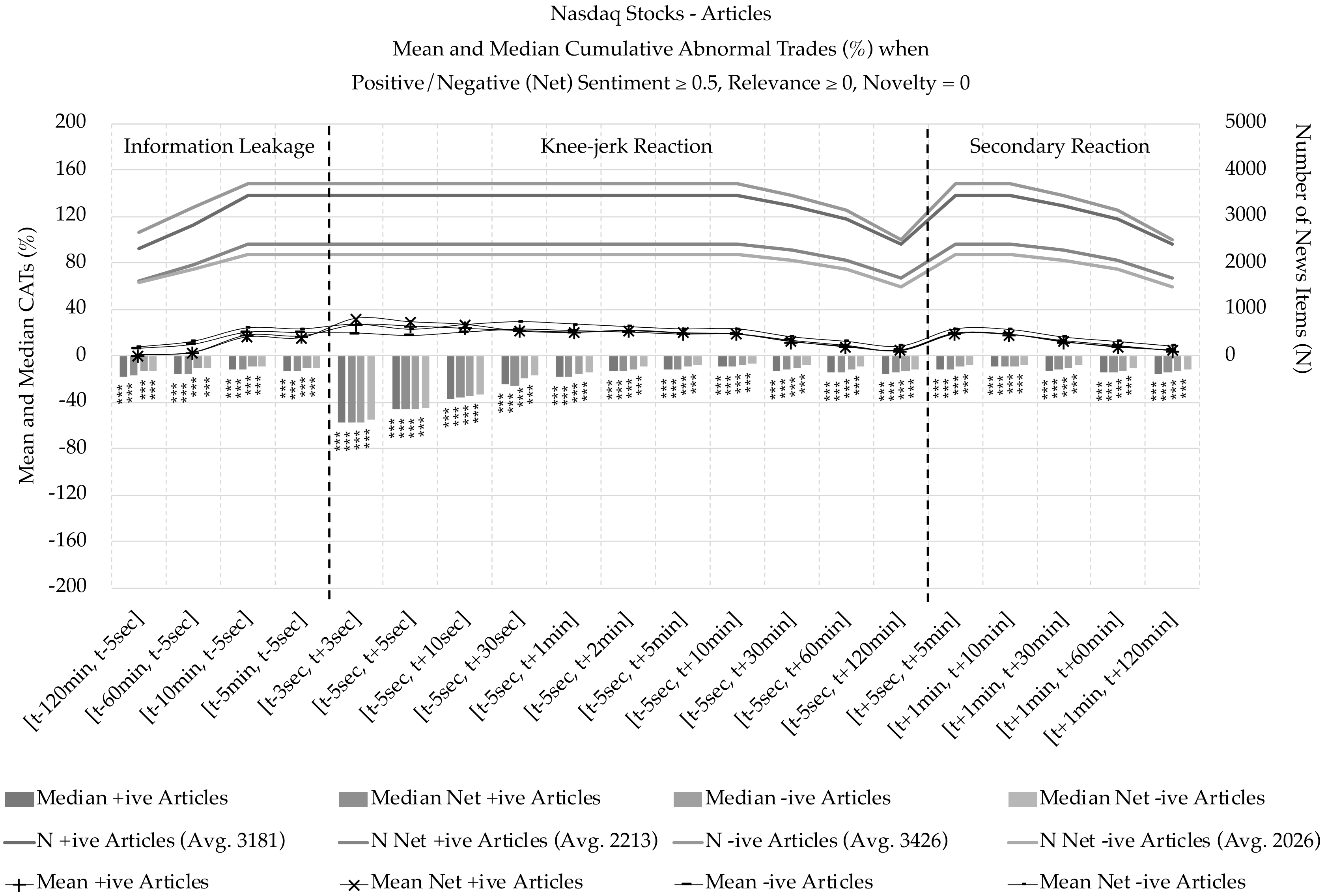
Mean and Median CATs for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

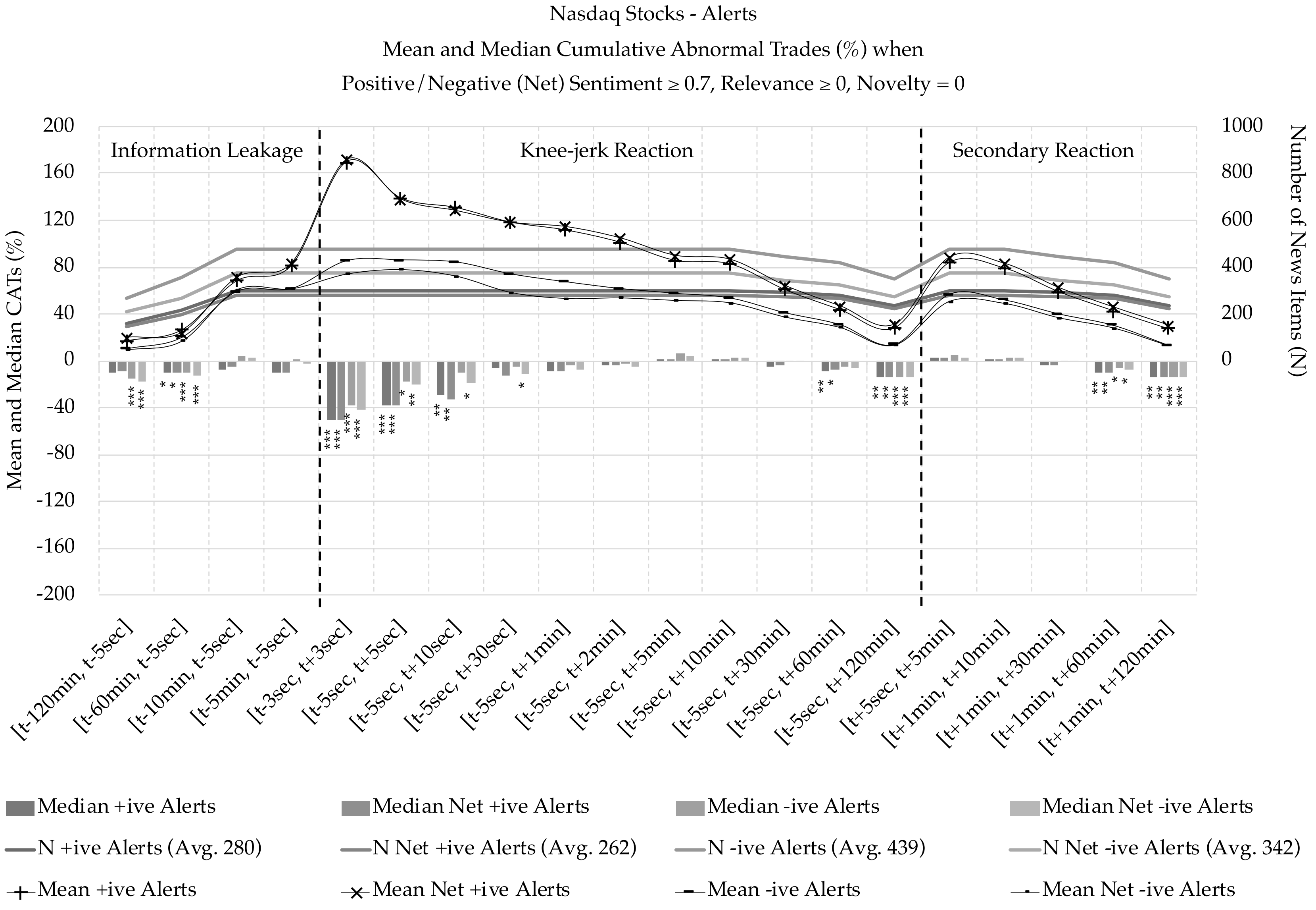
Mean and Median CATs for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

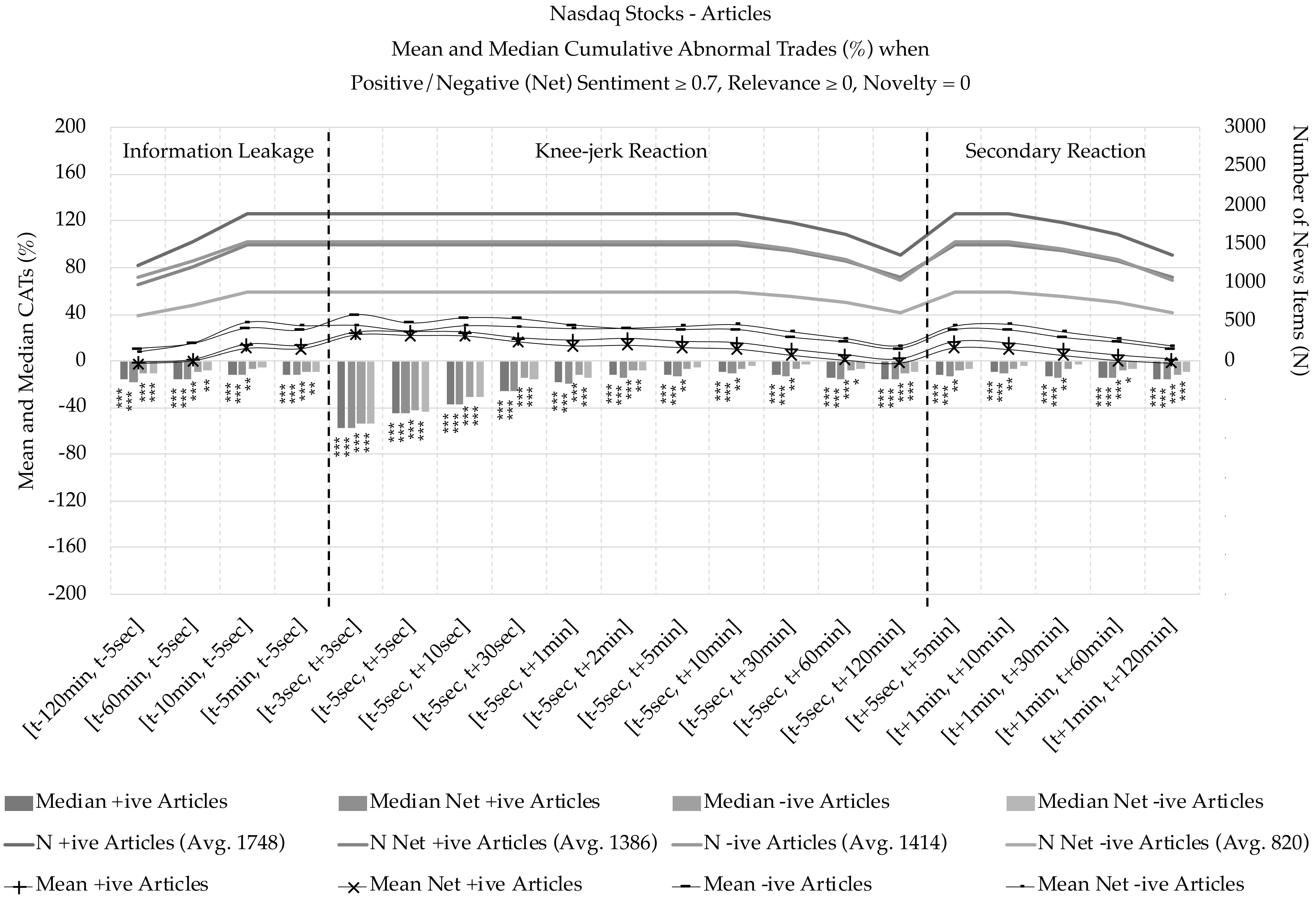
Mean and Median CATs for Novel Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CATs for Novel Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Appendix H – continued — *Additional Cumulative Abnormal Trade Results*

Mean and Median CATs for Nasdaq News across (Net) Sentiment Thresholds

---

The six charts that follow report the median and mean Cumulative Abnormal Trades (CATs) as well as the number of corresponding news items across all 20 event windows for: Positive News, Net Positive News, Negative News, and Net Negative News, respectively.

For this series of tests, no thresholds are set for Relevance and Novelty (all relevance and novelty scores are included), while absolute (Net) Sentiment thresholds are progressively increased from 0 to 0.5 (50 per cent) to 0.7 (70 per cent) for positive and negative news, per the flow chart below. Note that results for Alerts are reported in Figures A, C, and E, while Articles are reported in figures B, D, and F.

Mean and median CATs are measured in per cent above (below) the stock's 45-day moving average number of trades during market open. Significance is measured using the Sign Test: P < 0.05 *, P < 0.01 **, P < 0.001 ***.

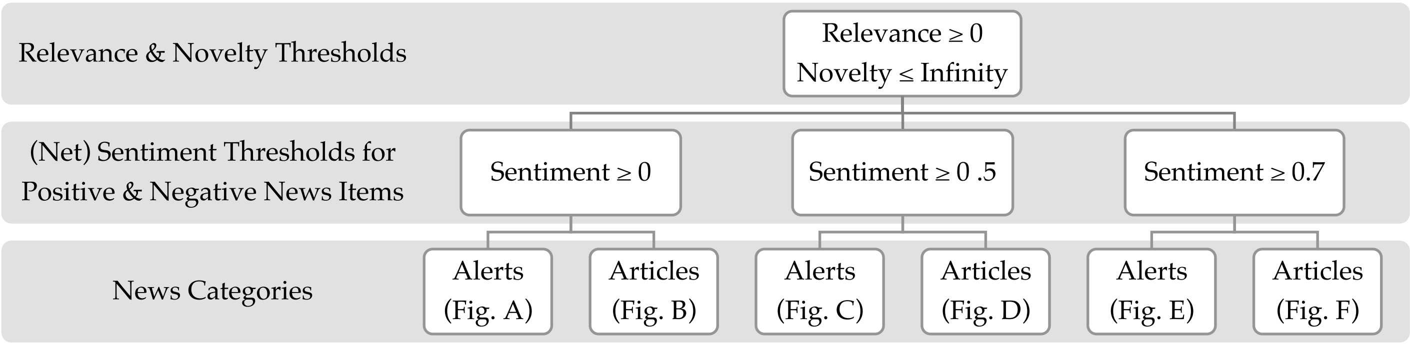
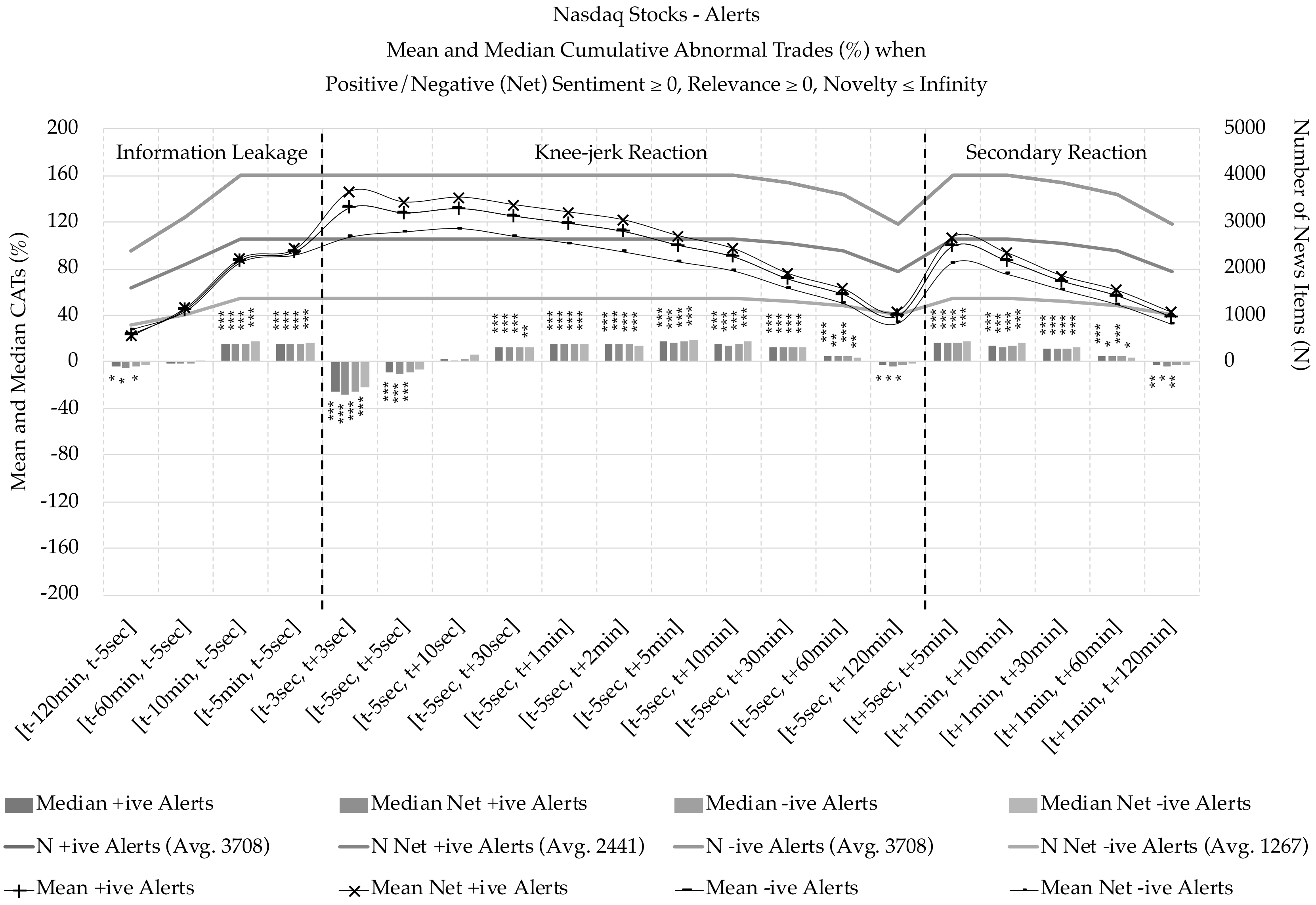


Figure A

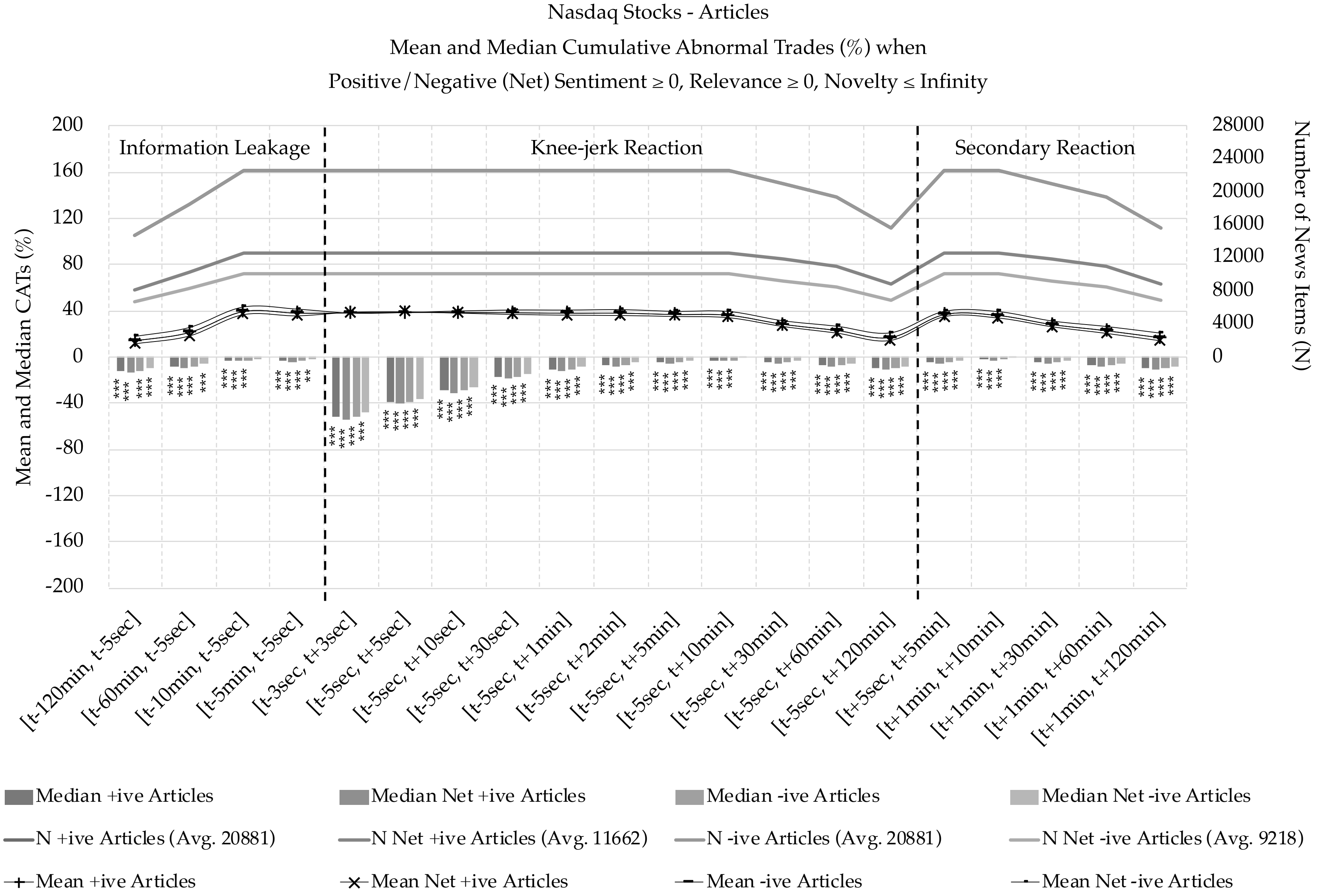
Mean and Median CATs for Nasdaq Alerts for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure B

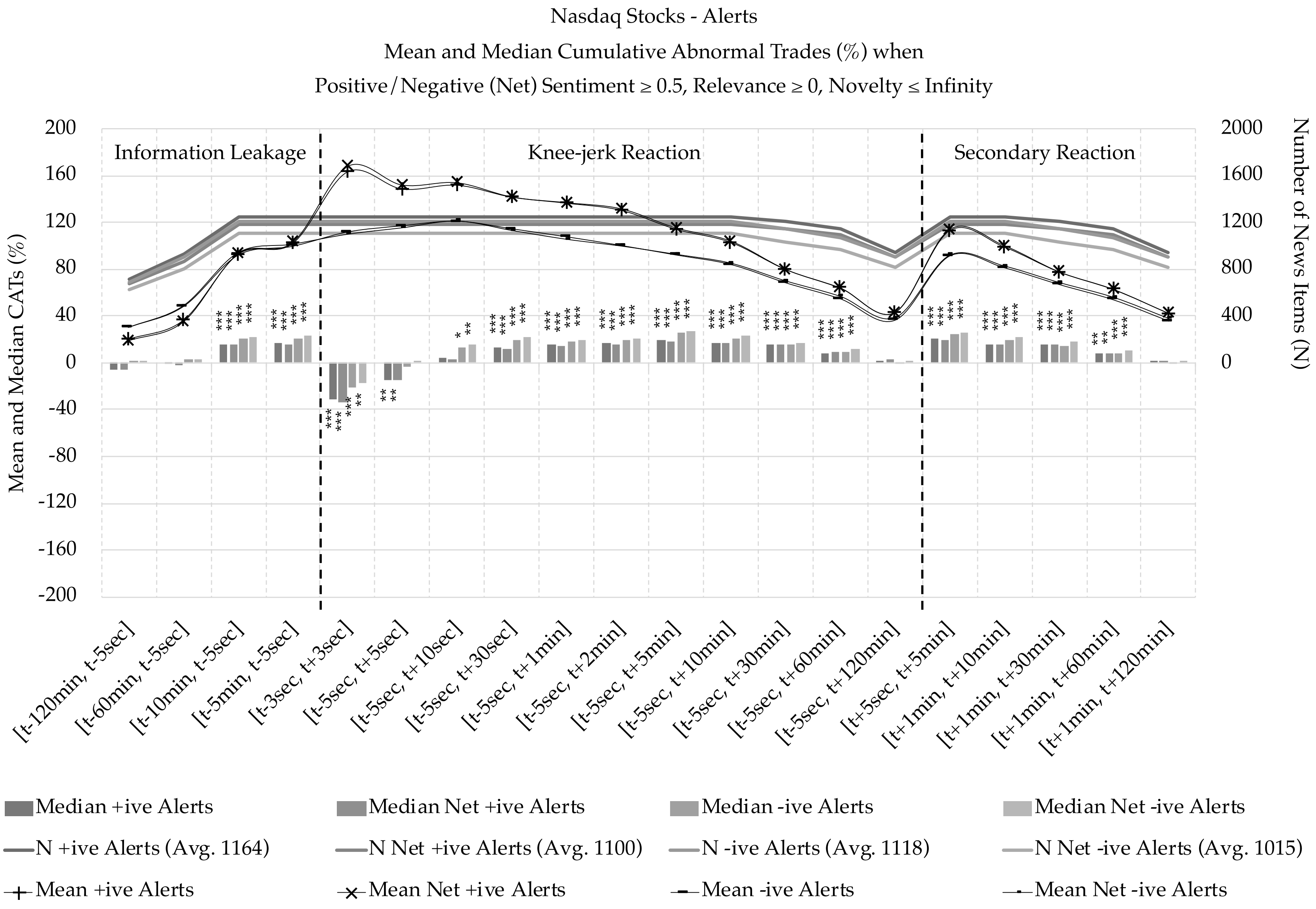
Mean and Median CATs for Nasdaq Articles for all (Net) Sentiment Values



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure C

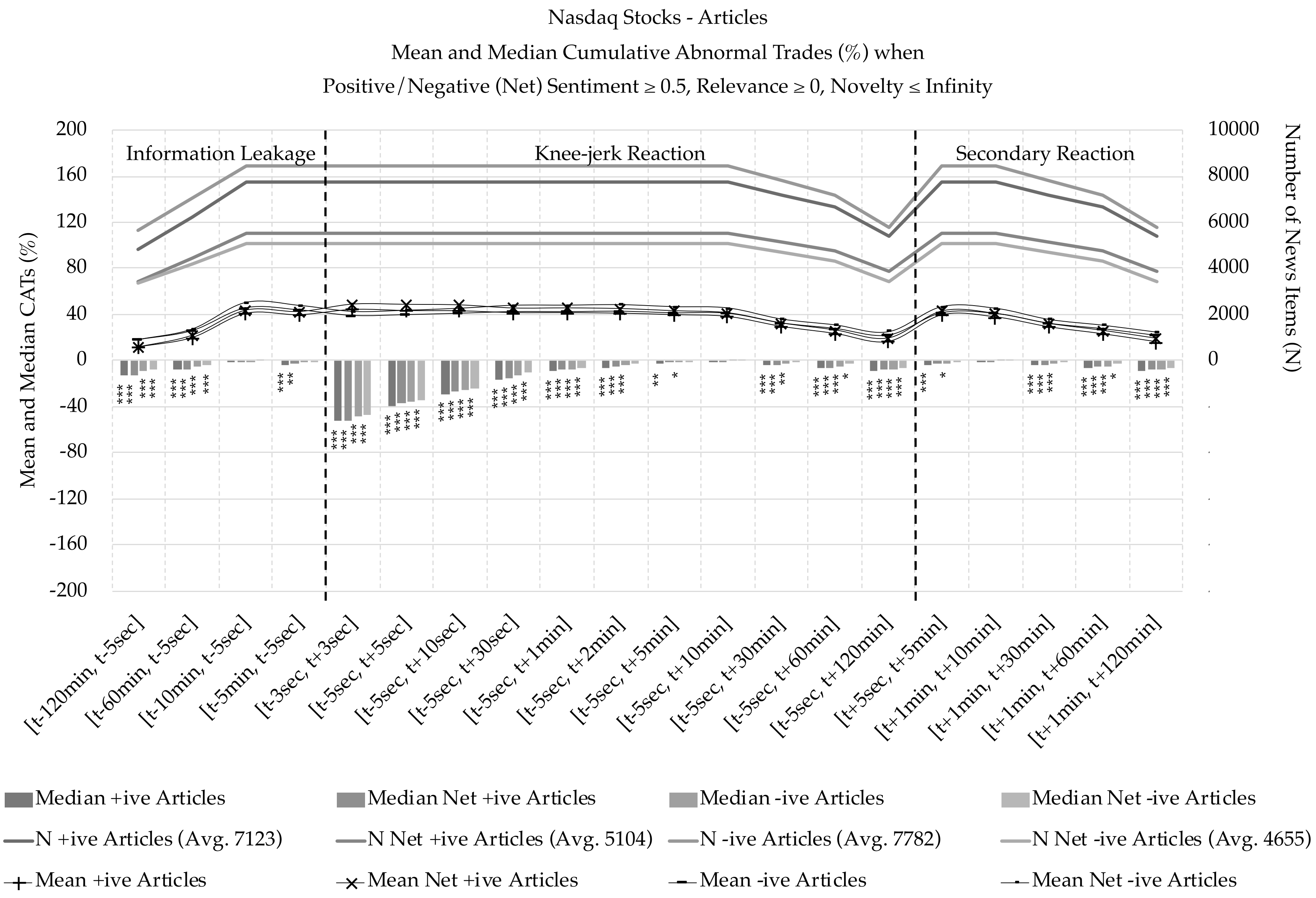
Mean and Median CATs for Nasdaq Alerts when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure D

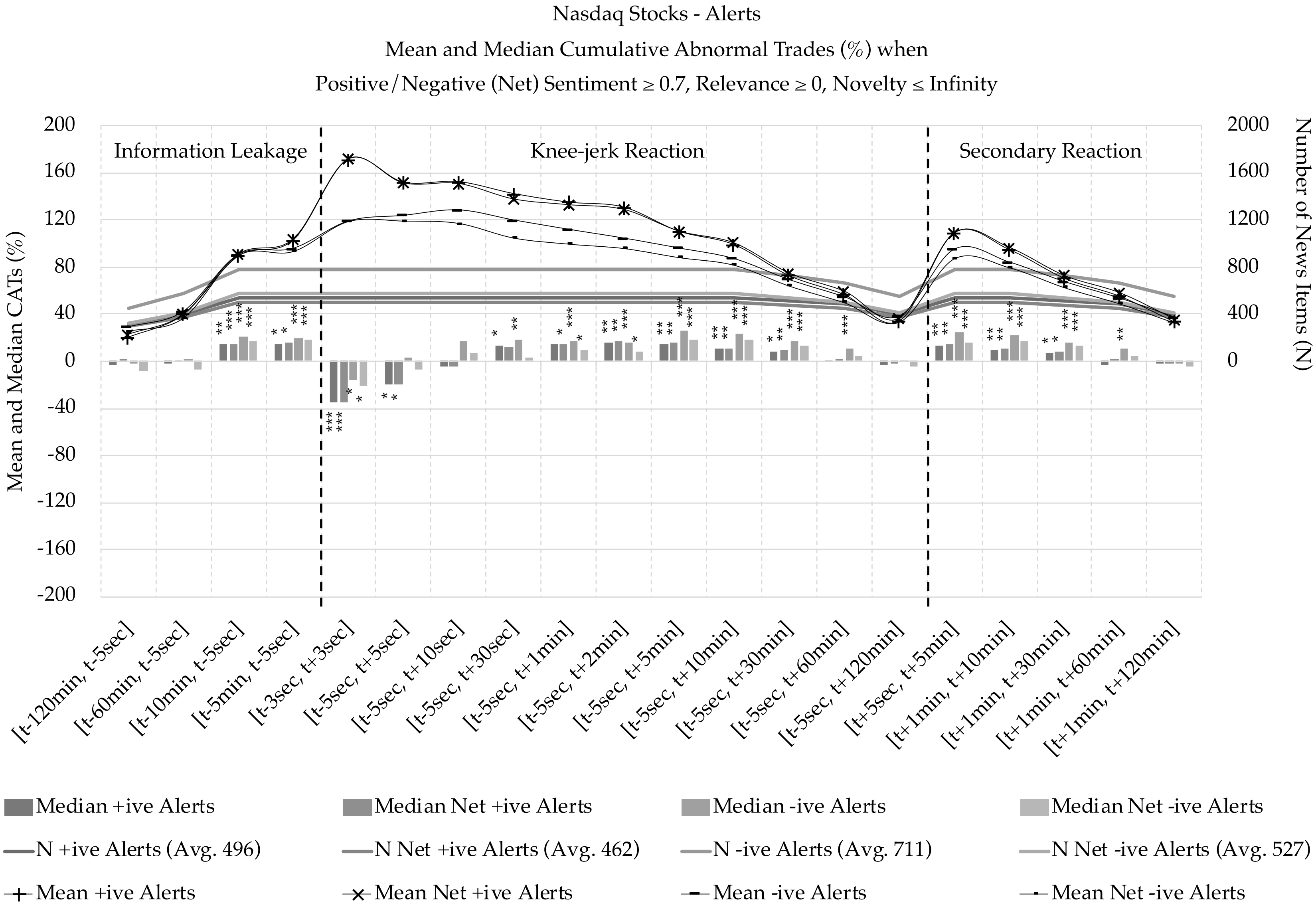
Mean and Median CATs for Nasdaq Articles when (Net) Sentiment ≥ 0.5



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure E

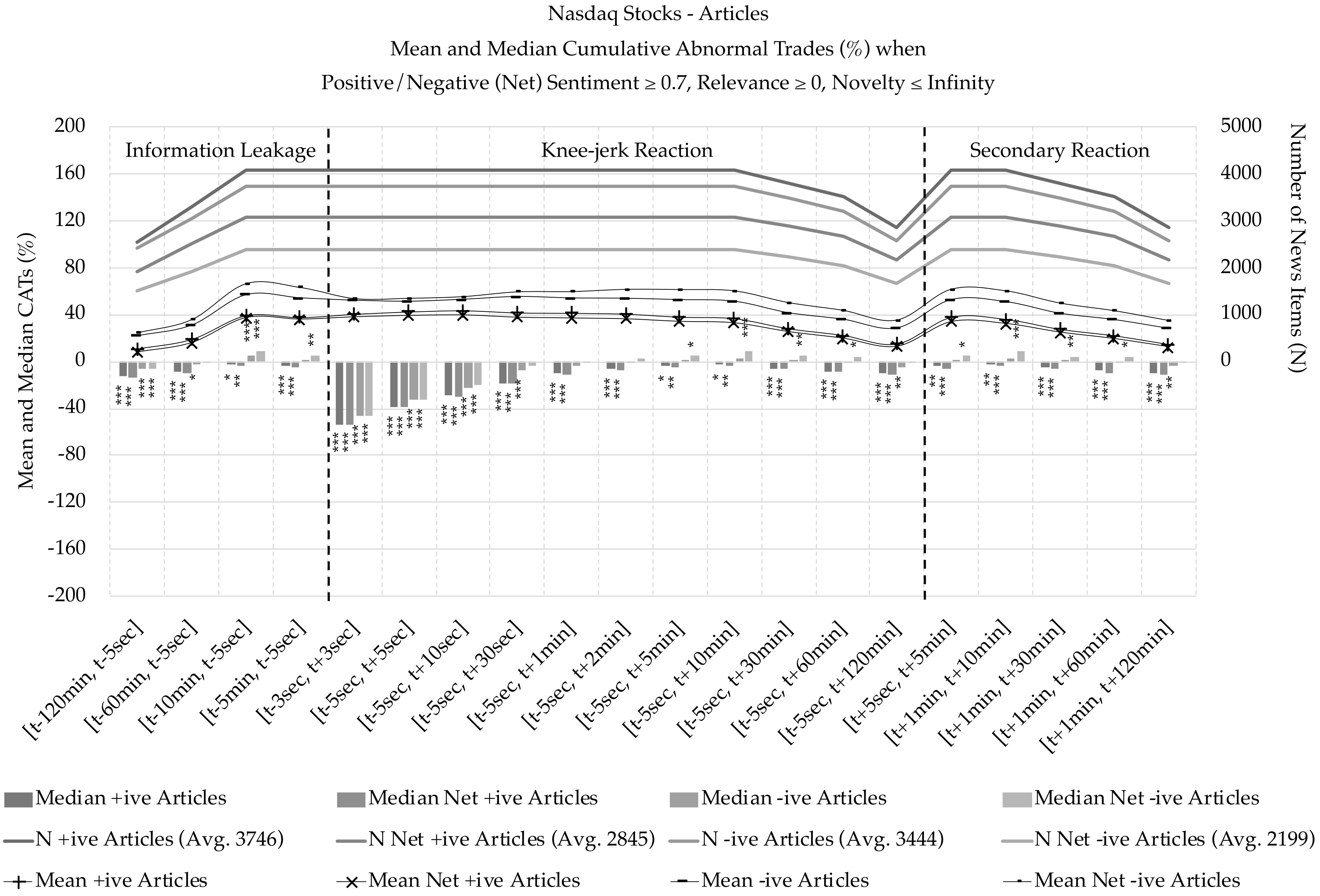
Mean and Median CATs for Nasdaq Alerts when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.

Figure F

Mean and Median CATs for Nasdaq Articles when (Net) Sentiment ≥ 0.7



Note significance is measured using the Sign Test where: $P < 0.05$ *, $P < 0.01$ **, $P < 0.001$ ***.