DISCUSSION PAPER SERIES

DP12871

CHASING PRIVATE INFORMATION

Marcin Kacperczyk and Emiliano Pagnotta

FINANCIAL ECONOMICS

ISSN 0265-8003

CHASING PRIVATE INFORMATION

Marcin Kacperczyk and Emiliano Pagnotta

Discussion Paper DP12871
Published 15 April 2018
Submitted 15 April 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CHASING PRIVATE INFORMATION

Abstract

Using over 5000 equity and option trades unequivocally based on nonpublic information about firm fundamentals, we find that widely used adverse selection signals display abnormal values on days with informed trading. Volatility and volume values are abnormally high, whereas illiquidity values are low, both in equity and options markets. Signals are more sensitive to informed trading in options markets and before unscheduled corporate announcements. We characterize cross-sectional responses based on the sign, type, and duration of private information. Evidence from the U.S. Securities and Exchange Commission (SEC) Whistleblower Reward Program addresses potential selection concerns.

JEL Classification: N/A

Keywords: N/A

Marcin Kacperczyk - m.kacperczyk@imperial.ac.uk
Imperial College London and CEPR

Emiliano Pagnotta - e.pagnotta@imperial.ac.uk
Imperial College London

Acknowledgements
For helpful comments and conversations we are grateful to Franklin Allen, Yakov Amihud, Alejandro Bernales, Bruno Biais, Utpal Bhattacharya, Ekkehart Boehmer, Alon Brav, Andrea Buraschi, Adrian Buss, Lauren Cohen, Slava Fos, Thierry Foucault, Ross Garon, John Griffin, Diane Del Guercio, Luigi Guiso, Terry Hendershott, David Hirshleifer, Travis Johnson, Andrei Kirilenko, Peter Koudijs, Craig Lewis, T-C Lin, Dong Lou, Albert Menkveld, Lubos Pastor, Maureen O'Hara, Lasse Pedersen, Ioanid Rosu, Gideon Saar, Norman Shurhoff, Zheng Sun, Paul Tetlock, Michela Verado, Pietro Veronesi, Marius Zoican, and seminar participants at the Tinbergen Institute, Imperial College Business School, MIT Sloan, Cass Business School, ISEAD, HEC Geneva, London Business School, Paris Dauphine University, Bocconi University, Loughborough University, University of Southampton, University of Nottingham, Tilburg University, the University of California, Irvine, Cheung Kong Business School, Tsinghua University, IDC Herzliya, the University of Texas at Austin, Einaudi Institute, SAC Capital, AQR Capital Management, and seminar participant at the following conferences: the Western Finance Association (Park City), the Society of Economic Dynamics (Toulouse), the CEPR First Annual Symposium in Financial Economics, the 2017 NBER Asset Pricing Summer Institute, the 2017 NBER New Developments in Long-Term Asset Management Conference, the 2017 FIRN Microstructure Conference (Sydney), the 2017 Conference of the Paul Woolley Centre for the Study of Capital Market Disfunctionality (London School of Economics), the Eighth Erasmus Liquidity Conference (Rotterdam), the 2017 FIRS Conference (Hong Kong), the Buy-Side Risk Europe, and the 2017 World Federation of Exchanges Conference (London). We are also grateful to Alon Brav, Hernan Ortiz-Molina, Amit Seru, and T-C Lin for their help with the data. Engin Iyidogan, Tianyu Wang, and Jingyu Zhang provided excellent research assistance.

Chasing Private Information*

Marcin Kacperczyk† and Emiliano S. Pagnotta‡

Abstract

Using over 5000 equity and option trades unequivocally based on nonpublic information about firm fundamentals, we find that widely used adverse selection signals display abnormal values on days with informed trading. Volatility and volume values are abnormally high, whereas illiquidity values are low, both in equity and options markets. Signals are more sensitive to informed trading in options markets and before unscheduled corporate announcements. We characterize cross-sectional responses based on the sign, type, and duration of private information. Evidence from the U.S. Securities and Exchange Commission (SEC) Whistleblower Reward Program addresses potential selection concerns.

This version: April 10, 2018

*The most recent version of this paper can be found here. For helpful comments and conversations we are grateful to Franklin Allen, Yakov Amihud, Alejandro Bernales, Bruno Biais, Utpal Bhattacharya, Ekkehart Boehmer, Alon Brav, Andrea Buraschi, Adrian Buss, Lauren Cohen, Slava Fos, Thierry Foucault, Ross Garon, John Griffin, Diane Del Guercio, Luigi Guiso, Terry Hendershott, David Hirshleifer, Travis Johnson, Andrei Kirilenko, Peter Koudijs, Craig Lewis, T-C Lin, Dong Lou, Albert Menkveld, Ľuboš Pàstor, Maureen O'Hara, Lasse Pedersen, Ioanid Rosu, Gideon Saar, Norman Shurhoff, Zheng Sun, Paul Tetlock, Michela Verado, Pietro Veronesi, Marius Zoican, and seminar participants at the Tinbergen Institute, Imperial College Business School, MIT Sloan, Cass Business School, ISEAD, HEC Geneva, London Business School, Paris Dauphine University, Bocconi University, Loughborough University, University of Southampton, University of Nottingham, Tilburg University, the University of California, Irvine, Cheung Kong Business School, Tsinghua University, IDC Herzliya, the University of Texas at Austin, Einaudi Institute, SAC Capital, AQR Capital Management, and seminar participant at the following conferences: the Western Finance Association (Park City), the Society of Economic Dynamics (Toulouse), the CEPR First Annual Symposium in Financial Economics, the 2017 NBER Asset Pricing Summer Institute, the 2017 NBER New Developments in Long-Term Asset Management Conference, the 2017 FIRN Microstructure Conference (Sydney), the 2017 Conference of the Paul Woolley Centre for the Study of Capital Market Disfunctionality (London School of Economics), the Eighth Erasmus Liquidity Conference (Rotterdam), the 2017 FIRS Conference (Hong Kong), the Buy-Side Risk Europe, and the 2017 World Federation of Exchanges Conference (London). We are also grateful to Alon Brav, Hernan Ortiz-Molina, Amit Seru, and T-C Lin for their help with the data. Engin Iyidogan, Tianyu Wang, and Jingyu Zhang provided excellent research assistance.

†Imperial College London and CEPR. Email: m.kacperczyk@imperial.ac.uk

‡Imperial College London. Email: e.pagnotta@imperial.ac.uk

1 Introduction

Asymmetric information is ubiquitous in financial markets, since investors have unequal knowledge of firm fundamentals. The presence of informed market participants is widely believed to affect the behavior of economic outcomes. At the same time, the challenge in empirically testing such links is that investors' information sets are almost never observable. Therefore, most tests rely on publicly observable proxies, information signals, such as volume or market prices, under the assumption that these signals bear a specific relation with the unobserved information asymmetry. For example, higher bid–ask spreads or trade price impact indicating higher risk of informed trading.[1] While information signals could provide useful guidance, the validity of such assumptions is not granted. Moreover, any statistical relationship one identifies could be spurious due to possible omitted variables. For example, changing prices could reflect time-varying risk premia or changing volume levels could be due to a systematic liquidity component or uninformed demand pressure. Hence, most empirical efforts to test the consequences of asymmetric information suffer from the joint hypothesis problem. Therefore, the empirical assessment of the reliability of information signals is difficult: it requires studying their distribution conditional on the actions of informed traders whose *unobserved* presence these signals *chase* in the first place.

We address this identification challenge by studying the conditional behavior of a broad set of information signals linked to trades that are *unequivocally based on nonpublic information* about firm fundamentals. Our inference is based on a hand-collected sample of 453 insider trading investigations of the U.S. Securities Exchange Commission (SEC) that document in detail how certain individuals trade on nonpublic and material information. For example, a hedge fund trader personally linked to a given firm's chief financial officer could privately learn about its exceptionally high quarterly earnings and acquire shares of the company in advance of the company's report.[2] These cases involve 5058 trades in 615 firms over the period 1995-2015. Hence, they are representative of a fairly large universe of assets and market conditions. We hypothesize that, if the presence of informed traders has a traceable impact on information signals, the signals should display abnormal behavior on precisely identified days with

[1] Various information-based trading theories argue that uninformed investors update their beliefs about the presence of informed trading based on publicly observed signals. Theories of learning from prices and trade flows originate from the seminal papers of Grossman (1976), Glosten and Milgrom (1985), and Kyle (1985).

[2] Our sample contains famous hedge fund trader cases that resemble this narrative (e.g., R. Rajaratnam at the Galleon fund in 2011 and M. Martoma at SAC Capital in 2012). The sample also contains individuals with small amounts of available capital, wealthy individual investors, and different institutional investors.

informed trading relative to a sample of random dates.

Guided by prior theoretical and empirical research, we consider three groups of information signals: those based on volatility, volume, and illiquidity. Although most empirical research has relied on stock market–based signals, informed traders could take advantage of options (Black (1975)). We additionally study the behavior of information signals in option markets.[3] We can identify not only the exact dates of informed trades, but also the trading instrument informed traders use. Thus, we can identify the presence of informed traders both across both time *and* markets. Standard theory also suggests that informed trading behavior could differ depending on whether private information has a random or a predictable end horizon (e.g., Back (1992); Caldentey and Stacchetti (2010)). We therefore hypothesize that the impact of informed trading on signals could differ with respect to scheduled or unscheduled events. Given that we observe the *content* of each trader's information set (e.g., earnings announcements or mergers), we can assess this hypothesis.

Our first result is that information signals largely display abnormal behavior on days with informed trading. The change in an average signal's value on days with informed trading is both statistically and economically significant. Moreover, we find that signals display stronger reactions on days with large informed trades. Second, information signals that originate in option markets are valuable. Not only do they exhibit abnormal behavior but also, relative to stock-based signals, they tend to be more sensitive to the presence of informed traders. Third, across stock and option markets, we observe consistent patterns in the direction of market response: volatility and abnormal volume increase while, contrary to common wisdom, illiquidity levels decrease. To illustrate the quantitative patterns and their economic significance, we show that, relative to their sample standard deviations, volatility measures—*Realized Variance* and *Price Range*—increase by 24.51% and 31.11% in the stock market and *Implied Volatility* for calls and puts increases by 6–7%. At the same time, *Quoted Spreads* decrease by about 10% in stock markets and by 20% in option markets. Fourth, information signals display stronger responses to the presence of informed traders during unscheduled events, in both stock and option markets.

We further evaluate salient economic characteristics of our sample. First, we calculate the informed volume and show that, on days when informed traders trade, their trades constitute more than 10% of

[3]Standard theory suggests that metrics reflecting volatility, volume, or illiquidity are impacted by the presence of informed traders and can thus act as information signals. Such signals can originate in stock markets (e.g., Kyle (1985); Glosten and Milgrom (1985); Easley and Hara (1987); and Wang (1993)) or in option markets (e.g., Back (1993); Biais and Hillion (1994); and Easley et al. (1998)).

the total amount for stocks, and more than 30% for options. Second, if price discovery takes place, one would expect prices to respond to the sign of informed trades. To this end, we compute the average raw and abnormal returns for the affected stocks on days with informed trading based on positive and negative news and find that the same-day average returns are 0.8% and -0.6%, respectively. Third, we assess the strength of the information individuals in our sample trade on by computing the hypothetical stock returns (excluding dividends) that an investor would realize if the investor initiated a trade at the opening price of the day of his or her first trade and closed the position at the opening price of the day following the public information disclosure. We show that, on average, such returns exceed 40% for private signals with a positive sign and 20% for those with a negative sign. The results are economically large,[4] given that they accrue over a relatively short period: The median number of days from the trade until the public announcement is seven, while the median number of days between the first and the last informed trade is eight. Overall, if information signals were not affected by the presence of such striking information asymmetries, one would arguably be even less inclined to hold the opposite view about the usefulness of a particular signal if asymmetries were smaller.

Since our sample includes solely trades originating from SEC investigations, one could be concerned about a potential sample selection bias. One pressing concern is that insider traders are exposed *only* when information signals display abnormal values, as if the SEC relied on detection technology that followed a similar set of results as those we document. If this were indeed the case, one would overestimate the ability of information signals to react to the presence of informed traders. We argue in Section 5 that such a scenario is unlikely. Many cases are investigated based on external referrals and not based on the SEC's screening of, say, illiquidity metrics. Even if such framework were in place, the results in Section 4 fail to support this view. For example, several signals display patterns that are arguably inconsistent with what economic reasoning would identify as patterns of informed trading, especially the fact that illiquidity signals have lower values when insiders trade. In other words, it is highly unlikely that the SEC is particularly sensitive to insider trading activity when markets look orderly and abnormally liquid. Besides, even if the agency intended to use public signals to flag an asset–date pair, it would be hard for its officials to identify systematically which individuals are breaching the law due to trade aggregation, netting, use of multiple accounts, and so forth.

[4] These figures likely underestimate the pre-fee profits from informed trading, since 30% of the trades in our sample are executed using options, not stocks.

We further address the problem of selection bias formally with three separate tests. The first test shows that the conditional behavior of information signals is not affected by the size of traders' profits that accrue over their trading horizon. Thus, selection based on the decision to litigate only highly profitable cases is unlikely. The second test indicates that the conditional behavior of signals is the same for investigations with a small and large number of firms. Using the argumentation of Meulbroek (1992), this result points against selection bias based on the origin of the investigation. Third, we research the origin of the investigation *directly* using the 2010 adoption of the SEC Whistleblower Reward Program (WRP), which offers monetary rewards to individuals who provide useful tips to uncover illegal insider trading. The identifying assumption of this test, as stipulated by the regulation, is that such tips cannot rely on publicly available data and must uncover independent new evidence. We show that the conditional behavior of public signals is mostly insensitive to the origin of investigations. Overall, our results strongly suggest that our findings are unlikely to merely reflect selection biases.

In Section 7, we additionally analyze the *content* of information sets, that is, whether informed traders obtained positive or negative news and whether information was related to scheduled (earnings) or unscheduled announcements, such as merger and acquisition (M&A) announcements. We find a similar conditional signal response for positive and negative private information, although stock-based (option-based) signals display a more pronounced decline in illiquidity for the case of negative (positive) information, which indicates that traders could rely relatively more heavily on short selling rather than on options when trading based on negative news. The differences across subsamples are more pronounced for different event types. We find that the stock- and option-based signals display stronger responses to the presence of informed traders ahead of unscheduled events, compared to scheduled events, such as earnings announcements.

To follow up on the last result, we conjecture that scheduled events, due to their predictability, could foster uninformed directional bets that introduce additional noise. To evaluate this possibility, we construct a sample of all earnings and M&A announcements covering our sample period. In Section 8, we study the time-series properties of stock- and option-based signals before and upon the announcement dates. We find that, with the exception of *Realized Variance* and *Price Range,* which increase the day before the announcement, stock- and option-based signals display virtually no abnormal behavior before public M&A announcements. For earnings announcements, we do not observe any pre-event reactions in stock-based signals. However, implied volatilities increase in the two-week period before

announcements, and option quoted spreads decrease one week before. To sum up, the results of this broader sample study suggest that (i) the patterns of signal behavior conditional on informed trades are largely different in the unconditional samples and (ii) for scheduled events, price-based option metrics are less reliable as information signals.

One of our key findings is the negative conditional response of illiquidity measures, in both stock and option markets. One explanation for the lower illiquidity on days of informed trading relies on the strategic behavior of informed investors. In particular, if trading costs are high due to temporary market illiquidity, an informed investor might want to time the execution of trades to minimize such costs, which would imply low illiquidity levels when better-informed investors trade (Admati and Pfleiderer (1988); Collin-Dufresne and Fos (2016)). To shed light on this hypothesis, we exploit a unique feature of our data: the ability to observe the dates *when private information is received.* We classify trades according to the length of their corresponding information horizons, that is, the time period between receiving the tip and the public announcement of that information. Arguably, for cases with short information horizons (at most three days), the ability to optimally time trades should be more constrained. The results in Section 9 support the timing hypothesis. Indeed, for short horizon cases, in contrast to the full sample, some illiquidity signals display near-zero or positive values. A second factor that could rationalize the conditional response of illiquidity measures is the use of limit orders by informed traders (e.g., Biais, Glosten, and Spatt (2005)). We screen the SEC investigations for the use of limit orders and find that, of the 85 cases with well-identified order types, 73% involve limit orders. Our statistical analyses suggest that the use of limit orders could be more prevalent for small-cap stocks, since their quoted spreads and order imbalance measures display significantly lower values.[5]

Taken together, our study has important implications for the literature that studies the economic consequences of asymmetric information for corporate finance and asset prices and for the literature that examines information content in stocks and options. We provide a more detailed discussion of these implications in Section 10. Very few studies have addressed the issue of whether different types of information signals help in the identification of informed traders across markets.[6] Our ability to

[5]This interpretation is consistent with the theoretical model of Baruch, Panayides, and Venkataraman (2017), who argue that limit orders are more likely to be used by informed traders in the case of stocks for which short selling is difficult (typically small-cap stocks).

[6]We note that not all empirical analyses in the literature rely on information signals of the type we analyze. An alternative approach that proved useful in several settings is to study the trades of a particular set of traders. For

observe the arrival of private information directly is in stark contrast to prior literature that infers the presence of informed trading indirectly, for example, by observing the trading behavior of certain groups, such as large activist shareholders, or analyzing trades immediately before corporate events. Collin-Dufresne and Fos (2015) identify a negative relation between stock-level trading volume of SEC Schedule13D filers and liquidity measures. Our findings regarding illiquidity measures in stocks are consistent with these authors' results.[7] Our ability to observe the arrival of private information allows us to provide direct evidence that strategic timing contributes to the prima facie counterintuitive illiquidity effect. Moreover, the granularity of our data enables us to study the use of informed limit orders. In contrast to trades by 13D filers, we can observe the content of information sets and document the specific responses of signals to both positive and negative information and before both scheduled and unscheduled events. Meulbroek (1992) was the first to use the information in SEC insider trading investigations and studied stock market efficiency. Our focus, instead, is on the joint distribution of information signals and the (usually unobserved) presence of informed traders, in both stock and option markets. In contrast to Collin-Dufresne et al. (2017), we document that option markets are used extensively by informed traders over the 20-year period we analyze. Therefore, our results provide support to the conjecture of Black (1975) and the theoretical analyses of Back (1993), Biais and Hillion (1994), and Easley et al. (1998). Moreover, our results show that the conditional patterns of signal behavior in option markets are consistent with those observed in stock markets.

2 Main Test and Information Signals

This section first describes the empirical environment of our study and the specification of our main test. Next, it describes the information signals used in the analysis. For brevity, we relegate several details of the data implementation of each signal to Appendix A of the Online Supplement, where we

example, Boulatov, Hendershott, and Livdan (2013) and Hendershott, Livdan, and Schurhoff (2015) use institutional order flow and Cohen, Malloy, and Pomorski, 2012 study routine corporate insiders SEC filings.

[7] We note, however, that informed traders in SEC litigation files and SEC Schedule 13D activist investors are not directly comparable. From an information structure viewpoint, activist investors may act on the belief that they are privately informed, but without specific knowledge of a particular corporate event or fundamentals different from their own equity position. Indeed, an average activist investor faces long-lasting uncertainty regarding whether the activist investor's efforts will be fruitful. Large stock purchases by activist investors, of course, could have a positive price impact on the stock return and even induce herding if other participants anticipate future positive price pressure. Second, from a strategic viewpoint, the incentives of activist investors may not be representative of the classical profit-maximizing individuals in informed trading models, but, arguably, are more closely tied to long-term corporate control. Consequently, for example, one can rationalize why 13D filers de-emphasize option markets.

provide summary sample statistics.

2.1 A Stylized Framework of Asymmetric Information in Empirical Studies

Consider a firm, a period, and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over which random variables $\{Y, I\}$ are defined, with $Y : \Omega \to \mathbb{R}$ denoting an economic variable of interest (e.g., a return or specific corporate decision) and $I : \Omega \to \{0, 1\}$ denoting the presence ($I = 1$) or absence ($I = 0$) of traders who are privately informed about that firm's fundamentals. A theory of asymmetric information can be seen as a proposition that implies that $\mathbb{P}(Y|I) \neq \mathbb{P}(Y)$, for example, $\mathbf{P}$, $\mathbb{E}(Y|I = 1) > \mathbb{E}(Y)$. An econometrician is interested in testing $\mathbf{P}$; however, I is not observable. The econometrician then considers one or more information signals $S : \Omega \to \mathbb{R}$ and makes an identification assumption, $\mathbf{A}$, that relates I and S. Typically, $\mathbf{A}$ can be expressed as $S = \text{constant} + \Delta I + \text{noise}$, $\Delta \in \mathbb{R}$. If $\Delta > 0$ is assumed, the econometrician could determine a threshold value $\hat{s}$ and replace I by $\hat{I}(\omega) = \mathbb{I}\{S(\omega) > \hat{s}\}$, where $\mathbb{I}$ is the indicator function. Proposition $\mathbf{P}$ is validated by the test if $\hat{\mathbb{E}}(Y|\hat{I} = 1) > \hat{\mathbb{E}}(Y)$, where $\hat{\mathbb{E}}(Y)$ is the empirical counterpart of $\mathbb{E}(Y)$.[8]

Ultimately, however, the empirical examination is a joint test of $\mathbf{P}$ and $\mathbf{A}$: One could observe $\hat{\mathbb{E}}(Y|\hat{I} = 1) > \hat{\mathbb{E}}(Y)$ while *both* $\mathbf{P}$ and $\mathbf{A}$ are false. Consider, for instance, $\mathbf{A}$ stipulating that $\Delta > 0$, while the actual population relations are $\mathbb{E}[Y|I = 1] < \mathbb{E}[Y]$ and $\Delta < 0$.

2.2 A Test on the Reliability of Information Signals

To shed light on the reliability of the identification assumption, $\mathbf{A}$, one would ideally design a test on $\mathbb{E}(S|I)$. Let S be the value of a given information signal. By the law of iterated expectations,

$$\mathbb{E}\left[S\right] = \mathbb{E}\left[S|I = 1\right] \times \mathbb{P}(I = 1) + \mathbb{E}\left[S|I = 0\right](1 - \mathbb{P}(I = 1)).$$

As is common in the literature (e.g., sEasley and O'Hara (1992)), we assume that privately informed traders are not permanently trading every single period, so $\mathbb{P}(I = 1) < 1$. Thus, the probability of informed trading (PIN) is less than one, which implies $\mathbb{E}\left[S|I = 1\right] \neq \mathbb{E}\left[S\right]$. We can then succinctly

[8] In a similar fashion, assumption $\mathbf{A}$ could be stated as $\frac{\partial}{\partial p_i}\mathbb{E}(S|p_i) > 0$, with $p_i := \mathbb{P}(I_i = 1)$. We note that an econometrician might be interested in the distribution of I itself, for example, evaluating the extent of adverse selection risk for the considered firm. Similarly, given the unobservability of I, the econometrician would rely on assumptions on the joint distribution of (S, I).

express the value of the signal as

$$\mathbb{E}\left[S|I=1\right] = \mathbb{E}\left[S\right] + \Delta, \tag{1}$$

where Δ represents the impact of informed trading on the signal value relative to its unconditional mean.[9] We hypothesize that, if the firm-specific signal S captures the presence of privately informed traders, it should display abnormal behavior on days when such traders enter the market. The baseline test on the reliability of $\mathbf{A}$ for signal S thus has a null hypothesis $\Delta = 0$ against an alternative hypothesis, $\Delta \neq 0$.

A test of Δ in equation (1) is generally not feasible, given the unobservability of I (which motivated the use of S in the first place). We fix the problem by utilizing a new sample in which we can observe the presence of informed traders for a given asset on a given day (see Section 3). We define the binary variable *Info Trade* to be equal to one when an informed trade (sees Section 4 for details). The unique ability of the regulatory agency to document the use of material nonpublic information and to time-stamp it precisely for a given asset allows us to rely on a fundamental connection between observables and unobservables, that is, Info Trade $= 1 \Rightarrow I = 1$, and therefore to empirically assess the relative degree of a signal's reliability.

2.3 Signals of Private Information–Based Trading

Our tests study the behavior of the following types of signals: volatility, volume, and illiquidity. For volatility, we consider the realized variance using 30-minute returns, the daily price range, and price informativeness (e.g., Durnev et al. (2004)) in stock markets and call/put implied volatility and implied volatility skewness (e.g., Cremers and Weinbaum (2010)) in option markets. We construct a measure of abnormal volume for both stock and option markets as the residual from a prediction model of the daily volume. Moreover, based on Black's (1975) insight that informed traders value leverage, we compute the ratio of out-of-the-money (OTM) options to the total options volume. We use the following illiquidity signals: the percentage quoted bid–ask spread (*Quoted Spread*); the five-minute price impact (e.g., Goyenko et al. (2009)), *Price Impact*; the absolute order flow imbalance,[10] *Order*

[9] We note that $\mathbb{E}\left[S|I=0\right]$ is hard to identify in a non-laboratory setting, since identification would require access to *every* trader's information set.

[10] We follow Easley et al. (2008) and Holden and Jacobsen (2014) and use the absolute order imbalance as a proxy for the usual *PIN* measure. We do not use the *PIN* measure directly, since it is a cross-sectional estimate defined over a relatively long time period. Using the absolute order flow imbalance has two distinct advantages: First, it can be computed over short periods, such as a day. Second, it does not have the numerical overflow problems that can arise

Table I
Matrix of Information Signals

Signal Type/Market	Stocks	Stock options
Volatility	Price Range, Realized Variance Price Informativeness	Implied Volatility (calls and puts) IV Skewness
Volume	Abnormal S. Volume	Abnormal O. Volume Volume Ratio (OTM/all)
Illiquidity	Quoted Spread, Price Impact, Order Imbalance, Lambda, S. Illiq	Quoted Spread (all options, OTM) O. Illiq

Imb..; Kyle's lambda (e.g., Hasbrouck (2009)) in stock markets, *Lambda*; and the quoted bid–ask spread for all options as well as for OTM options. In addition, we use a daily version of Amihud's (2002) *Illiq* measure for both stocks and options (*S. Illiq* and *O. Illiq*, respectively). Table I summarizes the set of information signals. Unless otherwise stated, in the remainder of the paper we refer to an information signal as a signal, or simply S.

3 Insider Trading Sample

In this section, we provide background information on insider trading cases and discuss the construction of our data. We further relate instances of insider trading to aggregate market activity.

3.1 Background

The term *insider trading* refers to both legal and illegal conduct. The legal variety is when corporate insiders—officers, directors, large shareholders, and employees—buy and sell stock in their own companies and report their trades to the SEC. On the other hand, illegal insider trading refers to buying or selling a security in breach of a fiduciary duty or other relationship of trust and confidence while in possession of material nonpublic information about the security.

The legal framework prohibiting insider trading dates back to Rule 10b-5 of the Securities Exchange Act of 1934. Under the classical view of insider trading, a trader violates Rule 10b-5 if trading on material nonpublic information about a firm to which the trader owes a fiduciary duty, where information is deemed material if a reasonable investor would consider it important in deciding whether to buy or sell securities. Alternative interpretations of what constitutes illegal insider trading activity

when computing the *PIN* log-likelihood function.

continue to be made to this day. We do not seek to settle this debate here. In fact, whether a given trade is formally illegal or not is not important to us. Rather, our identification strategy relies on two conditions: (i) the trade under consideration is motivated by actual information, as opposed to, say, sentiment, and (ii) the material information is not widely available to market participants at the time of the trade. This approach allows us to concentrate on all investigations for which the SEC reported that conditions (i) and (ii) were met, regardless of their legal classification.[11]

3.2 Data Collection

We retrieve the list of SEC investigations from all SEC press releases that contain the term *insider trading* and use it to obtain all the available civil complaint files available on the SEC website. In cases in which a complaint file is not available on the SEC website, we rely on manual web searches and on information from the U.S. District Court where the case was filed. We collect all files starting from January 2001 until December 2015. We track all documents that provide updates on a previously released legal case. Whenever updated information is made available at a later date, we rely on the most recent version.

The resulting sample represents all SEC cases that were either litigated or settled out of court. Most complaint files include a detailed account of the allegations. Since the documents provide most of the relevant information in a textual form, the data files must be thoroughly read and summarized by hand. Available information typically includes the biographical records of the defendants, individual trades, a description of the leak to which the trades are linked, and the relationships between the tippers and the tippees.

We organize the information by characterizing trades and information events. A *trade* is any single transaction record for which we can observe a date and a trading instrument (e.g., stocks or options). For most trades, information about the price, trade direction, quantity, trading profits, the closing date of the position, as well as contract characteristics for options is also available. An *information event* is a collection of one or more trades that were motivated by a unique piece of private information, such as an earnings announcement or a merger. For our purpose, the key information event records include the firms involved, the nature of the leaked information (e.g., a new product), and the date the information

[11] Furthermore, for a significant proportion of investigations, legal resolution is a monetary settlement with the SEC. It is difficult to infer from such resolutions whether the defendant is guilty or would rather pay a fine than legally contest the regulator.

was released to the general public. We also collect information on the date of information transmission from the tipper to the tippee. This information allows us to test our hypotheses on strategic trading delays.

3.3 Descriptive Statistics

Our data cover 453 cases. The most frequent event types are M&As (55.90%) and earnings announcements (15.06%). The remaining cases correspond to several types of business events, such as information about products, firm projects, patents, Food and Drug Administration medical trials, corporate restructuring, bankruptcy, and fraud. The average number of cases per year in our sample is 30.83, with a maximum number of 46 filed in 2012. The distribution of the number of firms per case is highly asymmetric. Approximately 80% of the cases involve a single firm and 4% of the cases involve 10 or more firms.

Table II summarizes our data at the trade level. We identify s 5,058 unique trades involving 615 firms. Panel A shows the distribution of trades with respect to the trading instrument. The vast majority of trades are executed via stocks (67.06%) and options (31.83%). The remaining few are trades in American Depositary Shares and bonds. Panel B shows the breakdown of trades with regard to trade direction. There are 4,220 buys (83.43%) and 838 sells. Even though the SEC litigation files date back to 2001, they involve trades that took place earlier, spanning the period 1995–2015. The sample is quite evenly distributed over time, with over 100 trades in each year between 1999 and 2014, although we observe a smaller number of trades earlier, in the 1990s.Trades are dispersed across many different industries. The three most represented industry sectors in our sample are chemicals, business services, and electronic equipment, which account for more than 40% of all trades. We note that the trading involves companies from almost all industrial sectors.

A distinct feature of our data is the independent information on the date of information arrival and the date of its use. Panel C of Table II shows that the median time between the arrival and the use of information by insiders is two days, with a significant variation of 24 days in the data. In turn, the median number of days from a trade until the public announcement of information is seven. The median horizon between the first and the last trades is eight days. The median trader in the sample executes 10 trades, with a maximum of 97 trades. A median firm is traded 16 times and a median legal case involves two firms. The median age of tippers and traders is almost identical and equals 45

years; the vast majority are male.

Our sample contains both small retail investors and professional investors. We find that at least 60% of them have some finance background or work for financial firms and 30% of them are highly ranked corporate executives (vice-president level or higher). Hence, an economically meaningful fraction of them are capable of relatively sophisticated trading (e.g., using stock derivatives). The reported profits are highly skewed, with an average trade profit of $1.01 million and a median of $90,000. Over 49% of trades elicit at least $100,000 in profits.

3.4 Informed Volume and the Information Content of Trades

A relevant aspect of our data is the amount of trading carried out by informed traders. We construct this statistic by aggregating all informed trades in a given firm on a given day, separately for stocks and for options. Panel A of Table III shows that informed trades make up a significant percentage of the total trades in the market. On average, 10% of the daily volume in stocks and more than 30% of the option volume is traded by informed traders.

If price discovery takes place, one would expect prices to respond to informed trades. In particular, the direction of price movements should, on average, be consistent with the sign of private information that motivates the trades. To evaluate this connection, we compute the average raw and abnormal returns for the affected stocks on days of informed trading. Panel B of Table III shows the results. The average return on days with positive information is slightly over 0.8% and that on days with negative information is nearly -0.6%. The magnitudes are similar when we use market-adjusted returns, which suggests that market activity is not far from normal on such days. Therefore, consistent with previous findings by Meulbroek (1992), stock returns seem to respond to these information-motivated trades.

Besides the same-day impact of trades, we also explore how *material* the received information is. In other words, we evaluate the strength of the information content. To do so, for each information event, we compute the percentage change in the corresponding stock price from the opening of the day of the informed trade to the opening price immediately after the information becomes public. Setting the trading window in such a way ensures that the arrival of public information is contained within its range. Panel C of Table III shows the results for the aggregate sample and each news type. For positive news, the average and median returns are approximately 43.5% and 33.7%. These values are remarkably large, given that the median period from a trade to private information disclosure is merely

seven days. Arguably, one could treat these numbers as a lower bound of the true signal strength, since about 30% of trades are in options, thus embedding leverage. To put these numbers in perspective, we construct benchmark returns for a sample of SEC 13D filers, who are often regarded as informed, between 1994 and 2014.[12] The benchmark return is based on the return measured from the opening of the day when the 13D filer trades an asset until the opening of the day following the release of the trade information to the public. The trades of 13D filers represent large long positions in a security and have been shown to predict positive stock returns, so they can be interpreted as being based on positive news (e.g., Brav, Jiang, and Kim (2015); Collin-Dufresne and Fos (2015)). The mean and median returns for 13D filers are 4.9% and 2.4%, respectively.

4 Full-Sample Results

In this section, we present our baseline empirical results. We first describe the construction of our empirical counterparts to the components in equation (1) for stock- and option-based information signals. We then describe evidence on Δ using both a univariate time-series approach and a multivariate cross-sectional regression specification.

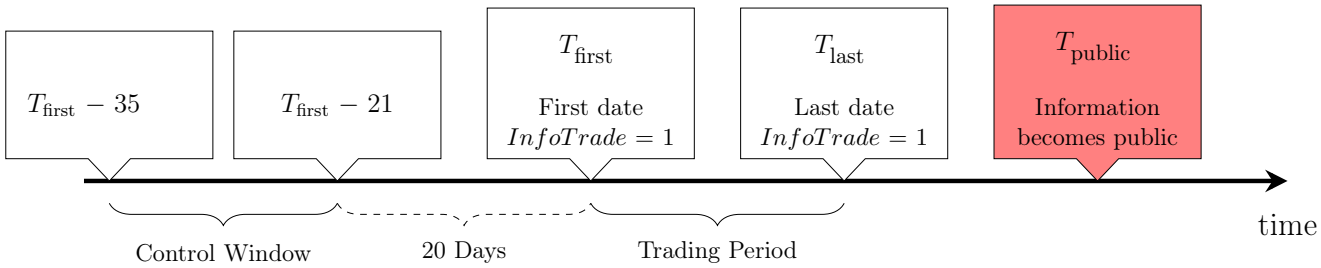
4.1 Test Design and Univariate Time-Series Evidence

Motivated by equation (1), we seek to compare the value of a given signal, S, on days with informed trading, $\mathbb{E}[S|I = 1]$, with its unconditional expected value, $\mathbb{E}[S]$. We begin by constructing a formal measure of informed trading, $InfoTrade$. For a given *trading instrument* i in the sample—an individual stock or a stock option—we identify the set of dates $\{T_{i,\text{first}}, ..., T_{i,\text{last}}\}$ on which trades motivated by the *same* piece of private information occurred. We set $InfoTrade_{it} = 1$ if and only if $t \in \{T_{i,\text{first}}, ..., T_{i,\text{last}}\}$. In cases in which only $T_{i,\text{first}}$ and/or $T_{i,\text{last}}$ are reported, we do not set $InfoTrade_{it} = 1$ for the dates in between, since, for those dates, the use of private information is not precisely verifiable. Next, we estimate $\mathbb{E}[S|I = 1]$ by conditioning the average value of S on $InfoTrade_{it} = 1$.

To define a set of *normal* dates for which $InfoTrade_{it} = 0$ to estimate $\mathbb{E}[S]$, we focus on a narrow window of 15 days close to T_{first}, which insulates us from any longer-term trends driving the data. In

[12] We thank Alon Brav for providing us the data.

Figure 1. Event Timeline



particular, we define a control period spanning 21–35 trading days before T_{first} (see Figure 1). That is, while we count each case of $\text{InfoTrade}_{it} = 1$ with $t \in \{T_{i,\text{first}}, ..., T_{i,\text{last}}\}$ as a separate observation, we use only the pre-event window that corresponds to the earliest of the trades, T_{first}. By skipping the last 20 days before T_{first} in the control window, we aim to reduce the likelihood that $\mathbb{E}[S|InfoTrade = 0]$ differs from the unconditional mean of S, as it would be the case if, for example, unidentified informed trades for the same stock occurred very close to T_{first}.[13] Last, we eliminate all cases in which informed trades occur less than four days prior to scheduled corporate events announcements (mostly earnings announcements) to avoid capturing the effect of (predictable) directional bets on the announcements that may not be motivated by private information.

To obtain a first glance of the main test results, we investigate the time-series behavior of the signals and the potential presence of pre-trends. In Figures 5 and 6, we display the average values of several stock- and option-based signals over the control window $[T_{\text{first}} - 35, T_{\text{first}} - 21]$ and on days with informed trading (date 0 in each figure). This univariate analysis offers several useful insights. First, we observe that information signals do display abnormal behavior on days with $InfoTrade = 1$. The abnormal behavior is observed in both stock and option markets. Across stock and option markets, we observe consistent patterns in signal response: volatility and abnormal volume measures increase while, perhaps surprisingly, illiquidity measures decrease. Second, we do not observe any clear time trends over the control window, which suggests that information-driven trades are unlikely to occur in the short period preceding informed trades. In other words, the average signal value in the control window can approximate the corresponding unconditional mean. We want to stress that the behavior

[13] If the likelihood of informed trades were indeed higher near dates with InfoTrade = 1, using a control window such as $[T_{\text{first}} - 20, T_{\text{first}} - 1]$ would likely downward bias the estimates of Δ in absolute terms. Although the specific choice of the control window is ultimately somewhat arbitrary, robustness tests suggest that our results are not highly sensitive to changes in its specification.

of signal values on $InfoTrade = 1$ dates is not a result of any particular announcement, since those occur strictly before the public release of information on date T_{public} and, of course, informed traders do not publicly announce their trades when they trade.

Even though the time-series results are indicative of an abnormal response for various signals, the patterns in the data may be influenced by firm- and time-specific cohort-invariant effects. Moreover, they do not provide a precise account of the statistical significance of the observed patterns. To address these limitations, we further investigate the role of privately informed trades and that of potential confounding factors in a formal regression test.

4.2 Results from a Regression Design

To conduct formal tests on Δ, we estimate the following multivariate regression model:

$$S_{it} = c \times Controls_{it} + d_i + e_t + \Delta \times InfoTrade_{it} + \varepsilon_{it}, \tag{2}$$

where *Controls* is a vector of firm-specific controls, including the natural logarithm of firm market capitalization ($LNSIZE$), the natural logarithm of trading volume ($LNVOL$), the equity price per share (PRC), and stock turnover ($TURNOVER$). To eliminate the problem of bad controls, we use their pre-determined values, defined as follows: for any given information event on firm i, containing observations on dates $[T_{i,\text{first}} - 35, T_{i,\text{first}} - 21] \cup \{T_{\text{first}}, ..., T_{\text{last}}\}$, the value of each control variable is fixed at its realization on date $T_{i,\text{first}} - 35$.[14] We also include firm fixed effects d to account for any source of time-invariant unobserved heterogeneity. The fact that our analysis is based on daily data within a short horizon makes it unlikely that any common time-series trend differentially affect a given signal on days with $InfoTrade$=1 and days in the control window. Nonetheless, to account for the possibility that information signals vary generically over time, we consider time fixed effects e. To avoid effects from extreme outliers, we winsorize all signals at 1%. We cluster standard errors at the firm level to account for serial correlation in residuals.

[14] Since informed trading is unlikely to be related to firm characteristics, the choice of control variables is largely meant to reduce the noise in our empirical estimates. In fact, as shown in Table C7 in the Online Supplement, we obtain essentially the same Δ estimates using a specification with no controls.

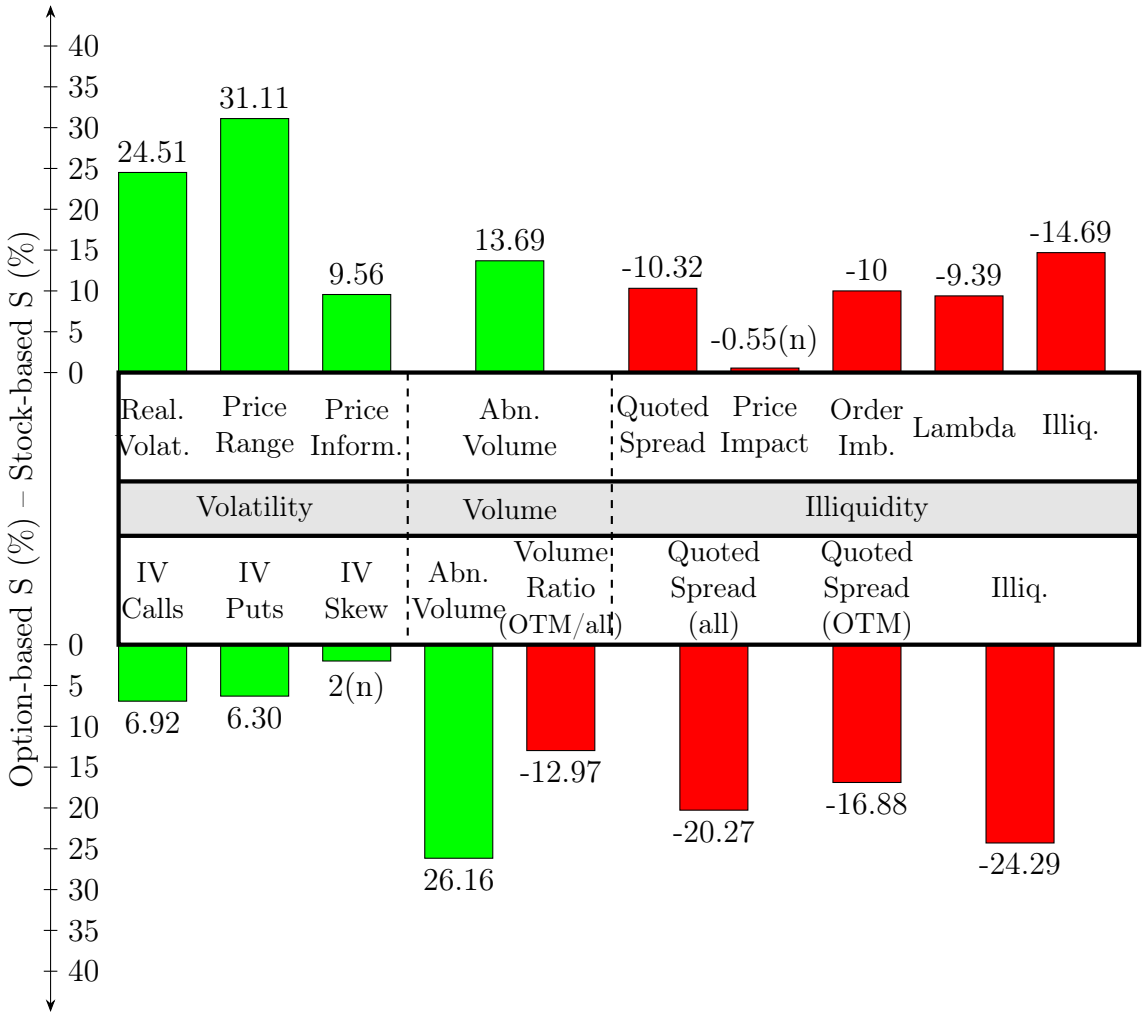


Figure 2. Main Empirical Design: Results Summary

For each information signal, the numbers displayed in the bars correspond to the percentage ratio between the estimated coefficient on days with precisely identified informed trading and the corresponding full-sample standard deviation. Green (red) columns correspond to positive (negative) effects on the signal value. Estimates that are not statistically significant are indicated by '(n)'.

In Table IV, we present the results from estimating the regression model (2) for stock-based signals. To provide some perspective on economic magnitudes, given the different units in which S are measured, the top panel of Figure 2 shows the estimated coefficients relative to the corresponding S sample standard deviation as a percentage. First, we confirm that information signals do display abnormal behavior on days when $InfoTrade = 1$. The picture shows that the changes are both statistically and economically significant. Second, all three volatility measures increase on days when informed traders enter the market.The variables *Realized Variance* and *Price Range* increase by 24.51% and 31.11%, respectively, and *Price Informativeness* also displays abnormally higher values. Third, *Abnormal S*.

Volume is 13.69% higher and, consistent with the univariate time-series observation, all illiquidity measures display lower values. For example, *Quoted Spread* is, on average, 10.32% lower. Two illiquidity measures that are based on order flow, *Order Imb* and *Lambda*, display values that are 10% and 9.4% lower, respectively. The quantitative negative effect for *S. Illiq* is larger, at -14.69%. We do not observe a significant change in the five-minute price impact measure.

Table V presents the results for option-based signals. The bottom panel of Figure 2 displays changes in signal values relative to their respective standard deviations. As in the case of stocks, most option-based signals display abnormal behavior on days with $InfoTrade = 1$. The qualitative patterns are consistent with those observed in stocks as well. The variable IV is higher for both calls and puts.[15] We observe no significant abnormality in *IV skewness*. In turn, *Abnormal O. Volume* is around 26.16% higher. Notably, the relative volume of OTM options is lower on $InfoTrade = 1$ days. Hence, *Vol. Ratio* values are approximately 13% lower. The lower values of *Quoted Spread* (all) and *O. Illiq* suggest that illiquidity is lower, on average, when informed traders trade.

Overall, our results indicate that information signals display traceably different behavior on days with $InfoTrade = 1$, in both stock and option markets. These changes are both statistically and economically significant. Finally, we observe common patterns of behavior across different signal types: volatility and volume signals display higher values while illiquidity signals display lower values.

4.3 Sensitivity to Trading Intensity

The evidence in Section 4.2 suggests that signals display different behaviors on days when privately informed investors trade. It is thus reasonable to hypothesize that a larger share of informed trading has a greater impact on such signals. To explore this relation, we split both stock and option trades into low-intensity (high-intensity) trades. We define trades as low (high) intensity if the respective informed trades are below (above) the within-asset class median. We next estimate the regression model (2) for all information signals conditional on low- and high-intensity trades. Table VI displays the results.

Even though the qualitative results for each subsample are not different from those of the uncondi-

[15] Our findings on volatility are based on the rich cross-section of option contracts with different maturities and moneyness. We aggregate these results by weighting observations with respect to the corresponding open interest. To establish the robustness of our results, we have also considered alternative specifications in which weights are proportional to option volume and option vega. The results from these tests are qualitatively similar in terms of their direction and economic magnitudes.

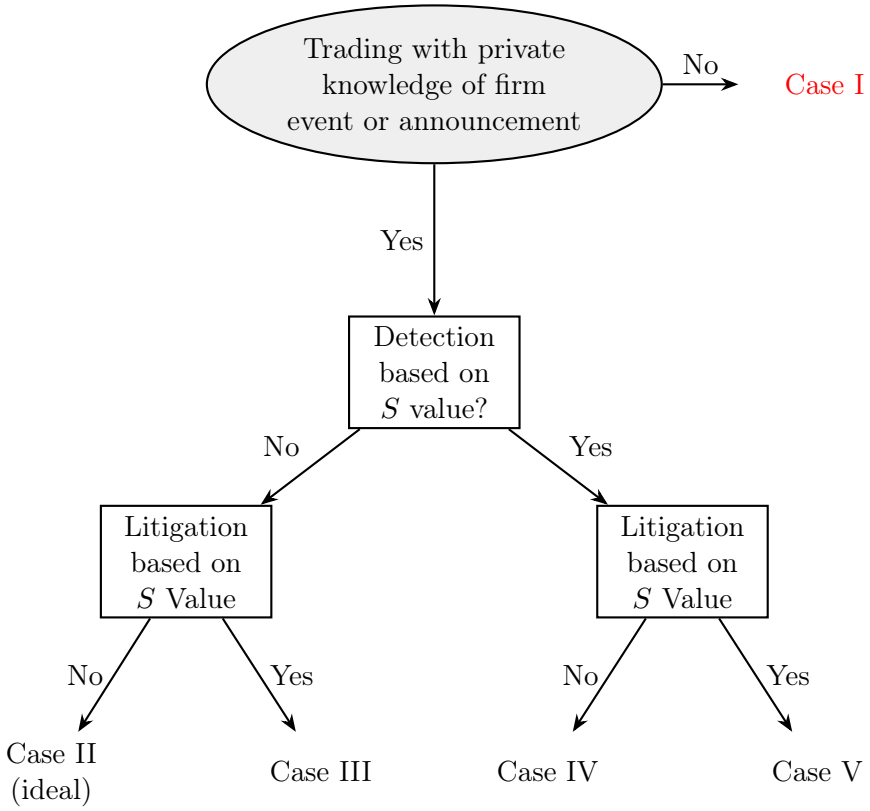
tional sample, the associated t-statistics are, generally, higher for high-intensity cases in both stock and option markets. In fact, some information signals are only statistically significant for high-intensity trades. We also observe that, with the exception of *Abnormal O. Volume*, when coefficients are statistically significant in both subsamples, their absolute values are larger for high-intensity trades. These results are consistent with the view that a larger informed trading participation has a greater impact on the signal value.

5 Evaluating Potential Sources of Selection Bias

One of the primary empirical challenges in evaluating the impact of informed trading on market signals is the unobservability of investors' information. The estimation of Δ in equation (1) using trades from SEC investigations has a distinct advantage that stems from the rather unique opportunity of accessing individual traders' information sets. Therefore, one can: (i) eliminate uncertainty about whether traders had private information about firm fundamentals, (ii) provide precise informed trading dates, (iii) identify in which market the trade took place, and (iv) connect each trade with the specifics of the information, that is, what is that the informed traders knew and when. In contrast, most alternative approaches in the literature do not rely on access to individual information sets but, rather, on assumptions that either certain groups of traders possess private information or certain information events are more likely related to informed trading. Under such circumstances, in the context of our framework, one can misclassify pairs of asset–date pairs (i, t) for which $InfoTrade_{it} = 1$.

However, we do not claim that our setting is free of any identification concerns. Instead, we consider the possibility of a selection bias explicitly and evaluate the potential effects that distinct selection sources could have on our baseline results. Figure 3 illustrates various conceptual cases of our analysis. As a first step, we note that we can precisely identify information-based trading, given the fact that regulators can document the use of advance knowledge regarding a given corporate event. This allows us to abstract from Case I. In the subsequent steps, we focus on the role of institutional processes, such as the decision to advance a litigation against a seemingly guilty trader and the detection of such a trader. The ideal benchmark is represented by Case II, a situation in which both institutional processes are exogenous with respect to the value of S. In the remainder of this section and the next, we analyze the possibility of selection biases in either the litigation or detection phase.

Figure 3. **Institutional Setting and Potential Selection Sources**



5.1 Litigation

Consider first the litigation decision, $\mathbb{I}_{\text{lit}} \in \{0, 1\}$. From the SEC files, for a given company and period, one can compute $\mathbb{E}[S|I = 1, \mathbb{I}_{\text{lit}} = 1]$, that is, values of information signals conditional on the case being litigated. By the definition of conditional expectation, these can be expressed as $\frac{\mathbb{E}[S \times \mathbb{I}_{\text{lit}}|I=1]}{\mathbb{P}(\mathbb{I}_{\text{lit}}|I=1)}$. One does not face an identification challenge if the information signal and litigation decision are unrelated, that is, $\mathbb{E}[S \times \mathbb{I}_{\text{lit}}|I = 1] = \mathbb{E}[S|I = 1]\mathbb{E}[\mathbb{I}_{\text{lit}}|I = 1]$. Otherwise, if cases included in SEC filings are litigated based on the behavior of information signals, as represented by Cases III and V in Figure 3, we could impound a bias. We argue that the decision to litigate is driven by factors other than the value of the considered signal. Among these, first, is the availability of external evidence—such as emails, phone calls, or wiretaps—confirming a violation of Rule 10b-5 of the Securities Exchange Act of 1934. Such evidence is likely to be uncorrelated with S. Second, we have factors related to individual characteristics, such as repeated rogue trading behavior or the fact that the individual involved committed other securities-related violations. If repeated offenses were indeed considered, $\mathbb{I}_{\text{lit}}$ could increase the representation of

investigations involving multiple firms, but not necessarily the estimates of Δ.

Third, $\mathbb{I}_{\text{lit}}$ could be influenced by the behavior of asset prices when information is publicly announced (date T_{public} in Figure 1). A large price jump upon the announcement could help the SEC officials to claim that the information was *material*. If this was the case for some investigations, the sample could over-represent cases for which the value of information is high. However, for any signal about fundamental firm value that a trader receives, whether one observes a large price reaction on T_{public} largely depends on the prior actions of the informed trader (e.g., Kyle (1985); Back and Baruch (2004)). Therefore, if this type of selection exists, we could be oversampling prudent traders, that is, those who could have traded more aggressively to impound information into prices. Under such circumstances, this type of selection would work against the considered signal set displaying abnormal behavior on days with $InfoTrade = 1$.

To obtain further perspective on whether the value of information affects our baseline results, we turn to regression analysis. Formally, we define a variable *Strength* that, as in Table III, measures the percentage returns from the opening price on T_{first} to the opening price on day $T_{\text{public}} + 1$. We estimate the model in equation (2) with an interaction term between *Strength* and $InfoTrade$ as the main variable of interest. If *Strength* has a monotonic effect on S, it should be captured by the interaction term. The estimation results in Table B3 of the Online Supplement suggest the opposite. We observe no systematic pattern in the coefficients and the vast majority of them are statistically insignificant. Hence, even if the SEC screens cases based on their profitability, that selection does not seem to correlate with our results.

Finally, the SEC could be more lenient against traders who made negligible profits or traded very small amounts, as in the theoretical model of DeMarzo, Fishman, and Hagerty (1998). However, we would not expect very small trades to influence market aggregates in the first place.

5.2 Detection

Consider now the detection outcome, $\mathbb{I}_{\text{det}} \in \{0, 1\}$. Can one rule out that S covaries with $\mathbb{I}_{\text{det}}$? In principle, the SEC could screen trades based on the signals we find informative (Cases IV and V in Figure 3). This concern would be particularly strong if informed traders were exposed *only* when these signals display abnormal values. If that were the case, one could then overestimate the signals' capacity to detect the presence of informed traders. To the best of our knowledge, during our sample period,

the SEC did not have a formal analytical framework to monitor the value of information signals of the type we consider in this paper.[16] In fact, even if such a framework were in place, the results in Section 4 do not support this view. Many stock-based signals display patterns that are generally inconsistent with what economic reasoning would suggest are patterns of informed trading. For example, illiquidity measures have lower values on days with informed trades. One would then need to believe that the SEC is particularly sensitive to insider trading activity when markets look orderly and abnormally liquid. Furthermore, even if the regulating agency intended to rely on public information such as liquidity measures to flag an asset–date pair, it is unlikely that officials would be able to systematically identify which specific individuals are breaching the law, due to lack of granularity, trade aggregation and netting, the use of multiple accounts, and so forth.

Instead, prior evidence suggests that a significant fraction of investigations originate not from monitoring various information signals but, rather, from external tips. Meulbroek (1992) studies a sample of cases filed by the SEC in the 1980s and reports that *public complaints*—a category of investigations initiated for reasons unrelated to direct screening by regulators—are the most important source of investigations (41% of cases). Another source of tipping is from parties such as exchanges or brokers observing suspicious portfolio activity in their clients' accounts. For example, an individual could buy an unusually large position in a company for the first time just before a merger or any other important corporate announcement. This detection category is thus related to trades but relies on access to traders' identities, a source of information that is non-public and not necessarily reflected in public signals.

To further investigate potential detection-driven selection, we follow Meulbroek (1992), who argues that the detection of insider trading in investigations involving multiple companies is less likely to originate in market signals such as returns. Intuitively, for a generic case with, say, 10 firms, it is unlikely that detection in *each* firm was based on independent publicly observed signal movements. Rather, even when the investigation originated from screening one firm's signal values, it is likely that trades in the remaining firms were uncovered as part of a subsequent investigation (e.g., involving access to individual brokerage accounts, phone conversations, etc.). Therefore, if detection bias is at work, we should expect the correlation with information signals to be stronger for single-company cases. Based on this idea, we classify the SEC investigations according to the number of firms involved,

[16] This notion is supported by interviews we conducted with SEC officials.

n. The majority of investigations involve only one or two companies (n_{low}). Others involve a greater number of firms, n_{high}, with up to $n = 25$ in the sample. We then estimate the model in equation (2) with two additional regressors: n_{high} and an interaction term between n_{high} and $InfoTrade$. The results are reported in Table B4 of the Online Supplement. We find no statistically significant effects in the interaction terms for 15 out of 17 information signals. Moreover, for the two signals for whose interaction term is significant, the value is positive, opposite the selection hypothesis we conjecture.

To summarize, our analysis suggests that it is unlikely that signals based on public information, such as volatility or illiquidity, trigger investigations. Nonetheless, the conceptual possibility still remains that, say, a given broker's tip to the SEC is based on an abnormally high individual volume and that such a volume impacts the aggregate figures. To evaluate the impact of such a potential selection, we examine the origin of the investigation directly in the next section.

6 Evidence from the SEC WRP

In this section, we exploit a change in the regulatory environment related to insider trading investigations. As part of the Dodd–Frank Act of 2010 (15 USC par. 78u-6), the SEC instituted the WRP. The program rewards whistleblowers for providing *original information* directly to the SEC or related agencies, which is defined as information that is (i) derived from the independent knowledge or analysis of a whistleblower, (ii) not known to the SEC from any other sources, and (iii) not exclusively derived from an allegation made in a judicial or administrative hearing, governmental report, hearing, audit, or investigation or from the news media. This definition makes it clear that the detection of such cases is uncorrelated with any SEC/government action and, thus, such cases are free of detection selection concerns based on public signals values. Hence, if selection bias drives our results, we would expect signals to display different dynamics for cases originating from the WRP.

Given the nature of the shock, our analysis is confined to the period of 2011–2015. In this period, our sample includes 166 different cases, of which 37 were investigated through the program and 129 have no precise source of investigation. Table VII summarizes various trading characteristics for the two types of cases, which are fairly similar along most dimensions, including time from news arrival to trade, number of trades per firm, and trades per trader. The only notable difference is that the WRP cases involve, on average, companies with greater market capitalization.

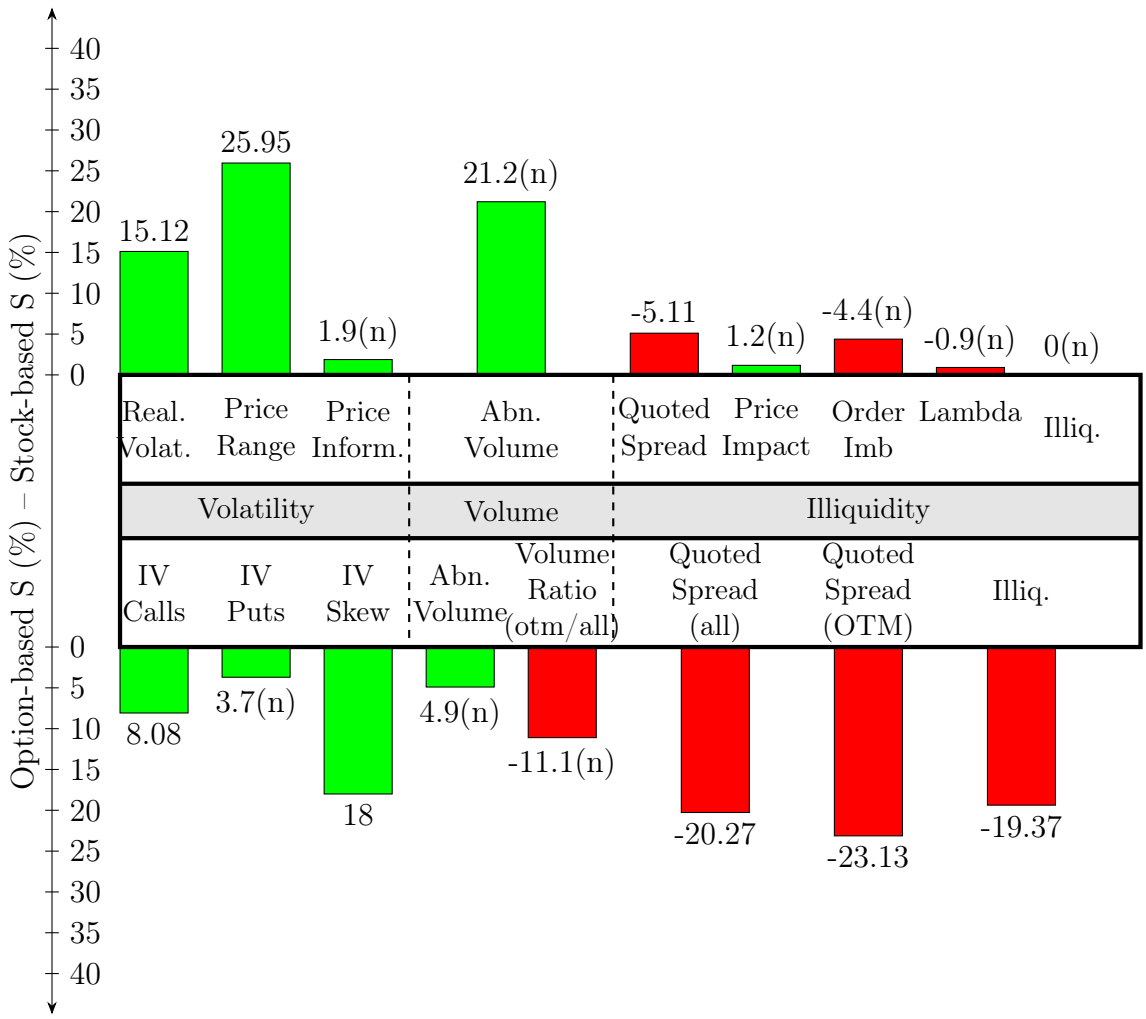


Figure 4. SEC WRP Cases Sample: Results Summary
For each information signal, the numbers displayed in the bars correspond to the percentage ratio between the estimated coefficient on days with precisely identified informed trading and the corresponding full-sample standard deviation. Green (red) columns correspond to positive (negative) effects on the signal value. Estimates that are not statistically significant are indicated by '(n)'.

Table VIII reports the results from estimating the regression model in equation (2) for the WRP sample and Figure 4 displays the value of the *InfoTrade* coefficient relative to the signal's corresponding standard deviation. The patterns for volatility, volume, and illiquidity remain essentially the same as for the full sample. All measures of volatility increase in both stock and option markets. The precision of the estimates is lower, although this is expected, given the significantly smaller sample. The effect on *IV* is stronger for call options, which further increases *IV skewness*. Regarding illiquidity, all measures in options markets, quoted spreads and *O. Illiq*, remain negative and statistically significant. *Quoted Spread* is also negative and statistically significant in stock market. *Lambda* and *Order Imbalance*

are negative but statistically insignificant. The illiquidity patterns are thus relatively stronger in option-based signals. *InfoTrade* has a positive coefficient of abnormal volume in both stock and option markets. The magnitude of the coefficient of stock abnormal volume is indeed higher in the WRP sample, albeit the t-statistics are smaller. For options, the coefficient is positive, but the magnitude is smaller.

Overall, the patterns of signal behavior in the WRP are consistent with our findings in the full sample in Section 4 and suggest that the detection methodology is not responsible for such patterns in the first place. Hence, we exploit the full sample to conduct cross-sectional tests on the content of information sets in Section 7 and to explore the role of specific trading strategies in Section 9.

7 The Cross-Section of Information Sets

In this section, we report the results from cross-sectional tests to shed light on how the specifics of the private information motivating the trades—information and corporate event types—affect the conditional behavior of information signals.

Information Type. Our sample contains instances of both positive and negative private information. From Section 3, we know that the use of positive and negative information displays a strong correlation with positive and negative returns, respectively, on the same days. In this section, we further examine whether the information sign generates distinct conditional signal responses. To this end, we estimate the regression model in (2) separately for each information sign. Table IX reports the results for stock-based signals (Panels A and B) and option-based signals (Panels C and D). Stock- and option-based volatility measures increase by significant amounts for both positive and negative news. We also observe an increase in abnormal volume, although the effect is only statistically significant for the larger subsample of positive news. Illiquidity measures also display negative coefficients in both subsamples. However, stock-based illiquidity measures display stronger negative responses when information has a negative sign, whereas, on the other hand, option-based illiquidity signals' responses are generally stronger for positive news. These differences could indicate that trading frictions play a role in the transmission of information across asset classes. We leave the details of such analysis for future research.

Corporate Event Types. We examine whether the behavior of information signals depends on the specific event category, that is, scheduled events (earnings announcements) and unscheduled events (M&As). This distinction is economically relevant given that scheduled announcements could trigger different reactions from investors than unscheduled ones would. For example, for scheduled events, investors could take directional bets on the event outcome, even when they do not act on actual information but on the sentiment or belief that they are informed. In contrast, making bets on unscheduled events is impossible without private information.

To this end, we estimate the regression model in (2) for both unscheduled and scheduled events. We note that the sample size of unscheduled events is more than twice as large as that of scheduled events for options and more than five times larger for stocks, which could affect the relative precision of each set of estimates. Table X reports the results for stock-based signals (Panels A and B) and option-based signals (Panels C and D). We observe that, for the stock-based signals, the effect on volatility is similar for both types of events and largely mirrors the unconditional results. We also observe an increase in abnormal volume in both subsamples, but the results are statistically significant only in the case of unscheduled events. The effect on stock- and option-based illiquidity signals, however, is stronger for unscheduled events: Both the magnitudes of the estimates and their corresponding t-statistics are lower in the case of scheduled events.

Overall, information signals display stronger abnormal behavior in the subsample of unscheduled events. Is the heterogeneity in behavior driven by the presence of noise traders, that is, uninformed traders acting as if they had information on a scheduled announcement? We further explore this possibility in Section 8 by analyzing a broader set of announcements.

8 Information Signals and Generic Corporate Events

One possible concern with the response of information signals to the presence of informed traders is that the observed abnormal patterns could reflect a systematic relation between information signals and corporate events. If the behavior of signals were the same for any corporate event, independent of whether it is investigated by the SEC or not, the usefulness of a given signal to help identify the presence of asymmetric information would be limited. To assess this possibility, we collect records for all earnings announcements from the Center for Research in Securities Prices (CRSP) and all M&A

announcements from SDC Platinum during the sample period of 1995–2015. We then compute the average daily value of each signal over the event window $[-20, 0]$, where $t = 0$ corresponds to the public announcement.

Figure 7 displays the results for earnings announcements and stock-based signals. We observe that *Realized Variance*, *Price Range*, *Abnormal Volume*, and *S. Illiq* display strong responses on the day of the *public* announcement. However, we find very little, if any, abnormal behavior in the signals *prior* to the announcement. Figure 8 shows the results for earnings announcements and option-based signals. In this case, we observe a traceable increase in the value of IV for both calls and puts over the event window and an increase in abnormal volume between one and two days before the announcement. Quoted spreads decline two days prior to the announcement and then increase during the announcement. Overall, the only signal that displays a predictable increase over the time window is IV. For all the others, the baseline results in Section 4 do not coincide with the predictable patterns just before earnings announcements.

We now turn to M&A announcements. Figure 9 displays the results for stock-based signals and Figure 10 those for option-based signals. With the exception of *Realized Variance* and *Price Range*, which increase the day before the announcement, stock-based signals display little abnormal behavior during the period considered. The same pattern is observed for option-based signals. Although most signals display strong responses on the announcement date, virtually no distinct behavior is observed prior to that date.

We conclude that, with the exception of implied volatility prior to earnings announcements, the documented abnormal behavior of information signals on informed trading dates cannot be simply attributed to proximity to a corporate event announcement.

9 How Do Traders' Strategies Impact Illiquidity Measures?

One of our key findings is the negative conditional response of illiquidity measures in both stock and option markets. This section explores this connection by analyzing the incidence of trading horizons, the use of limit orders, and firm size.

9.1 Trading Horizon Effects

One explanation for the negative relation between illiquidity and *InfoTrade* is based on the strategic behavior of informed investors. In particular, if trading costs are high due to temporary market illiquidity, an informed investor might want to time the execution of his or her trades to minimize such costs and, thus, one could observe low illiquidity levels when informed investors trade. This effect has been hypothesized by Collin-Dufresne and Fos (2015; 2016) for the case of of 13D filers. The question remains whether such a pattern results from the strategic behavior of traders or is driven by some other unobservable factor. We shed more light on this issue by taking advantage of a unique feature of our data, the fact that we can observe the date when traders *receive* private information about firm fundamentals. In particular, we compare the baseline results in Section 4 with those based on the subsample of cases with a short information horizon, that is, for which the number of days between receiving a private tip and the public announcement of the same information is no greater than three days. We argue that, within such a short horizon, it would be relatively difficult for the trader to time his or her trades well being constrained by information disclosure.

Table XI shows the estimated values of Δ for the illiquidity signals that display abnormal behavior in the full sample. Panels A to C display the results for all stocks, large caps, and small caps, respectively. We observe that the only signal that remains negative and statistically significant is *S. Illiq.* In the case of *Order Imbalance* and large stocks, the coefficient becomes positive. We also find that the coefficient of *InfoTrade* is positive for *Quoted Spread* for a sample of OTM options written on large caps, and for stock-based *Quoted Spread* and *Lambda* in the case of small caps, albeit the t-statistics are small. In most cases, the value of the Δ estimates on illiquidity measures is higher for short information horizons than for the full sample.

Overall, the test results support the notion that strategic timing plays a role in explaining the negative value of illiquidity signals. The fact that some signals remain negative even for relatively short horizons suggests that either traders still keep some ability to time trades within a three-day period, and/or that there are additional factors at play, such as the use of limit orders.

9.2 Use of Limit Orders

The effect on illiquidity measures can be affected by a second strategic dimension, the use of limit orders. If the use of limit orders by informed trader adds to market liquidity, that fact could help explaining why illiquidity measures decrease in value. But do informed traders use limit orders? To gain perspective on this issue, we screen all SEC litigation files for references to the use of specific order types. For the vast majority of trades, these files do not specify order types. Nonetheless, during the sample period, we identify 85 stock trades' order types, with 62 limit orders and 23 market orders.[17]

We exploit the subsample with identified order types to test the hypothesis that trades conducted with limit orders are associated with higher same-day liquidity. Table XII displays the estimated coefficient of *InfoTrade* from regression (2) for all trades (Panel A) and for informed trades that use market and limit orders (Panels B and C, respectively). We do not find any significant difference between the order types in the case of the *Quoted Spread* and *Lambda*. On the other hand, the values of Δ for *Order Imbalance* and *S. Illiq.* are significantly lower when limit orders are used, consistent with the aforementioned hypothesis.

We note that using trade order types can provide perspective on their correlation with illiquidity, but establishing causality is more difficult. Unlike the test in Section 9.1, in which we exploit ex ante heterogeneity in information horizons, the informed trader could demand more immediacy, by using market orders, precisely when the trader observes that market illiquidity is low. To gain additional perspective on the direction of the causal link, we resort next to an ex ante source of illiquidity, namely, the market capitalization of the stock about which the informed trader receives a tip.

9.3 Illiquidity and Market Capitalization

Similar in spirit to previous tests, we estimate the empirical model in equation (2) for the subsample of firms with market capitalization below and above the median value in the sample. The results, presented in Panels D and E of Table XII, reveal interesting facts. First, we do find that the behavior of illiquidity measures is strongly related to equity size. The negative relation with *InfoTrade* is particularly strong for the subset of firms with capitalization below the market median. The negative and statistically significant coefficients of *Quoted Spread* and *Order Imbalance* for small stocks suggests

[17] The number of option trades for which the order type was identified is significantly smaller.

that informed traders could use relatively more limit orders when trading these stocks. Both of these coefficients are smaller in absolute value for large-cap stocks.

Overall, the results in this section are consistent with the view that informed traders are more likely to strategically time trades, especially for assets suffering from relatively high liquidity costs (i.e., small caps). Informed traders are also more likely to trade such assets using limit orders. Although it is difficult to empirically disentangle the specific contribution of each channel, it is reasonable to expect that a rational informed trader would consider both.

10 Discussion of Implications

A large literature examines the predictions of asymmetric information theories based on the statistical power of publicly available information signals.[18] Our finding that many of these signals display abnormal conditional behavior on asset–day pairs with precisely identified informed trading provides direct support for their use in empirical studies. At the same time, our research sheds new light on *how* these signals perform and offers new insights for future investigations. In this section, we provide a brief discussion of implications for several strands of related literature.

10.1 Implications for Empirical Analyses of Asymmetric Information

Option Signals. The results suggest strong information content in both stock and option markets. Given that option-based signals are not used as frequently in tests of asymmetric information theories, one can argue that they should be further emphasized (e.g., Johnson and So (2017)). Our findings are consistent with those of Chakravarty et al. (2004). Also consistent is the finding of Chan, Chung, and Fong (2002) that information in option markets manifests in quote revisions.

Volume and Signed Order Flow. We find that the signal contained in abnormal volumes is positively correlated with the presence of informed trading, even for those investigations that originate in the WRP. Given that much of the empirical research to date has relied on bid–ask spread constructs and/or order flow imbalances as signals of information, this result calls for more emphasis on

[18] Notable examples from the asset pricing and corporate finance literature include Easley, Hvidkjaer, and O'Hara (2002), Chae (2005), Chen, Goldstein, and Jiang (2007), Chen et al. (2007), Ferreira and Laux (2007), Roll et al. (2009), among many others. Examples can also be found in the accounting literature, e.g., Frankel and Li (2004).

volume. This fact seems to be more relevant in the current market environment, given the disruption of high-frequency trading (e.g., Easley, de Prado, and O'Hara (2016)). High-frequency trading and algorithmic trading introduce large amounts of noise in quote activity due to large cancellation ratios, spoofing, and so forth, but they do not necessarily disrupt volume in equal measure. Moreover, the use of sophisticated algorithms that routinely use limit orders across markets implies that the trade classification rule of the aggressor side may bear little correlation with information flows. Consistent with this view is our finding that *Order Imbalance* does not increase in the presence of informed traders. Structural empirical models that exploit volume, such as that of Back, Crotty, and Li (2018) and the volume-based imbalance measure of Easley, de Prado, and O'Hara (2016), are promising steps in this direction.

Illiquidity, Timing Options, and Market Making. Abnormally low illiquidity values when informed traders trade are consistent with models of optimal liquidity timing (e.g., Admati and Pfleiderer (1988), Collin-Dufresne and Fos (2016)), but they challenge the common use of illiquidity metrics in empirical studies, that is, linking higher levels of illiquidity with greater adverse selection risk. As noted in Section 2, this fact could affect the economic interpretation of test results that rely on such approach.

Low conditional values of illiquidity measures could also be the equilibrium outcome of traditional market-making models. To illustrate this fact, consider the model of Glosten and Milgrom (1985), in which bid–ask spreads are the difference between two conditional expectations, and assume that asymmetric information, if present, is revealed at the end of each day. On a given day, the evolution of the order flow depends on the unobservable presence of informed traders $I \in \{0, 1\}$. On days with $I = 1$, as the informed trader causes order imbalances, the market makers experience progressive resolution of uncertainty through Bayesian learning. Therefore, depending on parameter values, the daily average of the bid–ask spread may be actually lower on days with $I = 1$.[19]

Overall, we conclude that adverse-selection risk inference based on illiquidity measures require

[19] More formally, let Y_u denote the aggregate order flow at time u and T denote the time period between the opening and closing times. Depending on the specifics of the environment, one could indeed observe that

$$\frac{1}{T}\int_{\text{open}}^{\text{close}} \mathbb{E}_u(\text{bid-ask spread}_u | I = 1) dY_u < \frac{1}{T}\int_{\text{open}}^{\text{close}} \mathbb{E}_u(\text{bid-ask spread}_u | I = 0) dY_u,$$

where $\mathbb{E}_u$ denotes the conditional expectation relative to the available public information available at time u.

a structural approach that explicits timing options, as well as the learning process of uninformed participants.

Volatility. The footprints of information are reflected not only in the conditional responses of volume and illiquidity but also in those of volatility measures. Thus, our results advocate a broader use of the volatility-based signals in empirical studies. Their use in the literature to help identify the presence of informed traders, on the other hand, seems infrequent. Related to our findings are the results of An, Ang, Bali, and Cakici (2014), who suggest that patterns of IV could indicate informed trading. The authors find that an increase in call (put) IV predicts higher (lower) returns the following month.

Scheduled versus Unscheduled Events. We find that the ability of certain signals to reveal the presence of informed traders is mitigated in the case of scheduled events, especially IV in option markets. Therefore, time-series analyses of adverse selection risk around earnings announcements require addressing the fact that the conditional behavior of these signals resembles the unconditional behavior for the average event. One could, for example, develop an empirical strategy that exploits the insights of the model of Banerjee and Green (2015) where both informed and difference-in-opinion trade motives are present.

On the other hand, we do observe an increase in *Option Abnormal Volume* prior to those M&As announcements where informed traders were present, a fact that is consistent with the information-driven excess volume conjectured by Cao, Chen, and Griffin (2005) and Augustin, Brenner, and Subrahmanyam (2015) and the posterior PIN estimation before mergers of Brennan et al. (2018). Also related is the study by Chae (2005), who studies patterns of trading volume before scheduled and unscheduled events and finds that the trading volume before scheduled announcements is negatively correlated with levels of ex ante information asymmetry.

Combining Signals. One promising area for future work is the creation of reliable indexes of adverse selection risk, that is, methods to optimally aggregate the power from multiple signals. Bharath, Pasquariello, and Wu (2009) develop an index of asymmetric information by combining seven signals using principal component analysis.[20] This approach is plausible because it does not rely on a single signal to draw economic conclusions. However, the choice of signals in such exercises—how many, which

[20] Also related is the composite liquidity approach of Korajczyk and Sadka (2008).

ones, with what weights?—is likely to be ad hoc unless one has a source of external validation. Our study offers direct external evidence on the potential usefulness of volatility, volume, and illiquidity metrics in both equity and option markets when traders exploit information about firm fundamentals. In this regard, we hope that future research can benefit from our results to construct enhanced indexes, especially at frequencies that are closer to real-life decision making (e.g., weekly, daily, or intra-daily).

10.2 Implications for the Legal Enforcement of Insider Trading Laws

A regulatory agency may be interested in investigating insider trading. Such investigations yield an outcome $G_j \in \{0, 1\}$ that represents whether trader j is guilty of violating insider trading laws. One can consider the regulator as trying to estimate $\mathbb{P}(I = 1, G = 1)$ for a given asset over a given period. Investigations are costly, so the agency may target a certain estimation accuracy and try to minimize the expected cost. For example, in the theoretical model of DeMarzo, Fishman, and Hagerty (1998), regulators consider only trade volume as an information signal and trigger an investigation if volume $> \overline{v}$, for a given threshold $\overline{v}$. Generally, a regulator could consider a vector of signals S, and design a rule that triggers an investigation if $\mathbb{P}(I = 1, G = 1|\mathrm{S}) > \overline{P}$. If an investigation occurs, the regulator learns whether $I \times G = 1$ or $I \times G = 0$. Armed with a sample that includes such false positives, a researcher could evaluate what rule is optimal. We hope that our results provide guidance in future endeavors of this spirit.

10.3 Implications for Theories of Informed Trading

Multimarket Approach. We document that informed traders regularly use both stocks and options. This fact highlights the importance of modeling the information transmission process from a multi-asset perspective. For example, recent work by Back and Crotty (2015) addresses the interaction between the stock market and the corporate bond market and Johnson and So (2017) address the interaction between the stock market and option markets.

Information Structure. The structure of the PIN model has been enriched and extended by Easley, Engle, O'Hara, and Wu (2008), Odders-White and Ready (2008), and Duarte and Young (2009), among others. Most of these authors assume that informed traders do not respond to price changes. In contrast, Back, Crotty, and Li (2018) analyze a dynamic model with a PIN-like information structure

but in which a single informed trader acts strategically, as in Back (1992), and conclude that information asymmetries cannot be identified using order flow alone. Our results support the notion that structural models should relate both prices and volume to measures of adverse selection risk. On the other hand, they indicate that less reliance on signed order flow in their estimation is desirable.

In addition, an exciting avenue for future research is the study of more realistic information structures to better understand the interaction between informed investors, market makers, and other market participants. For example, the results of Wang and Yang (2017) show that inference based on the Kyle-type model of Back and Baruch (2004) is sensitive to the introduction of the possibility that the informed trader is not present in the first place. The results of Banerjee and Breon-Drish (2017) suggest that inference is also sensitive to the introduction of an information acquisition decision.

References

Admati, A. R. and P. Pfleiderer (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3–40.

Amihud, Y. (2002): "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5, 31–56.

An, B. J., A. Ang, T. G. Bali, and N. Cakici (2014): "The joint cross section of stocks and options," *Journal of Finance*, 69, 2279–2337.

Augustin, P., M. Brenner, and M. Subrahmanyam (2015): "Informed Options Trading Prior to MA Announcements: Insider Trading?" *Working Paper*.

Back, K. (1992): "Insider Trading in Continuous Time," *The Review of Financial Studies*, 5, 387–409.

——— (1993): "Asymmetric Information and Options," *The Review of Financial Studies*, 6, 435–472.

Back, K. and S. Baruch (2004): "Information in Securities Markets: Kyle Meets Glosten and Milgrom," *Econometrica*, 72, 433–465.

Back, K. and K. Crotty (2015): "The informational role of stock and bond volume," *Review of Financial Studies*, 28, 1381–1427.

Back, K., K. Crotty, and T. Li (2018): "Identifying Information Asymmetry in Securities Markets," *The Review of Financial Studies*, forthcomin.

Banerjee, S. and B. Breon-Drish (2017): "Dynamic Information Acquisition and Strategic Trading," *Working Paper*.

Banerjee, S. and B. Green (2015): "Signal or noise? Uncertainty and Learning about Whether Other Traders are Informed," *Journal of Financial Economics*, 117, 398–423.

Baruch, S., M. Panayides, and K. Venkataraman (2017): "Informed trading and price discovery before corporate events," *Journal of Financial Economics*, 125, 561–588.

Bharath, S. T., P. Pasquariello, and G. Wu (2009): "Does asymmetric information drive capital structure decisions," *Review of Financial Studies*, 22, 3211–3243.

Biais, B., L. Glosten, and C. Spatt (2005): "Market Microstructure: A survey of Microfoundations, Empirical Results, and Policy Implications," *Journal of Financial Markets*, 8, 217–264.

Biais, B. and P. Hillion (1994): "Insider and Liquidity Trading in Stock and Options Markets," *Review of Financial Studies*, 7, 743–780.

Black, F. (1975): "Fact and Fantasy in the Use of Options," *Financial Analysts Journal*, 31, 36–41.

Boulatov, A., T. Hendershott, and D. Livdan (2013): "Informed Trading and Portfolio Returns," *Review of Economic Studies*, 80, 35–72.

Brav, A., W. Jiang, and H. Kim (2015): "The Real Effects of Hedge Fund Activism: Productivity, Asset Allocation, and Labor Outcomes," *Review of Financial Studies*, 28, 2723–2769.

Brennan, M. J., S.-W. Huh, and A. Subrahmanyam (2018): "High-Frequency Measures of Informed Trading and Corporate Announcements," *Review of Financial Studies*, forthcomin.

Caldentey, R. and E. Stacchetti (2010): "Insider Trading With a Random Deadline," *Econometrica*, 78, 245–283.

Cao, C., Z. Chen, and J. M. Griffin (2005): "Informational Content of Option Volume Prior to Takeovers," *The Journal of Business*, 78, 1073–1109.

Chae, J. (2005): "Trading Volume, Information Asymmetry, and Timing Information," *The Journal of Finance*, 60, 413–442.

Chakravarty, S., H. Gulen, and S. Mayhew (2004): "Informed Trading in Stock and Option Markets," *Journal of Finance*, 59, 1235–1257.

Chan, K., Y. P. Chung, and W.-M. Fong (2002): "The Informational Role of Stock and Option Volume," *Review of Financial Studies*, 15, 1049–1075.

Chen, Q., I. Goldstein, and W. Jiang (2007): "Price Informativeness and Investment Sensitivity to Stock Price," *Review of Financial Studies*, 20, 619–650.

Cohen, L., C. Malloy, and L. Pomorski (2012): "Decoding Inside Information," *Journal of Finance*, 67, 1009–1043.

Collin-Dufresne, P. and V. Fos (2015): "Do Prices Reveal the Presence of Informed Trading?" *The Journal of Finance*, 70, 1555–1582.

——— (2016): "Insider Trading, Stochastic Liquidity and Equilibrium Prices," *Econometrica*, 84, 1441–1475.

Collin-Dufresne, P., V. Fos, and D. Muravyev (2017): "Informed Trading in the Stock Market and Option Price Discovery," *Working Paper*.

Cremers, M. and D. Weinbaum (2010): "Deviations from Put–Call Parity and Stock Return Predictability," *Journal of Financial and Quantitative Analysis*, 45, 335–367.

DeMarzo, P. M., M. J. Fishman, and K. M. Hagerty (1998): "The optimal enforcement of insider trading regulations," *Journal of Political Economy*, 106, 602–632.

Duarte, J. and L. Young (2009): "Why Is PIN priced?" *Journal of Financial Economics*, 91, 119–138.

Durnev, A., R. Morck, and B. Yeung (2004): "Value-Enhancing Capital Budgeting and Firm-Specific Stock Return Variation," *Journal of Finance*, 59, 65–105.

Easley, D., M. L. de Prado, and M. O'Hara (2016): "Discerning Information from Trade Data," *Journal of Financial Economics*, 120, 269–285.

Easley, D., R. F. Engle, M. O'Hara, and L. Wu (2008): "Time-Varying Arrival Rates of Informed and Uninformed trades," *Journal of Financial Econometrics*, 6, 171–207.

Easley, D. and O. Hara (1987): "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, 19, 69–90.

Easley, D., S. Hvidkjaer, and M. O'Hara (2002): "Is Information Risk a Determinant of Asset Returns?" *The Journal of Finance*, 57, 2185–2221.

Easley, D. and M. O'Hara (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.

Easley, D., M. O'Hara, and P. S. Srinivas (1998): "Option volume and stock prices: Evidence on where informed traders trade," *Journal of Finance*, 53, 431–465.

Ferreira, M. A. and P. A. Laux (2007): "Corporate governance, idiosyncratic risk, and information flow," *The Journal of Finance*, 62, 951–989.

Frankel, R. and X. Li (2004): "Characteristics of a firm's information environment and the information asymmetry between insiders and outsiders," *Journal of Accounting and Economics*, 37, 229–259.

Glosten, L. R. and P. R. Milgrom (1985): "Bid, Ask and Transaction prices in a Specialist Market with Heterogeneously Informed traders," *Journal of Financial Economics*, 14, 71–100.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009): "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics*, 92, 153–181.

Grossman, S. (1976): "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," *The Journal of Finance*, 31, 573–585.

Hasbrouck, J. (2009): "Trading costs and returns for U.S. Equities: Estimating Effective Costs from Daily Data," *Journal of Finance*, 64, 1445–1477.

Hendershott, T., D. Livdan, and N. Schurhoff (2015): "Are Institutions Informed about News?" *Journal of Financial Economics*, 117, 249–287.

Holden, C. W. and S. E. Jacobsen (2014): "Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions," *Journal of Finance*, 69, 1747–1785.

Johnson, T. L. and E. C. So (2017): "A Simple Multimarket Measure of Information Asymmetry," *Management Science*, 1909, 1–26.

Korajczyk, R. A. and R. Sadka (2008): "Pricing the commonality across alternative measures of liquidity," *Journal of Financial Economics*, 87, 45–72.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

Meulbroek, L. K. (1992): "An Empirical Analysis of Illegal Insider Trading," *Journal of Finance*, 47, 1661–1699.

Odders-White, E. R. and M. J. Ready (2008): "The Probability and Magnitude of Information Events," *Journal of Financial Economics*, 87, 227–248.

Roll, R., E. Schwartz, and A. Subrahmanyam (2009): "Options trading activity and firm valuation," *Journal of Financial Economics*, 94, 345–360.

Wang, J. (1993): "A Model of Intertemporal Asset Prices Under Asymmetric Information," *The Review of Economic Studies*, 60, 249–282.

Wang, Y. and M. Yang (2017): "Insider trading when there may not be an insider," *Working Paper*, 1–26.

Table II
Trade Characteristics: Descriptive Statistics

The unit of observation is the insider trade. **Panel A** classifies trades by trading instrument. **Panel B** classifies trades by the direction of trading. **Panel C** reports various trading statistics.

Panel A: Distribution of Trading Instruments	**Number of trades**	**Percentage of trades**
Stocks	3,392	67.06
Options	1,610	31.83
ADS	44	0.87
Bonds	12	0.33
Total	5,058	100
Panel B: Distribution of Buys and Sells		
Buys	4,220	83.43
Sales	838	16.57

Panel C: Trading Statistics

Characteristic	mean	median	st. dev.	min	max
Number of days from receiving a tip to an informed trade	8.05	2	23.88	0	417
Number of days from a trade to information disclosure	24.77	7	61.59	0	998
Number of days from the first to the last informed trade	19.23	8	73.34	1	738
Firms per case	4.72	2	5.32	1	25
Traders per case	5.06	3	4.55	1	18
Trades per firm	31.47	16	45.17	1	231
Trades per trader	20.26	10	24.05	1	97
Reported profit ($1,000s)	1013.6	90.00	7926.8	4.0	27500

Table III
Informed Volume and the Information Content of Trades

Panel A shows the share of informed trading volume as a percentage of total volume on days when $InfoTrade = 1$, as defined in Section 4. **Panel B** shows the average stock returns on days when $InfoTrade = 1$ for positive and negative news. Market is a value-wighted portfolio of all stocks in CRSP. **Panel C** shows stock returns (excluding dividends) computed from the opening price on the first insider trading day to the opening price on the day following the information disclosure date. The returns are split according to positive and negative news. The aggregate return considers the absolute value of each return. The returns for SEC 13D filers are measured from the opening price of the day 13D filers trade to the opening price of the day following the public disclosure of the trade. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Panel A: Volume Share of Informed Traders

Security	Stocks	Calls	Puts
Mean (%)	10.2	38.1	31.5
Median (%)	2.8	23.3	13.9
Standard deviation (%)	22.2	39.9	38.5

Panel B: Returns on Informed Trading Days

Information/ Adj. Portfolio	Positive	Negative	Positive Market	Negative Market	Positive S&P500	Negative S&P500
Return (%)	0.814***	-0.584*	0.815***	-0.683*	0.825***	-0.681**
	(0.163)	(0.304)	(0.165)	(0.299)	(0.166)	(0.299)
#Obs	2,397	506	2,397	506	2,397	506

Panel C: Private Information Signal Strength

	Positive	Negative	Aggregate	
	SEC Insider Trading Cases			SEC 13D Filers
Mean Return (%)	43.510***	-18.564***	38.271***	4.927***
	(4.199)	(2.142)	(3.389)	(0.638)
Median Return (%)	33.690***	-15.322***	29.427***	2.401***
	(2.348)	(2.545)	(2.275)	(0.173)
#Obs	2,351	696	3,055	2,628

Table IV
Stock-based Signals: Estimation Results

The dependent variables are daily stock-based signals at the firm level over the period 1995-2015. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume	Illiquidity				
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
InfoTrade	0.723***	1.199***	826.182**	413.083***	-0.097***	-0.095	-0.016***	-0.031***	-0.423***
	(0.120)	(0.134)	(399.446)	(107.995)	(0.016)	(0.482)	(0.005)	(0.007)	(0.072)
LNSIZE	-1.036**	-1.814***	-629.587	1,135.726	-0.312***	-5.460***	-0.054***	-0.066***	0.089
	(0.443)	(0.434)	(625.417)	(781.669)	(0.043)	(1.202)	(0.012)	(0.017)	(0.061)
LNVOL	0.794***	0.997***	-421.947	-290.165	-0.231***	-0.637	-0.018	-0.053***	-0.341***
	(0.298)	(0.281)	(710.046)	(329.994)	(0.048)	(1.494)	(0.011)	(0.012)	(0.092)
TURNOVER	-6.742	13.689	-2,759.961	35,178.276	28.155***	198.180**	1.544**	5.219***	25.092***
	(18.167)	(23.791)	(49,493.401)	(27,462.341)	(3.354)	(91.654)	(0.752)	(0.787)	(5.874)
PRC	0.121***	0.113***	4.393	-82.744*	-0.007**	0.017	-0.002*	0.001	-0.002
	(0.042)	(0.028)	(46.656)	(42.437)	(0.003)	(0.069)	(0.001)	(0.001)	(0.004)
Constant	0.470*	0.315*	-464.547	101.828	-0.293***	-2.068***	-0.045***	-0.037***	0.077*
	(0.271)	(0.174)	(389.087)	(160.028)	(0.025)	(0.616)	(0.006)	(0.005)	(0.047)
#Obs	10,157	10,196	10,087	10,156	10,079	10,029	10,029	10,014	10,121

Table V
Option-based Signals: Estimation Results

The dependent variables are daily option-based signals at the firm level over the period 1995-2015. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (OTM)	O. Illiq.
InfoTrade	0.018***	0.017**	0.001	2,514.978***	-5.578***	-0.075***	-0.081***	-0.153***
	(0.006)	(0.007)	(0.003)	(561.111)	(2.047)	(0.015)	(0.021)	(0.029)
LNSIZE	-0.028	-0.060	0.304	10,303.422**	-29.546	0.840***	1.026***	0.193
	(0.091)	(0.114)	(0.244)	(5,200.853)	(47.109)	(0.219)	(0.250)	(0.281)
LNVOL	-0.145***	-0.119***	0.439	22,235.474***	10.112	0.036	0.068	0.069
	(0.048)	(0.037)	(0.429)	(7,866.591)	(16.670)	(0.063)	(0.082)	(0.059)
TURNOVER	11.551**	5.723	-42.666	-2001624.461***	-579.399	-39.375***	-55.142***	-7.615
	(5.411)	(4.306)	(40.118)	(742,170.275)	(1,962.582)	(7.476)	(9.372)	(7.072)
PRC	0.003*	0.004	-0.007	-215.008*	0.443	-0.019***	-0.024***	-0.005
	(0.002)	(0.002)	(0.005)	(112.945)	(0.927)	(0.004)	(0.005)	(0.006)
Constant	0.061	0.071	-0.369	-16,241.190***	13.553	-0.471***	-0.601***	-0.088
	(0.040)	(0.057)	(0.334)	(5,926.633)	(23.009)	(0.107)	(0.121)	(0.151)
#Obs	2,721	2,709	2,442	2,728	2,728	2,673	2,673	2,590

Table VI
Conditioning on Trade Intensity

The dependent variables are information signals over the period 1995-2015. This table presents results for low- and high-intensity trades, as defined in Section 4. **Panels A and B** report the results for stock-based signals and **Panels C and D** for option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume		Illiquidity			
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
Panel A: Stock-based Signals: Low Intensity									
InfoTrade	0.685***	0.895***	1,068.047	668.337*	-0.044	-0.388	0.011	-0.027*	-0.080
	(0.202)	(0.224)	(701.195)	(389.987)	(0.030)	(0.630)	(0.009)	(0.015)	(0.059)
(0.003)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	4,123	4,123	4,122	4,120	4,123	4,123	4,123	4,123	4,123
Panel B: Stock-based Signals: High Intensity									
InfoTrade	0.711**	1.514***	782.190	155.128	-0.145***	-0.271	-0.041***	-0.042**	-0.850***
	(0.290)	(0.360)	(758.012)	(132.801)	(0.053)	(0.942)	(0.012)	(0.017)	(0.227)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	4,687	4,704	4,633	4,667	4,679	4,641	4,641	4,627	4,653

	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (OTM)	O. Illiq.
Panel C: Option-based Signals: Low Intensity								
InfoTrade	0.020	0.033	-0.009	6,367.787*	-1.261	-0.033	-0.072	-0.031*
	(0.023)	(0.029)	(0.010)	(3,268.428)	(2.830)	(0.047)	(0.060)	(0.017)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	852	848	819	857	857	856	856	825
Panel D: Option-based Signals: High Intensity								
InfoTrade	0.016*	0.008	0.003	2,421.320**	-8.100**	-0.080**	-0.067	-0.207**
	(0.008)	(0.009)	(0.004)	(952.145)	(3.976)	(0.039)	(0.048)	(0.080)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,378	1,372	1,169	1,378	1,378	1,344	1,344	1,319

Table VII
SEC Whistleblower Reward Program Cases: Summary Statistics

Characteristic/Sample	WRP Cases	Non-WRP Cases
Number of cases	55	373
Number of days from receiving a tip to an informed trade	17.16	20.97
Number of days from a trade to information disclosure	13.95	26.13
Number of days from the first to the last informed trade	21.72	19.04
Trades per firm	26.46	23.58
Trades per trader	21.05	19.8
Reported profits ($ millions)	1.58	0.8

Table VIII
SEC Whistleblower Reward Program Cases: Estimation Results

This table presents results for the subsample of SEC WRP cases. The dependent variables are information signals. **Panel A** reports results for stock-based signals and **Panel B** the results for option-based signals. The dependent variables are daily stock-based signals at the firm level over the period 1995-2015. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Panel A: Stock-based Signals

Based on	Volatility			Volume	Illiquidity				
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
InfoTrade	0.446**	1.000***	162.565	640.212	-0.048**	0.203	-0.007	-0.003	0.000
	(0.192)	(0.267)	(629.517)	(478.682)	(0.021)	(1.075)	(0.007)	(0.015)	(0.032)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,527	1,527	1,522	1,526	1,507	1,507	1,507	1,507	1,603

Panel B: Option-based Signals

Based on	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (OTM)	O. Illiq.
InfoTrade	0.021**	0.010	0.009*	472.835	-4.776	-0.075**	-0.111**	-0.122*
	(0.010)	(0.014)	(0.005)	(1,256.698)	(4.206)	(0.034)	(0.047)	(0.065)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	432	430	414	432	432	415	415	420

Table IX
Conditioning on Information Direction

This table presents separate results for positive and negative news. The dependent variables are information signals. **Panels A and B** report the results for stock-based signals and **Panels C and D** the results for option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume		Illiquidity			
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
Panel A: Stock-based Signals: Positive News									
InfoTrade	0.673***	1.018***	421.313	362.335**	-0.076**	0.260	-0.011	-0.016	-0.473***
	(0.194)	(0.219)	(432.467)	(182.982)	(0.030)	(0.616)	(0.008)	(0.010)	(0.123)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	7,569	7,596	7,514	7,557	7,512	7,463	7,463	7,451	7,529
Panel B: Stock-based Signals: Negative News									
InfoTrade	0.698**	1.553***	2,166.924	521.596	-0.146**	-1.337	-0.025*	-0.096***	-0.850***
	(0.296)	(0.491)	(1,415.416)	(469.736)	(0.073)	(1.166)	(0.013)	(0.031)	(0.227)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,751	2,763	2,736	2,762	2,730	2,729	2,729	2,726	4,653

	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (OTM)	O. Illiq.
Panel C: Option-based Signals: Positive News								
InfoTrade	0.022*	0.017	0.007	3,355.232**	-4.171	-0.069**	-0.082**	-0.135***
	(0.012)	(0.013)	(0.008)	(1,363.562)	(2.905)	(0.031)	(0.040)	(0.048)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,076	2,064	1,881	2,079	2,079	2,040	2,040	1,985
Panel D: Option-based Signals: Negative News								
InfoTrade	0.018*	0.016	-0.005	1,889.943	-4.331	-0.077	-0.074	-0.091
	(0.010)	(0.011)	(0.005)	(1,157.605)	(3.472)	(0.068)	(0.084)	(0.070)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	524	524	468	526	526	510	510	518

Table X
Conditioning on Unscheduled and Scheduled Corporate Events

This table presents separate results for unscheduled and scheduled corporate events. The dependent variables are information signals. **Panels A and B** report the results for stock-based signals. **Panels C and D** present the results for option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume	Illiquidity				
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
Panel A: Stock-based Signals: Unscheduled Events									
InfoTrade	0.727***	1.251***	753.514	432.798**	-0.107***	-0.102	-0.019**	-0.032***	-0.504***
	(0.186)	(0.222)	(499.504)	(173.602)	(0.033)	(0.621)	(0.008)	(0.012)	(0.134)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	8,306	8,329	8,241	8,289	8,229	8,179	8,179	8,164	8,255
Panel B: Stock-based Signals: Scheduled Events									
InfoTrade	0.719*	1.024*	1,480.837	374.543	-0.002	0.264	0.011	-0.033	-0.026
	(0.410)	(0.588)	(1,494.538)	(724.786)	(0.058)	(1.572)	(0.012)	(0.035)	(0.035)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,404	1,419	1,403	1,419	1,404	1,404	1,404	1,404	1,419

	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (OTM)	O. Illiq.
Panel C: Option-based Signals: Unscheduled Events								
InfoTrade	0.018	0.017	0.002	2,276.295***	-7.299**	-0.078**	-0.078*	-0.140***
	(0.012)	(0.013)	(0.006)	(694.660)	(3.258)	(0.033)	(0.039)	(0.052)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,858	1,847	1,654	1,864	1,864	1,863	1,863	1,747
Panel D: Option-based Signals: Scheduled Events								
InfoTrade	0.024***	0.021	-0.002	3,513.298	2.034	-0.068	-0.083	-0.206*
	(0.007)	(0.015)	(0.004)	(4,207.452)	(4.977)	(0.045)	(0.060)	(0.108)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	789	788	714	790	790	736	736	771

Table XI
Illiquidity Signals: Tests on Strategic Timing

This table presents separate results for long and short information horizons. A long (short) horizon is defined as one containing more than (less than or exactly) three trading days between the time in which information is received by the trader and the date of its public disclosure. The dependent variables are information signals. **Panel A** reports the results for short horizons and all assets and **Panels B and C** the results for short horizons and large and small caps, respectively. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Market	**Stock**				**Option**	
	Quoted Spread	Order Imbalance	Lambda	S. Illiq.	Quoted Spread otm	O. Illiq.
Panel A: Short Information Horizon: All Market Caps						
InfoTrade	-0.041	-0.029	-0.012	-0.736**	0.011	-0.018
	(0.147)	(0.041)	(0.049)	(0.348)	(0.021)	(0.035)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	671	668	665	669	123	117
Panel B: Short Information Horizon: Large Caps						
InfoTrade	-0.112	0.051**	-0.042	-0.158*	0.011	-0.018
	(0.131)	(0.019)	(0.035)	(0.081)	(0.020)	(0.035)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	290	290	290	290	118	112
Panel C: Short Information Horizon: Small Caps						
InfoTrade	0.008	-0.085	0.009	-1.138**	n/a	n/a
	(0.230)	(0.061)	(0.078)	(0.518)		
Controls	Yes	Yes	Yes	Yes		
#Obs	381	378	375	379		

Table XII
Stock Illiquidity Signals: Tests on Order Types and Market Capitalization

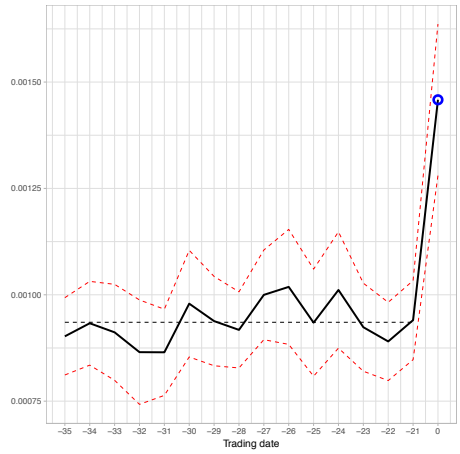
The dependent variables are information signals. **Panel A** presents the results for the sample of all trades. We consider subsamples for informed trades executed with limit or market orders (**Panels B and C**) and those based on different market capitalizations (**Panels D and E**). $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

	Quoted Spread	Order Imbalance	Lambda	S. Illiq.
Panel A: All Trades				
InfoTrade	-0.097***	-0.016***	-0.031***	-.423***
	(0.016)	(0.005)	(0.007)	(0.0072)
Controls	Yes	Yes	Yes	Yes
#Obs	10,079	10,029	10,014	10,121
Panel B: Only Market Orders				
InfoTrade	-0.210**	-0.053	-0.016	0.063
	(0.088)	(0.035)	(0.015)	(0.169)
Controls	Yes	Yes	Yes	Yes
#Obs	117	117	117	119
Panel C: Only Limit Orders				
InfoTrade	-0.181	-0.099**	0.045	-2.574***
	(0.181)	(0.046)	(0.043)	(0.741)
Controls	Yes	Yes	Yes	Yes
#Obs	238	240	238	252
Panel D: All Order Types: Small Caps				
InfoTrade	-0.163***	-0.027**	-0.053***	-0.755***
	(0.047)	(0.012)	(0.018)	(0.195)
Controls	Yes	Yes	Yes	Yes
#Obs	6,022	5,972	5,956	6,048
Panel E: All Order Types: Large Caps				
InfoTrade	-0.016	0.001	-0.004	-0.850***
	(0.024)	(0.005)	(0.008)	(0.227)
Controls	Yes	Yes	Yes	Yes
#Obs	4,236	4,236	4,237	4,653

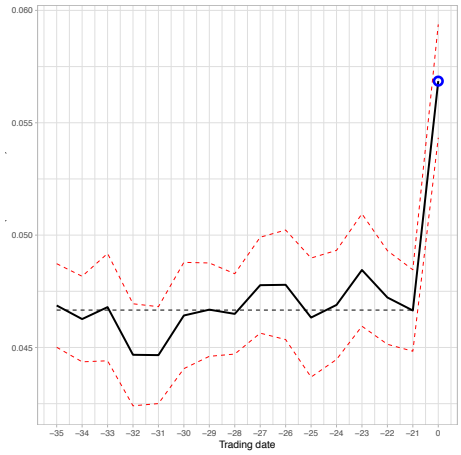
Figure 5. Stock-based Signals: Daily Mean Values Around Informed Trading

The figure presents the average values (aggregated across all trades) of stock-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days when $InfoTrade = 1$. The black dotted line is the mean value in the control window period of trading dates [-35,-21].

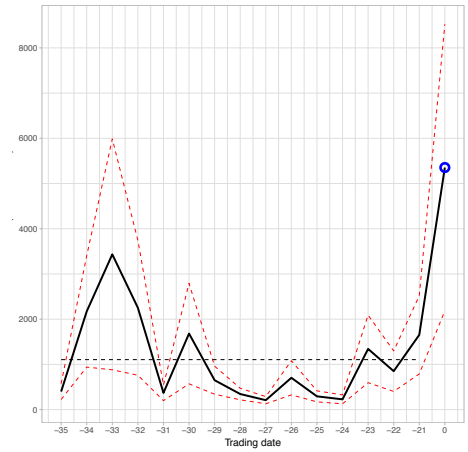
(a) Realized Volatility



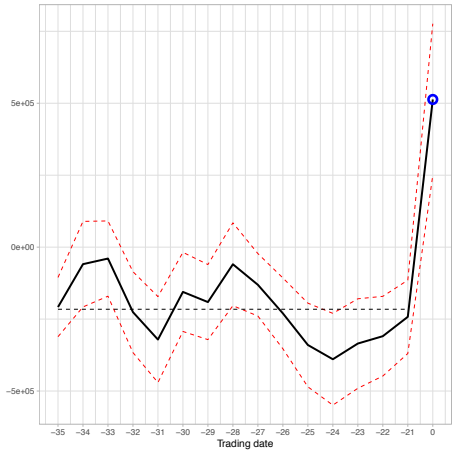
(b) Price Range



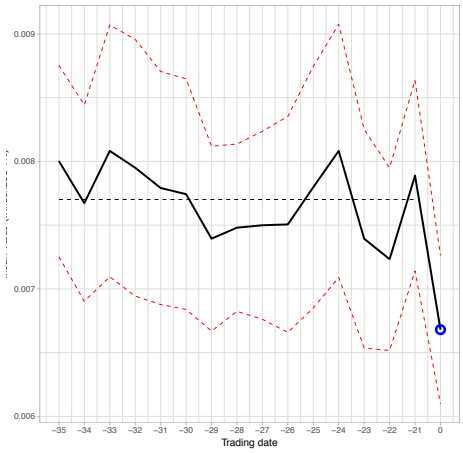
(c) Price Informativeness



(d) Abnormal Volume



(e) Quoted Spread



(f) Illiquidity

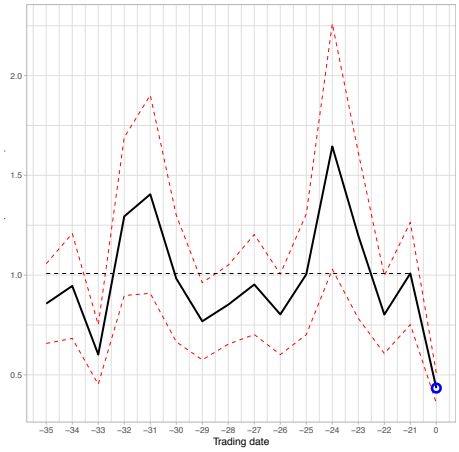
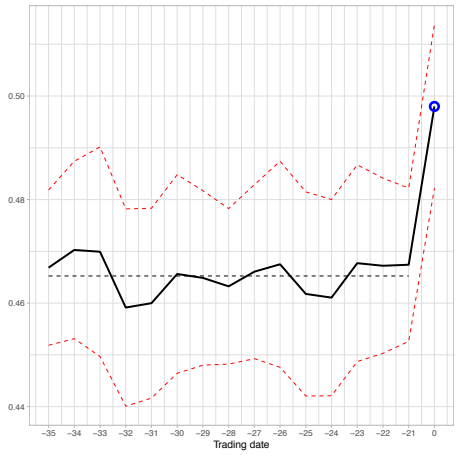


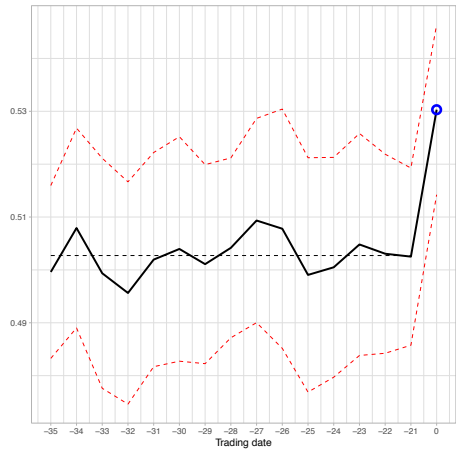
Figure 6. Option-based Signals: Daily Mean Values Around Informed Trading

The figure presents the average values (aggregated across all trades) of option-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days when $InfoTrade = 1$. The black dotted line is the mean value in the control window period of trading dates [-35,-21].

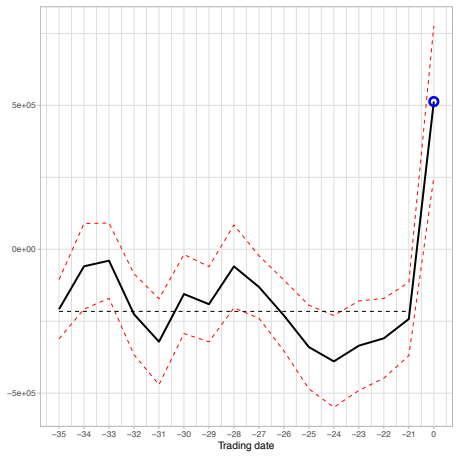
(a) IV Calls



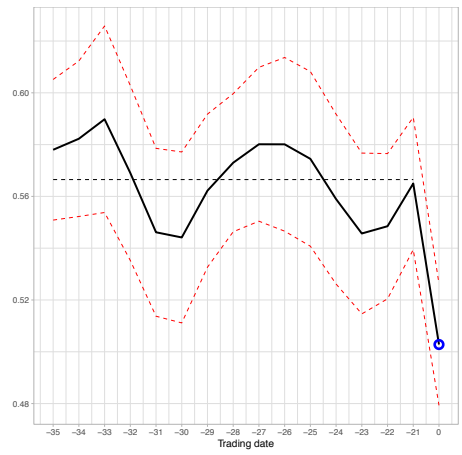
(b) IV Puts



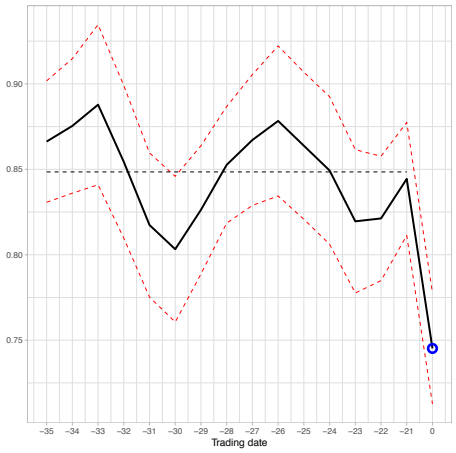
(c) Abnormal Volume



(d) Quoted Spread



(e) Quoted Spread - otm options



(f) Illiquidity

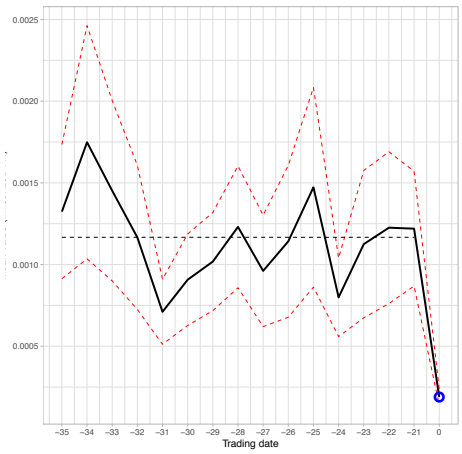
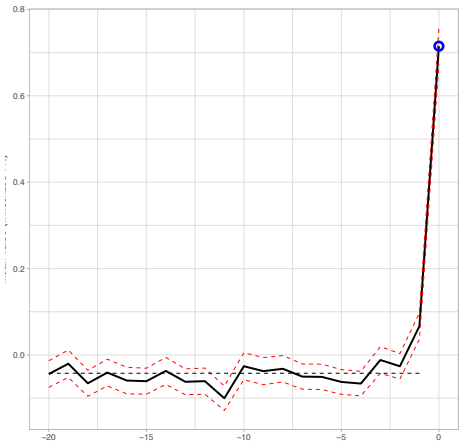


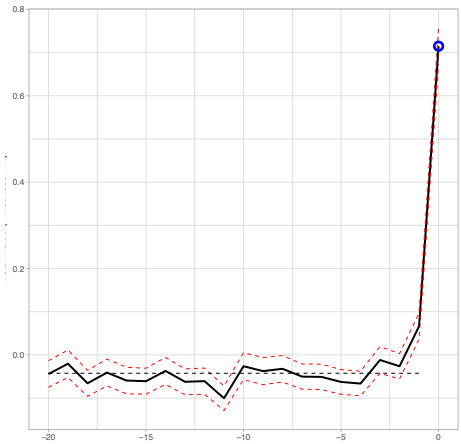
Figure 7. Stock-based Signals: Daily Mean Values Around Earnings Announcements

The sample is all earnings announcements for the period 1995–2015. The figure presents the average values (aggregated across all trades) of stock-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days of earnings announcements. The black dotted line is the mean value in the control window period of trading dates [-35,-21].

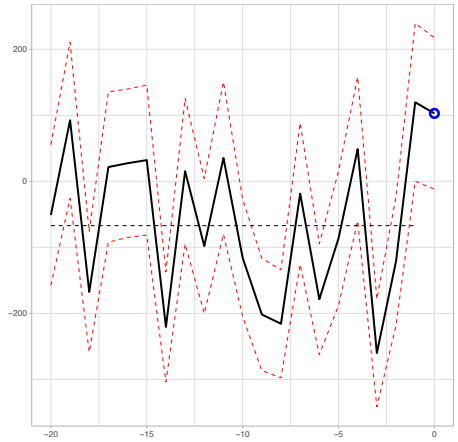
(a) Realized Volatility



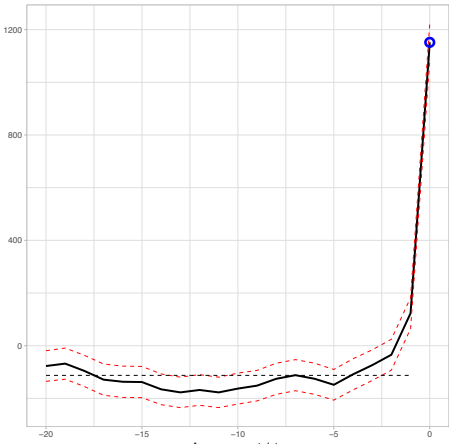
(b) Price Range



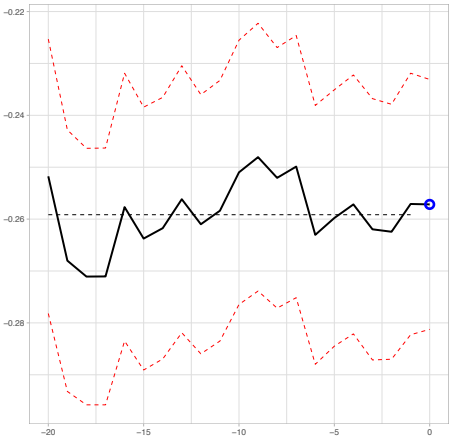
(c) Price Informativeness



(d) Abnormal Volume



(e) Quoted Spread



(f) Illiquidity

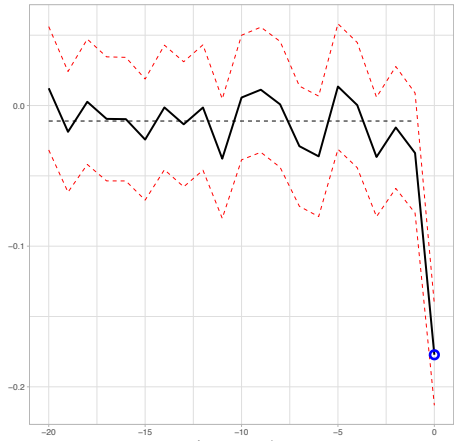
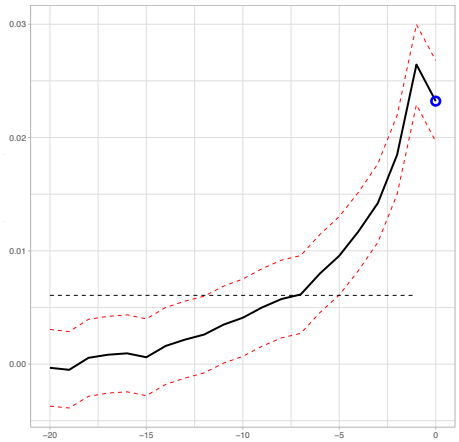


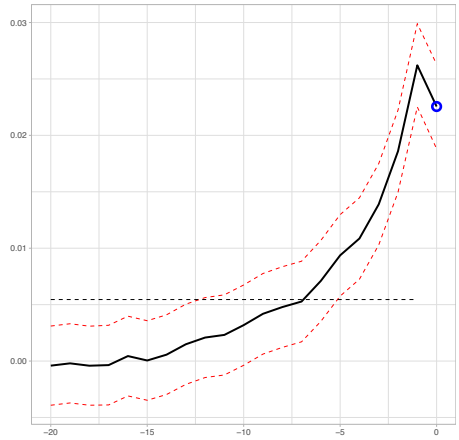
Figure 8. Option-based Signals: Daily Mean Values Around Earnings Announcements

The sample is all earnings announcements for the period 1995–2015. The figure presents the average values (aggregated across all trades) of option-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days of earnings announcements.The black dotted line is the mean value in the control window period of trading dates [-35,-21].

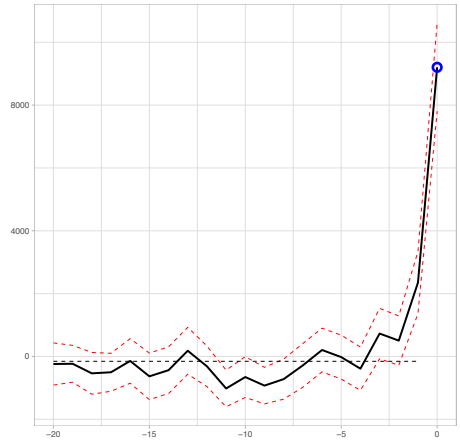
(a) IV Calls



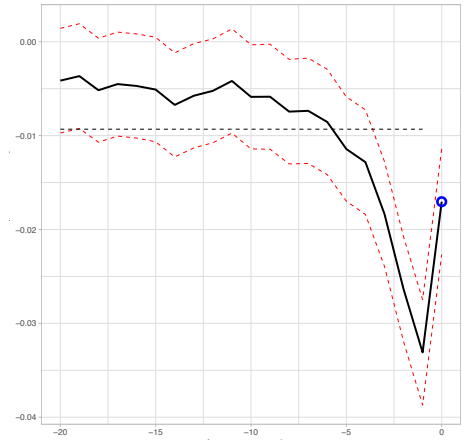
(b) IV Puts



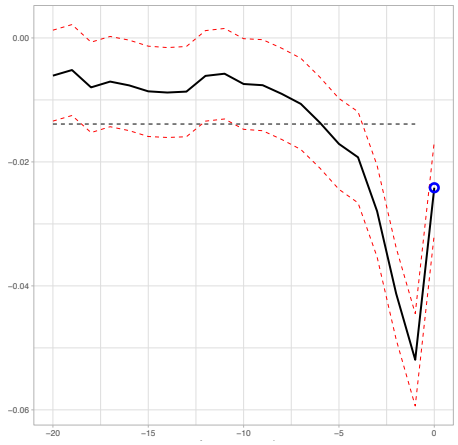
(c) Abnormal Volume



(d) Quoted Spread



(e) Quoted Spread - otm options



(f) Illiquidity

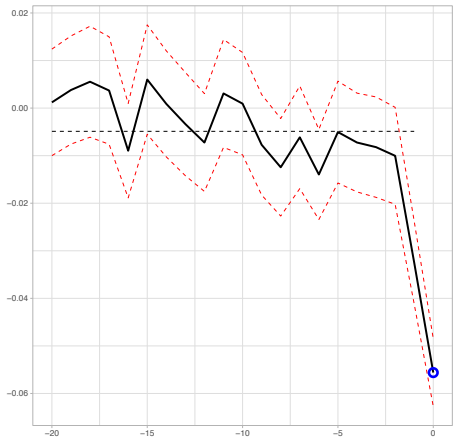
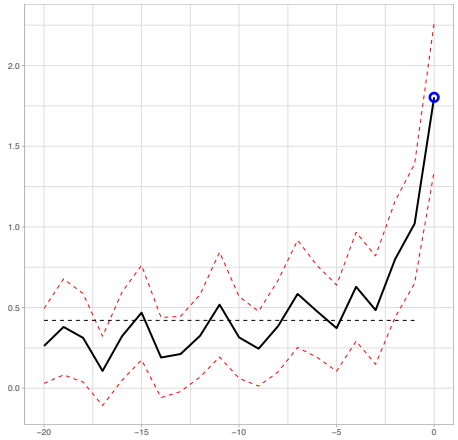


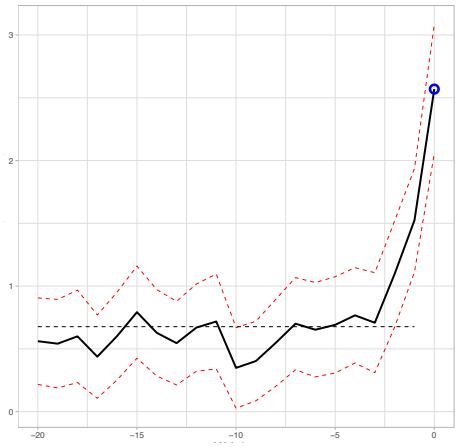
Figure 9. Stock-based Signals: Daily Mean Values Around M&As Announcements

The sample is all M&A announcements for the period 1995–2015. The figure presents the average values (aggregated across all trades) of stock-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days of earnings announcements. The black dotted line is the mean value in the control window period of trading dates [-35,-21].

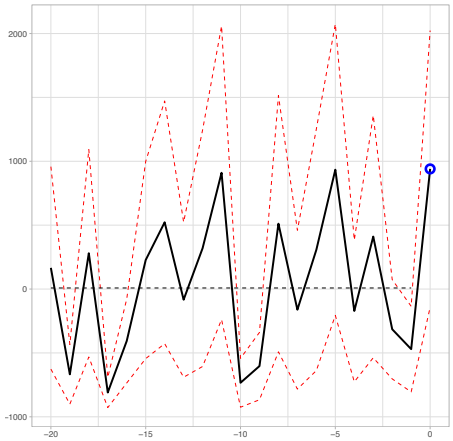
(a) Realized Volatility



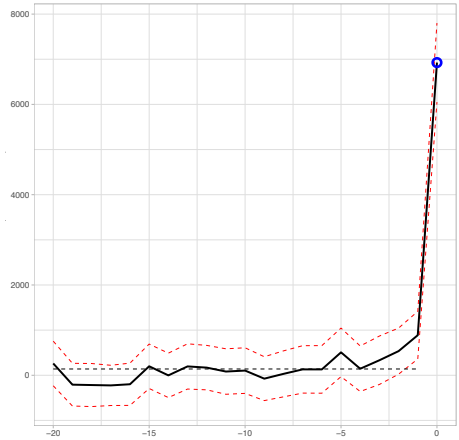
(b) Price Range



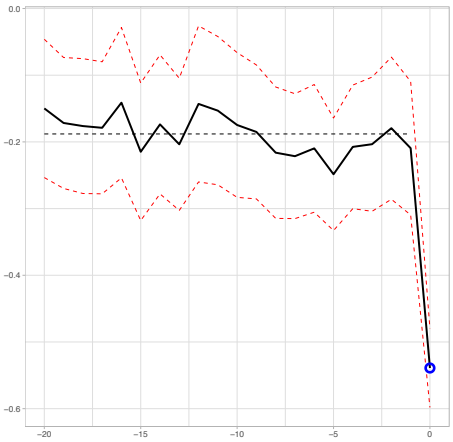
(c) Price Informativeness



(d) Abnormal Volume



(e) Quoted Spread



(f) Illiquidity

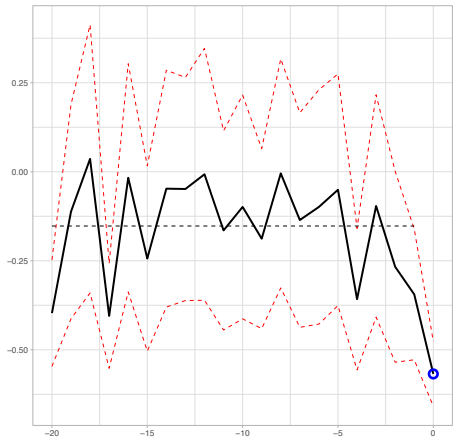
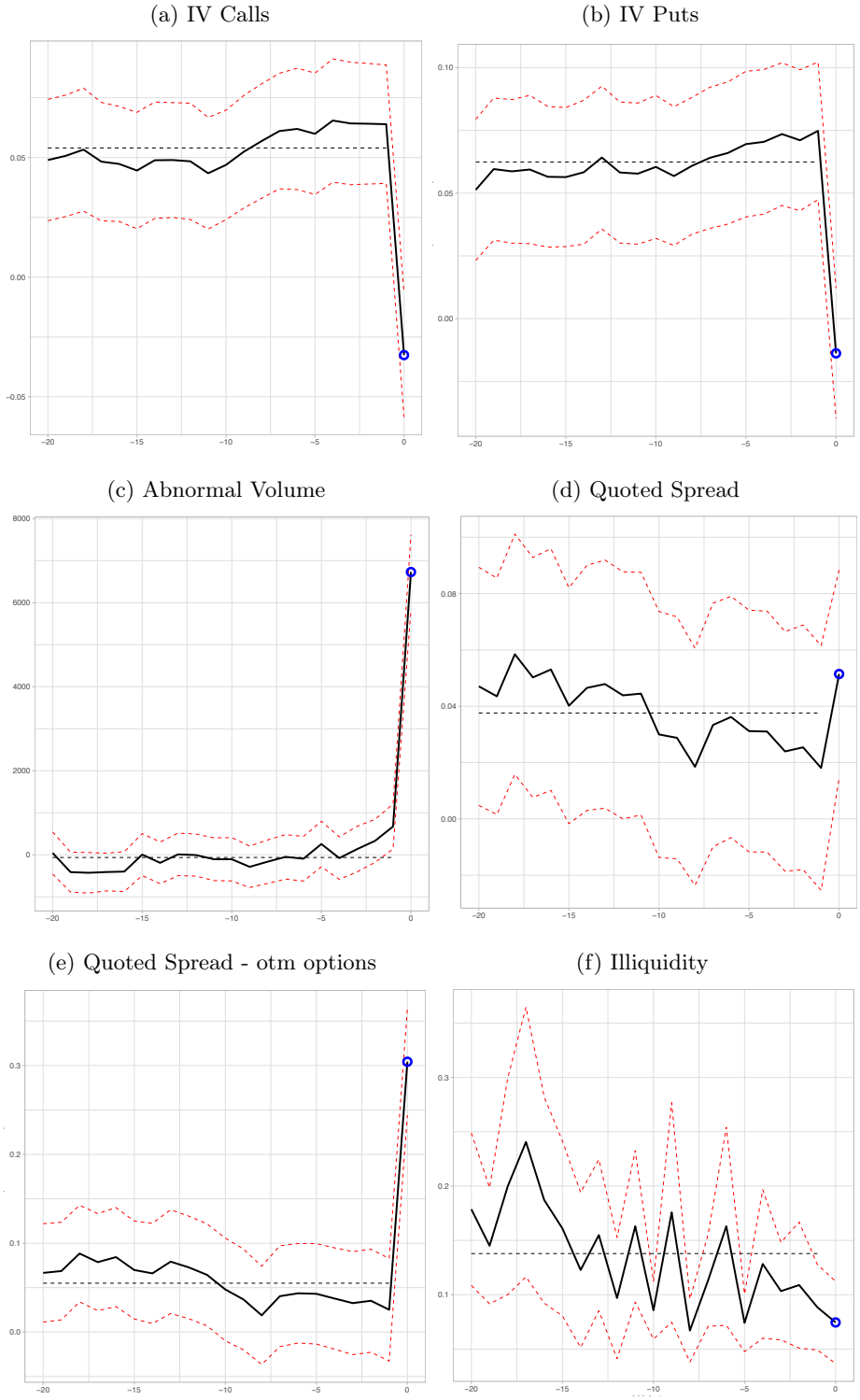


Figure 10. Option-based Signals: Daily Mean Values Around M&As Announcements

The sample is all M&A announcements for the period 1995–2015. The figure presents the average values (aggregated across all trades) of option-based signals, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 4. Date 0 corresponds to days of earnings announcements. The black dotted line is the mean value in the control window period of trading dates [-35,-21].



Online Supplement to "Chasing Private Information"

Marcin Kacperczyk*and Emiliano S. Pagnotta†

This version: April 10, 2018

A The Construction of Information Signals

In this appendix, we provide formal definitions of information signals and discuss their empirical implementation. For clarity, we group signals into volatility, volume, and illiquidity, and based on whether they originate in stock or option markets. Tables B1 and B2 provide summary statistics for the number of trades and information signals, respectively.

A.1 Information Signals in Stock Markets

Stock-based signals rely on high- and low-frequency data and are computed using monthly Trade and Quote (TAQ) and Center for Research in Security Prices (CRSP) data, respectively.

We compute the intra-day NBBO prices for each stock using the algorithm provided by Holden and Jacobsen (2014). The algorithm developed by these authors, first adjusts for withdrawn quotes and applies filters that eliminate nonsensical states due to data errors that could otherwise affect the precision of the NBBO quotes. Second, given the lack of intra-second time stamps in the monthly TAQ files, the algorithm exploits the *order* of trades and quotes within a given second and, through a process of interpolation, makes an educated guess about in which millisecond each event happened Holden et al. (2014) show that the so-called Interpolated Time method enhances inference based on monthly TAQ files.

We winsorize all signals at the 1% level to mitigate the influence of outliers. In addition to dollar-weighted averages, we also compute intraday stock-based signals using the number of shares as weights, obtaining similar results.

Stock Volatility

Realized Variance. We use a standard realized variance (RV) specification based on 30-minute squared log-returns. Log-returns are calculated using the prevailing mid-quote.

*Imperial College London and CEPR. Email: m.kacperczyk@imperial.ac.uk
†Imperial College London. Email: e.pagnotta@imperial.ac.uk

Price Range. We define the daily price range as

$$\text{Price Range}_t = \frac{a_{\text{max},t} - b_{\text{min},t}}{m_t},$$

where $a_{\text{max,t}}$ and $b_{\text{min,t}}$ denote the day-t maximum offer price and the minimum bid price, respectively; m_t is the arithmetic average of the two quantities. *Price Range* is a measure of price dispersion and can be affected by the value of the bid-ask spread, especially if prices are stale.

Price Informativeness. Roll (1988) argues that firm-specific variation is largely unassociated with public announcements and therefore largely due to trading by investors with private information. Extending the author's argument, we hypothesize that greater firm-specific variation indicates more intensive informed trading and, consequently, more informative pricing. We compute the measure of price informativeness $\left(1 - R_{it}^2\right)/R_{it}^2$ for each stock i on date t using the S&P500 exchange-traded fund (SPY) as the market index and considering 30-minute intervals during the trading day. Because of infrequent trades (a concern especially before 2001), we average prices over two minutes for each national BBO (NBBO) quote, for example, averaging over 9:59 a.m.–10:01 a.m. for 10 a.m.

Stock Volume

Abnormal Stock Volume. We compute the abnormal volume signal as

$$\text{Abn. S. Volume}_t = \text{S. Volume}_t - \text{Predicted S. Volume}_t,$$

where *S. Volume* is the total trading stock volume on day t. The variable *Predicted S. Volume* is computed using a linear regression model with *S. Volume* as a dependent variable and the following contemporaneous controls: median daily cross-sectional volume of all stocks, the Chicago Board Options Exchange Volatility Index (VIX), the excess return of the value-weighted market portfolio, and the daily stock return.[1]

Stock Illiquidity

Quoted Spread. Let t and k index trading dates and generic intra-day observations, respectively. The quoted bid-ask spread for a given stock is given by

$$\text{Quoted Spread}_t = \sum_{k=1:K} \omega_k \left(\frac{a_k - b_k}{m_k}\right),$$

where b and a denote the best bid and offer (BBO) quotes, $m \equiv \frac{1}{2}(a+b)$ denotes the midpoint, and ω_k represents a weight that is proportional to the amount of time that observation k is in-force.

[1] The predictive model's coefficients are computed over a time window of [-55,-15] trading days prior to the informed trade.

Price Impact. The five-minute price impact is given by

$$\text{Price Impact}_t = \sum_{k=1:K} 2\omega_k d_k \left[\ln\left(m_{k+5}\right) - \ln\left(m_k\right)\right],$$

where m_{k+5} is the midpoint of the consolidated BBO quotes prevailing five minutes after the k-th trade, d_k is the buy-sell trade direction indicator (+1 for buys, –1 for sells), and ω_k represents a dollar weight for the k-th trade. This signal represents the permanent component of the effective spread and, intuitively, it measures gross losses of liquidity demanders due to adverse selection costs.[2]

Absolute Order Imbalance. The absolute order imbalance is defined as

$$\text{Order Imb.}_t = \left|\frac{\text{Buys}_t - \text{Sells}_t}{\text{Buys}_t + \text{Sells}_t}\right|,$$

where Buys_t and Sells_t are the numbers of buys and sells over a given trading day t, respectively. We consider alternative trade-typing conventions to determine whether a given trade is sell or buy initiated. For brevity, we report the results using the Lee-Ready algorithm (1991) only.

Lambda. We follow Hasbrouck (2009) and Goyenko et al. (2009) and compute lambda as the slope coefficient in the following regression:

$$\text{Lambda}_t \text{ (slope): } r_n = \lambda \times (\sum_k d_k \sqrt{|vol_k|})_n + \text{error}_n$$

where, for the n-th time interval period on date t, r_n is the stock return, vol_k is transaction k-th's dollar volume, and the bracketed term represents the signed volume over interval n. Intuitively, the slope of the regression measures the cost of demanding a certain amount of liquidity over a given time period. We report the results based on five-minute intervals.[3]

Stock Illiq. For a given day t, it is given by the ratio between the absolute price return to dollar volume

$$\text{S. Illiq}_t = \frac{|\text{Stock Return}_t|}{\text{S. Volume}_t}.$$

Intuitively, a liquid stock is one that experiences small price changes per unit of trading volume. Amihud's (2002) *ILLIQ* can be seen as a monthly average of the daily measure.

A.2 Information Signals in Option Markets

We obtain option data from the Ivy DB OptionMetrics database, which provides end-of-day information for all exchanged-listed options on U.S. stocks, including option prices, volumes, and IVs. For a given underlying stock, we consider all call and put option contracts and construct each daily metric using open-interest-weighted averages. All signals are winsorized at the 1% level to mitigate the influence of

[2] Two related signals are the effective spread and the realized spread. We tested these measures and the results are very similar to those of the price impact and are thus omitted.

[3] We also computed *Lambda* and the realized variance based on slightly different intervals, obtaining similar results.

outliers. Panel B of Table B2 provides summary statistics.

Option Volatility

IV Calls and IV Puts. Let $j = 1, ..., J$ denote a strike-maturity combination for calls and puts on the same underlying stock. For both calls and puts, the daily IV is computed as an open-interest-weighted average (with weight ω_j for option j) of OptionMetrics' IVs, respectively:

$$\text{IV Calls}_t = \sum_{j=1:J} \omega_j OMIV_j^{CALL},$$
$$\text{IV Puts}_t = \sum_{j=1:J} \omega_j OMIV_j^{PUT}.$$

IV Skew. Following Cremers and Weinbaum (2010), we compute the *IV skewness* measure on a given day t as

$$\text{IV Skew}_t = \sum_{j=1:J} \omega_j \left| OMIV_j^{CALL} - OMIV_j^{PUT} \right|.$$

Only pairs with IV and open interest records are included in the calculation.

Option Volume

Abnormal Volume. We follow Augustin et al. (2015) and compute a measure of abnormal volume in options. For all active contracts in a given underlying company, we calculate

$$\text{Abn. O. Volume}_t = \text{Volume}_t - \text{Predicted Volume}_t,$$

where *Volume* is the number of traded contracts on day t and *Predicted Volume* is computed using a linear regression model with *O. Volume* for the same underlying and the following contemporaneous controls: the median volume in all equity options, the VIX, the excess return of the value-weighted market portfolio, and the daily return of the underlying stock.[4]

Volume Ratio otm/all. We compute the ratio of the volume in otm options to total option volume. Specifically, for all options with the same underlying stock, we have

$$\text{Vol. Ratio (otm/all)}_t = \frac{\text{otm Volume}_t}{\text{Volume}_t}.$$

Option Illiquidity

Quoted Spreads. The daily quoted bid-ask spread is defined as

$$\text{Quoted Spr. (all)}_t = \sum_{j=1:J} \omega_j \left(\frac{a_{jt} - b_{jt}}{m_{jt}} \right),$$

[4]The predictive model coefficients are computed over a time window of [-55,-15] trading days prior to the informed trade.

where the bid and ask quotes correspond to values at the end of the day.

We also consider a version that concentrates on highly levered (out-of-the-money) options, Quoted Spr. (otm).

Option Illiquidity. We extend the reach of the illiquidity measure to options as follows

$$\text{O. Illiq}_t = \frac{|\text{Option Return}_t|}{\text{Option Volume}_t},$$

where *Option Volume* accounts for the volume on dat t in all options of the same underlying and Option Return is computed as the percentage daily change in the OMIV of a particular contract. We believe this is a reasonable approximation to option returns over a short period of one trading day.

B Supplemental Summary Statistics and Results

Table B1

Daily Trades in the U.S. Stock Market: Descriptive Statistics

This table reports summary statistics for stock trades calculated across time and firms for small (**Panel A**) and large (**Panel B**) market caps, respectively. The sample period is 1995–2015.

	Panel A: Small Cap Stocks			**Panel B: Large Cap Stocks**		
Year	mean	median	st.dev.	mean	median	st.dev.
1995	20.3	6	57.7	173.8	52	498.1
1996	24.2	8	75.9	204.7	59	630.7
1997	26.4	9	81.5	259.8	66	1,025.0
1998	38.0	11	230.7	390.0	89	1,726.6
1999	61.9	14	439.0	856.2	184	2,942.7
2000	73.8	18	299.5	1,478.3	320	5,090.1
2001	61.0	14	230.0	1,804.5	322	6,001.8
2002	71.7	16	268.9	1,984.6	479	5,867.5
2003	120.8	24	468.3	2,278.7	692	5,942.2
2004	191.1	36	992.6	2,392.9	796	6,839.3
2005	204.6	42	890.8	2,583.1	869	7,555.5
2006	204.7	43	731.9	3,004.6	1054	8,333.6
2007	232.2	53	792.6	3,721.1	1298	9,917.8
2008	273.5	57	671.5	6,146.2	1904	16,896.5
2009	350.2	78	1,248.5	7,226.2	2100	16,842.2
2010	333.1	78	1,149.6	6,084.3	1758	15,168.4
2011	346.9	67	1,073.8	6,328.3	1909	15,449.4
2012	312.7	61	1,049.8	5,565.8	1656	13,624.9
2013	346.8	69	1,255.9	5,058.4	1603	12,900.0
2014	425.1	96	1,845.7	5,869.4	2056	14,848.7
2015	409.2	90	1,926.8	6,013.8	2174	15,318.5

Table B2
Information Signals: Descriptive Statistics

Panel A reports the mean, median, and standard deviation calculated across time and firms of stock-based information signals for the period 1995–2015. **Panel B** refers to option-based measures. The information signals in Panels A and B are defined in Appendix A. All information signals are winsorized at 1%.

Variable	mean	median	st.dev.
Panel A: Stock-based signals			
Realized volatility (30 min; % per day)	1.34	0.45	2.95
Price range	0.499	0.395	3.85
Price informativeness (30 min)	959.27	6.11	8,641.59
Abnormal volume	68.73	(11.40)	3,017.85
Quoted spread *100	0.56	0.20	0.94
Price impact *100	10.67	4.68	17.29
Absolute order imbalance	0.15	0.11	0.16
Lambda	0.15	0.02	0.33
Illiquidity	0.57	0.03	2.88
Panel B: Option-based signals			
Implied volatility calls	0.55	0.48	0.26
Implied volatility puts	0.59	0.50	0.27
Implied volatility skewness	(0.01)	(0.01)	0.05
Abnormal volume	370.88	(26.93)	9,615.37
Volume ratio (otm/all options)	28.55	2.27	43.00
Quoted spread (all options)	0.58	0.48	0.37
Quoted spread (otm options)	0.84	0.74	0.48
Illiquidity *100	0.13	0.00	0.63

Table B3
Conditioning on Signal Strength

This table presents the results conditioning on the strength of the information tip received by the trader. Strength corresponds to the stock return (excluding dividends) computed from the opening price on the first insider trading day to the opening price on the day following the information disclosure. The dependent variables are information signals. **Panel A** reports the results for stock-based signals, **Panel B** the results for option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Panel A: Stock-based Signals

Based on	Volatility			Volume		Illiquidity			
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
InfoTrade	0.388***	0.978***	1,009.591*	451.890*	-0.112***	0.283	-0.012	-0.044***	-0.302**
	(0.149)	(0.218)	(570.684)	(232.099)	(0.034)	(0.697)	(0.007)	(0.013)	(0.130)
Firms	-0.002	-0.002	3.643	-0.620	0.002***	0.015***	0.000**	0.001***	0.001
	(0.001)	(0.001)	(3.119)	(0.562)	(0.000)	(0.005)	(0.000)	(0.000)	(0.001)
InfoTrade*Strength	0.009**	0.005	-3.947	-1.299	0.000	-0.010	-0.000	0.000	-0.003*
	(0.003)	(0.004)	(4.051)	(2.054)	(0.001)	(0.011)	(0.000)	(0.000)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	9,475	9,493	9,424	9,453	9,467	9,421	9,421	9,407	9,436

Panel B: Option-based Signals

Based on	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (otm)	O. Illiq.
InfoTrade	0.007	0.006	0.018	2,853.107*	1.001	-0.015	-0.022	-0.090
	(0.020)	(0.019)	(0.012)	(1,518.655)	(3.961)	(0.058)	(0.069)	(0.075)
Firms	-0.001	-0.001**	0.001	-58.506	0.240	0.003	0.003	0.007
	(0.001)	(0.001)	(0.000)	(47.705)	(0.154)	(0.004)	(0.004)	(0.005)
InfoTrade*Strength	0.001	0.000	-0.000	16.035	-0.175	-0.002	-0.002	-0.001
	(0.000)	(0.000)	(0.000)	(28.353)	(0.136)	(0.001)	(0.002)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,326	2,314	2,147	2,331	2,331	2,276	2,276	2,232

Table B4
Conditioning on the Number of Firms in the Investigation

This table presents the results conditioning on the number of firms in the investigation. The dependent variables are information signals. **Panel A** reports the results for stock-based signals and **Panel B** option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Panel A: Stock-based Signals

Based on	Volatility			Volume	Illiquidity				
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
InfoTrade	0.736***	1.197***	745.068	625.598**	-0.114***	-0.348	-0.019**	-0.060***	-0.234**
	(0.199)	(0.249)	(619.008)	(260.443)	(0.029)	(0.880)	(0.008)	(0.015)	(0.098)
n_{high}	-0.676*	-0.795	-699.473	-284.722	0.038	-0.727	-0.017	-0.008	0.048
	(0.373)	(0.516)	(750.913)	(418.557)	(0.145)	(1.820)	(0.015)	(0.050)	(0.143)
InfoTrade*n_{high}	-0.082	-0.123	78.623	-449.563	0.028	0.497	0.007	0.051**	-0.349
	(0.328)	(0.387)	(899.766)	(307.110)	(0.055)	(1.116)	(0.013)	(0.021)	(0.214)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,336	10,375	10,266	10,335	10,258	10,208	10,208	10,193	10,300

Panel B: Option-based Signals

Based on	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (otm)	O. Illiq.
InfoTrade	0.013	0.011	0.006	1,497.659	-3.693	-0.080*	-0.099	-0.210**
	(0.014)	(0.017)	(0.005)	(968.797)	(4.079)	(0.045)	(0.061)	(0.083)
n_{high}	-0.003	-0.208	-0.143***	-43,929.191**	22.646	-4.136*	-5.216*	-1.717**
	(0.442)	(0.270)	(0.034)	(17,485.635)	(118.984)	(2.427)	(3.046)	(0.711)
InfoTrade*n_{high}	0.016	0.009	-0.002	3,303.264	-0.806	0.024	0.041	0.161*
	(0.021)	(0.022)	(0.014)	(2,390.400)	(4.999)	(0.056)	(0.071)	(0.084)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,600	2,588	2,349	2,605	2,605	2,550	2,550	2,503

Table B5
Conditioning on Market Capitalization: SEC WRP Cases

This table presents results for firms with small and large stock market capitalization. The sample is the SEC WRP cases. The dependent variables are information signals. **Panels A and B** report the results for stock-based signals and **Panels C and D** the results for option-based signals. $InfoTrade$ is an indicator variable equal to one for asset–day pairs with informed trading. The definitions of $InfoTrade$ and control variables are in Section 4. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. ***, **, * denote 1%, 5%, and 10% level of statistical significance, respectively.

Based on	Volatility			Volume		Illiquidity			
	Realized Volatility	Price Range	Price Inform.	Abn. S. Volume	Quoted Spread	Price Impact	Order Imb.	Lambda	S. Illiq.
Panel A: Stock-based Signals: Small Caps									
InfoTrade	0.534	0.974**	690.414	-92.984	-0.041	0.194	-0.012	0.021	-0.014
	(0.375)	(0.470)	(1,469.032)	(91.725)	(0.037)	(2.508)	(0.014)	(0.038)	(0.074)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	703	703	698	702	683	683	683	683	701
Panel B: Stock-based Signals: Large Caps									
InfoTrade	0.184	0.836***	-334.999	975.210	-0.072	0.268	0.002	-0.020***	0.019
	(0.000)	(0.275)	(212.411)	(0.000)	(0.000)	(0.533)	(0.000)	(0.007)	(0.015)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	950	950	950	950	950	950	950	950	950

	Volatility			Volume		Illiquidity		
	IV Calls	IV Puts	IV Skew	Abn. O. Volume	Vol. Ratio (otm/all)	Quoted Spr. (all)	Quoted Spr. (otm)	O. Illiq.
Panel C: Option-based Signals: Small Caps								
InfoTrade	-0.037	-0.054***	-0.028	440.022	4.435	-0.384*	0.018	-0.808**
	(0.034)	(0.013)	(0.020)	(1,827.332)	(30.538)	(0.216)	(0.446)	(0.385)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	49	47	49	49	49	49	49	43
Panel D: Option-based Signals: Large Caps								
InfoTrade	0.019*	0.003	3,614.958**	-3.146	-0.127***	-0.178***	-0.009	-0.119**
	(0.010)	(0.004)	(1,814.710)	(5.469)	(0.038)	(0.057)	(0.014)	(0.046)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	347	324	347	347	330	330	330	2,087

References

Amihud, Y. (2002): "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5, 31–56.

Augustin, P., M. Brenner, and M. Subrahmanyam (2015): "Informed Options Trading Prior to MA Announcements: Insider Trading?" *Working Paper*.

Cremers, M. and D. Weinbaum (2010): "Deviations from Put–Call Parity and Stock Return Predictability," *Journal of Financial and Quantitative Analysis*, 45, 335–367.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009): "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics*, 92, 153–181.

Hasbrouck, J. (2009): "Trading costs and returns for U.S. Equities: Estimating Effective Costs from Daily Data," *Journal of Finance*, 64, 1445–1477.

Holden, C. W., S. Jacobsen, and A. Subrahmanyam (2014): "The Empirical Analysis of Liquidity," *Foundations and Trends in Finance*, 8, 263–365.

Lee, C. M. C. and M. J. Ready (1991): "Inferring Trade Direction from Intraday Data," *The Journal of Finance*, 46, 733–746.

Roll, R. (1988): "R Squared," *Journal of Finance*, 43, 541–566.