12-2019

Chasing private information

Marcin KACPERCZYK

Emiliano Sebastian PAGNOTTA
Singapore Management University, epagnotta@smu.edu.sg

Citation

KACPERCZYK, Marcin and PAGNOTTA, Emiliano Sebastian. Chasing private information. (2019). *Review of Financial Studies*. 32, (12), 4997-5047.

Available at: https://ink.library.smu.edu.sg/lkcsb_research/7027

Published in Review of Financial Studies, 2022 Dec, 32 (12), 4997–5047. DOI: 10.1093/rfs/hhz029

Chasing Private Information

Marcin Kacperczyk
Imperial College London and CEPR

Emiliano S. Pagnotta
Imperial College London

Using over 5,000 trades unequivocally based on nonpublic information about firm fundamentals, we find that asymmetric information proxies display abnormal values on days with informed trading. Volatility and volume are abnormally high, whereas illiquidity is low, in equity and option markets. Daily returns reflect the sign of private signals, but bid-ask spreads are lower when informed investors trade. Market makers' learning under event uncertainty and limit orders help explain these findings. The cross-section of information duration indicates that traders select days with high uninformed volume. Evidence from the U.S. SEC Whistleblower Reward Program and the FINRA involvement addresses selection concerns. (*JEL* D82, D83, G10, G12, G14)

Asymmetric information is ubiquitous in financial markets, because investors have unequal knowledge of firms' fundamentals. The literature widely accepts that the presence of informed market participants affect the behavior of

For helpful, detailed, and insightful comments we are grateful to Andrew Karolyi (editor) and two anonymous referees. We also benefitted from conversations with Franklin Allen, Yakov Amihud, Alejandro Bernales, Utpal Bhattacharya, Bruno Biais, Ekkehart Boehmer, Alon Brav, Andrea Buraschi, Adrian Buss, Lauren Cohen, Slava Fos, Thierry Foucault, Ross Garon, John Griffin, Diane Del Guercio, Luigi Guiso, Joel Hasbrouck, 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 Roşu, Gideon Saar, Norman Shurhoff, Zheng Sun, Paul Tetlock, Michela Verardo, Pietro Veronesi, and Marius Zoican; seminar participants at the Tinbergen Institute, ICBS, MIT Sloan, Cass Business School, INSEAD, 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 participants 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 (LSE), 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. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Emiliano Pagnotta, ICBS, Tanaka Building, South Kensington Campus, London SW7 2AZ; telephone: +44 (0)20 7594 0939. E-mail: e.pagnotta@imperial.ac.uk.

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 asymmetric information proxies (AIPs), such as volume or market prices, under the assumption that these bear a specific relation with the unobserved information asymmetry. For example, higher bid-ask spreads or trade price impact could indicate higher risk of informed trading.[1] While AIPs 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 premiums, 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 AIPs is difficult: it requires studying their distribution conditional on the actions of informed traders whose *unobserved* presence these AIPs *chase* in the first place.

We address this identification challenge by studying the conditional behavior of a broad set of AIPs linked to trades that are *unequivocally based on nonpublic information* about firms' 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 manager 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 5,058 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 materially affects AIPs, the AIPs should display abnormal behavior on precisely identified days with informed trading relative to a sample of random dates.

Guided by prior theoretical and empirical research, we consider three groups of AIPs: based on volatility, volume, and illiquidity. Although most empirical research has relied on stock market–based AIPs, informed traders could take advantage of options (Black 1975). We additionally study the behavior of AIPs in option markets.[3] We can specify not only the exact dates of informed trades

---

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

2 As an example, our sample contains famous hedge fund trading cases that resemble this narrative (e.g., R. Rajaratnam at the Galleon fund in 2011 and M. Martoma at SAC Capital in 2012). It also contains individuals with small amounts of available capital, wealthy individual investors, and different institutional investors.

3 Standard theory suggests that metrics reflecting volatility, volume, or illiquidity are affected by the presence of informed traders and thus can act as AIPs. Such AIPs can originate in stock markets (e.g., Kyle 1985;

but also the trading instrument informed traders use. Thus, we can identify the presence of informed traders both across time *and* markets.

Our first result is that AIPs largely display abnormal behavior on days with informed trading: the changes in the average values on days with informed trading are statistically significant. Moreover, AIPs exhibit stronger reactions on days with a large proportion of informed trades. Second, AIPs that originate in option markets are valuable. Relative to stock-based AIPs, these AIPs 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, whereas, 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, *Bid-ask spread* decreases by about 10% in stock markets and by 20% in 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 the total amount for stocks, and more than 30% for options. Second, if price discovery takes place, one would expect prices to respond in the direction 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 informed traders' information sets by computing the hypothetical stock returns (excluding dividends) that such a trader would realize if he initiated a trade at the opening price of the day of his 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 AIPs 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, and the median number of days between the first and the last informed trade is eight. Overall, if AIPs 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 AIP if asymmetries were smaller.

Our sample solely includes trades originating from SEC investigations, so one could be concerned about a potential selection bias. One pressing concern

---

Glosten and Milgrom 1985; Easley and O'Hara 1987; Wang 1993) or in option markets (e.g., Back 1993; Biais and Hillion 1994; Easley, O'Hara, and Srinivas 1998).

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

is that insider traders are *only* exposed when AIPs 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 AIPs to react to the presence of informed traders. We argue in Section 4 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 a framework were in place, the results in Section 3 fail to support this view. For example, several AIPs display patterns that are arguably inconsistent with what economic reasoning would identify as patterns of informed trading, especially the fact that illiquidity AIPs 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 AIPs to flag an asset-date pair, it would be difficult for its officials to identify systematically which individuals are breaching the law due to trade aggregation, netting, use of multiple accounts, and so forth.

We buttress the problem of selection bias formally with three separate tests. First, we classify the origin of each investigation during the period of 2011-2015 *directly* using the 2010 adoption of the SEC Whistleblower Reward Program (WRP). This program 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 AIPs is mostly insensitive to the origin of investigations. Second, we exploit legal investigation heterogeneity in the involvement of self-regulatory organizations, such as the Financial Industry Regulatory Authority (FINRA) and the Options Regulatory Surveillance Authority (ORSA). Unlike the SEC, SROs continuously monitor trade data feeds and search for a potentially illegal activity. We find that cases in which SROs are not involved exhibit patterns similar to our full sample results. The third test indicates that the conditional behavior of AIPs is the same for investigations with a small and a large number of firms. Using the argumentation of Meulbroek (1992), this result points against selection bias based on the origin of the investigation. Overall, our results strongly suggest that our findings are unlikely to reflect selection bias.

One of our key findings is the negative response of illiquidity measures to informed trading, in stock and option markets. We explore three plausible explanatory channels: the strategic timing of informed trades, the use of limit orders by informed traders, and the learning of market makers under event uncertainty. 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, first, we decompose the trading volume into informed and uninformed components. Consistent with the timing channel, we find that uninformed volume accounts for a significant fraction of the abnormal stock volume on days with informed traders, 51.7%. For options, this proportion is higher at 84.6%, consistent with the view that insiders have stronger incentives to time the trades of the more illiquid trading instrument. Second, 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 (up to 3 days), the ability to optimally time trades should be more constrained. The results in Section 5 support the timing hypothesis. Indeed, for short-horizon cases, in contrast to the full sample, some illiquidity AIPs display near-zero or positive values.

Our second mechanism relates to the use of limit orders by informed traders (e.g., Biais, Glosten, and Spatt 2005), the force which could lead to lower bid-ask spreads. 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, which is significantly more than what one would expect to find in a sample of uninformed investors. Further, we show that a sample of small-cap stocks exhibit significantly lower values of bid-ask spreads and order imbalance measures, which could be rationalized by insiders' greater use of limit orders in this sample.[5]

Our third mechanism explores the role of market makers' learning process in explaining our results. First, we study the implications of learning under uncertainty about the presence of informed traders (Easley and O'Hara 1992) on the behavior of the bid-ask spread. We hypothesize that if the trades of the informed lead to a rapid resolution of uncertainty, in principle, the bid-ask spread could decrease fast enough on days with informed trading and generate lower average values. To evaluate the plausibility of this conjecture and better understand its reach, we match the parameters of the model to several market characteristics during our sample period and simulate market sessions. We show that the conditional response of the average bid-ask spread crucially depends on prior beliefs about the probability of informed trading (PIN). When PIN is sufficiently high, the average value of the bid-ask spread is *lower* on days with private information. We relate the quote responses to how the market maker processes innovations in abnormal volume and order imbalance. The model's implications are corroborated by our empirical findings. We find that the conditional response of the bid-ask spread is negative on average, but the effect is significantly stronger for small-cap stocks. Arguably, small stocks can be seen as stocks with relatively higher PIN values. In a second test of the

---

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).

market making channel, we evaluate our empirical results conditioning on the directional response of prices relative to the sign of the private signal. We find that the responses of volatility and illiquidity measures are stronger when price movements reflect private information, consistent with a learning channel.

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 the literature that examines information content in stocks and options. We provide a detailed discussion of these implications in Section 9. Very few studies have addressed the issue of whether different types of AIPs help to identify informed traders across markets.[6] Our ability to observe the arrival of private information directly is in stark contrast to prior literature that infers the presence of informed trading indirectly by, for example, 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 and, in contrast to trades by 13D filers, we can observe the content of information sets. Finally, we provide a more rigorous evidence on the role of market makers and their learning for the illiquidity effect. Meulbroek (1992) was the first to use the information in SEC insider trading investigations and studied stock market efficiency.[8] Our focus, instead, is on the joint distribution of AIPs and the (usually unobserved) presence of informed traders, in both stock and option markets. We document that option markets are used extensively by informed traders over the 20-year period we analyze. Therefore, our results provide support for the conjecture of

---

[6] We note that not all empirical analyses in the literature rely on AIPs 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 example, Boulatov, Hendershott, and Livdan (2013) and Hendershott, Livdan, and Schurhoff (2015) use institutional order flow, Cohen, Malloy, and Pomorski (2012) study routine corporate insiders SEC filings, and Kacperczyk, van Nieuwerburgh, and Veldkamp (2016) use a model with endogenous information acquisition to infer private information in a sample of mutual funds.

[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 perspective, 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 about 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 deemphasize option markets.

[8] More recently, Del Guercio, Odders-White, and Ready (2017) study the effect of a changing regulatory environment on price discovery and Kacperczyk and Pagnotta (2018) study the relation between insider's trading strategies, enforcement risk, and information aggregation into prices.

Black (1975) and several theoretical analyses such as Back (1993). Moreover, our results show that the conditional patterns of AIPs behavior in option markets are consistent with those observed in stock markets.

1. Main Test and AIPs

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

1.1 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 $\mathbb{E}(Y|I) \neq \mathbb{E}(Y)$, for example, **P**, $\mathbb{E}(Y|I = 1) > \mathbb{E}(Y)$. An econometrician is interested in testing **P**; however, I is not observable. The econometrician then considers one or more proxies, $AIP: \Omega \to \mathbb{R}$, and makes an identification assumption, **A**, that relates I and the AIP. Typically, **A** can be expressed as $AIP = \text{constant} + \Delta I + \text{noise}$, $\Delta \in \mathbb{R}$. If $\Delta > 0$ is assumed, the econometrician could determine a threshold value π and replace I by $\hat{I}(\omega) = \mathbb{I}\{AIP(\omega) > \pi\}$, where $\mathbb{I}$ is an indicator function. Proposition **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 $\mathbb{E}(Y)$.[9]

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

1.2 A test of the reliability of AIPs

To shed light on the reliability of the identification assumption, **A**, one would ideally design a test on $\mathbb{E}(AIP|I)$. By the law of iterated expectations,

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

As is common in the literature (e.g., Easley and O'Hara 1992), we assume that privately informed traders are not trading every single period, so $\mathbb{P}(I = 1) < 1$.

[9] In a similar fashion, assumption **A** could be stated as $\frac{\partial}{\partial p_i}\mathbb{E}(AIP|p_i) > 0$, 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 (AIP, I).

Thus, the PIN is less than one, which implies $\mathbb{E}[API|I=1]\neq\mathbb{E}[AIP]$. We can then succinctly express the value of the AIPs as

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

where Δ represents the impact of informed trading on the AIPs value relative to its unconditional mean.[10] We hypothesize that, if the firm-specific proxy 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 **A** for a given AIPs thus has a null hypothesis $\Delta=0$ against an alternative hypothesis, $\Delta\neq0$.

A test of Δ in Equation (1) generally is not feasible, given the unobservability of I (which motivated the use of the AIPs 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 2). We define the binary variable $InfoTrade$ to be equal to 1 when an informed trades (see Section 3 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, $InfoTrade=1\Rightarrow I=1$, and therefore to empirically assess the relative degree of an AIPs's reliability.

1.3 AIPs of private information-based trading

Our tests investigate the behavior of three types of AIPs: 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, Morck, and Yeung 2004) in stock market and call/put implied volatility and implied volatility skewness (e.g., Cremers and Weinbaum 2010) in option market. We construct a measure of abnormal trading 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 daily volume ratio between out-of-the-money (OTM) options and all options. We use the following illiquidity AIPs: the percentage quoted bid-ask spread (*Bid-ask spread*); the 5-minute price impact (e.g., Goyenko, Holden, and Trzcinka 2009), *Price impact*; the absolute order flow imbalance,[11] *Order 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,

---

[10] We note that $\mathbb{E}[AIP|I=0]$ is difficult to identify in a nonlaboratory setting, because identification would require access to *every* trader's information set.

[11] 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, because 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 when computing the *PIN* log-likelihood function.

Table 1
Matrix of AIPs

AIP type/Market	Stocks	Stock options
Volatility	Price range, Realized variance Price informativeness	Implied vol. (calls and puts) IV skewness
Volume	Abnormal s. volume	Abnormal o. volume Volume ratio (OTM/all)
Illiquidity	Bid-ask spread, Price impact, Order imb., Lambda, S. illiq	Bid-ask spread (all, OTM) O. illiq

Table 2
Descriptive statistics for AIPs

AIP	Mean	Median	SD
A. Stock-based AIPs			
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
Bid-ask 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
B. Option-based AIPs			
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
Bid-ask spread (all options)	0.58	0.48	0.37
Bid-ask spread (OTM options)	0.84	0.74	0.48
Illiquidity *100	0.13	0.00	0.63

Panel A reports the mean, median, and standard deviation calculated across time and firms of stock-based AIPs for the period 1995–2015. Panel B refers to option-based measures. Section A of the Online Appendix defines the AIPs in panels A and B. All AIPs are winterized at 1%.

we use a daily version of Amihud's 2002 *Illiq* measure for both stocks and options (*S. Illiq* and *O. Illiq*, respectively). Table 1 summarizes the set of AIPs, and Table 2 presents summary statistics.

2. 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.

2.1 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 securities 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 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 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: (a) the trade under consideration is motivated by actual information, as opposed to, say, sentiment, and (b) 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 (a) and (b) were met, regardless of their legal classification.[12]

2.2 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 Web site. In cases in which a complaint file is not available on the SEC Web site, 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 resultant sample represents all SEC cases either litigated or settled out of court. Most complaint files include a detailed account of the allegations. The documents provide most of the relevant information in a textual form, so 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.

[12] 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.

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 was released to the general public. We also collect information on the date of information transmission from the tipper to the tippee.

2.3 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 one or two firms, and 4% of the cases involve 10 or more firms.

Table 3 summarizes our data at the trade level. We identify 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

Table 3
Trade characteristics: descriptive statistics

	Number of trades	Percentage of trades
A. Distribution of trading instruments		
Stocks	3,392	67.06
Options	1,610	31.83
ADS	44	0.87
Bonds	12	0.33
Total	5,058	100
B. Distribution of buys and sells		
Buys	4,220	83.43
Sells	838	16.57

C. Trading statistics

Characteristic	Mean	Median	SD	min	Max
Days from receiving a tip to an informed trade	8.05	2	23.88	0	417
Days from a trade to information disclosure	24.77	7	61.59	0	998
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)	1,013.6	90.00	7,926.8	4.0	27,500

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.

via stocks (67.06%) and options (31.83%).[13] 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. To provide perspective, Table B1 of the Online Appendix shows the mean and median number of all stock trades in the U.S. stock market during our sample period. 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 3 shows that the median time between the arrival and the use of information by insiders is 2 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 trade is 8 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 or higher). Hence, an economically meaningful fraction of them are capable of relatively sophisticated trading (e.g., using stock derivatives). Table B2 in the Online Appendix provides detailed statistics of the different job titles. 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.

2.4 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. We find 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

[13] Within the option sample, in-the-money options constitute approximately15% and 34% for calls and puts, respectively. Similarly, at-the-money options make up 16% and 19% of the sample, whereas out-of-the-money options make up 69% and 47% of the sample.

Table 4
Information content of trades

	SEC insider trading cases			
	Positive	Negative	Aggregate	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

This table 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. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

volume is traded by informed traders. Panel A of Table 11 provides more details on this distribution.

Apart from the volume of informed trading, 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 the day 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. Table 4 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 7 days. Arguably, one could treat these numbers as a lower bound of the true signal strength, because 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.[14] 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.

3. Full-Sample Results

In this section, we present our baseline results. We first describe the construction of our empirical counterparts to the components in Equation (1) for stock- and

14 We thank Alon Brav for providing us the data.

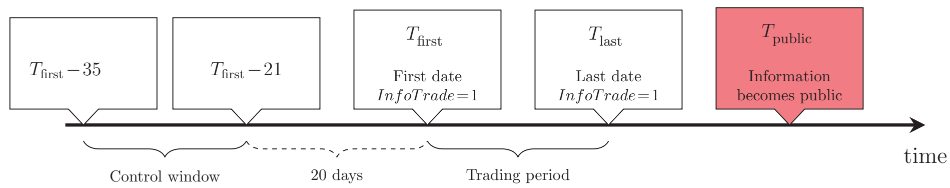


Figure 1
Event time line

option-based AIPs. We then describe evidence on Δ using both a univariate time-series approach and a multivariate cross-sectional regression specification.

3.1 Test design and univariate time-series evidence

Motivated by Equation (1), we seek to compare the value of a given AIPs on days with informed trading, $\mathbb{E}[AIP|I=1]$, with its unconditional expected value, $\mathbb{E}[AIP]$. 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. $T_{i,\text{first}}$ ($T_{i,\text{last}}$) corresponds to the first (last) date on which an insider trades a given instrument. 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 $\mathbb{E}[AIP|I=1]$ by conditioning the average value of AIP on $InfoTrade_{it} = 1$.

To define a set of *normal* dates for which $InfoTrade_{it} = 0$ so that to estimate $\mathbb{E}[AIP]$, we focus on a narrow window of 15 trading days close to T_{first}, which insulates us from any longer-term trends driving the data. In 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 $\mathbb{E}[AIP|InfoTrade=0]$ differs from the unconditional mean of AIP, as it would be the case if, for example, unidentified informed trades for the same stock occurred very close to T_{first}.[15] Last, we eliminate all cases in which informed trades occur less than 4 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 AIPs and the potential presence of pre-trends.

---

15 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.

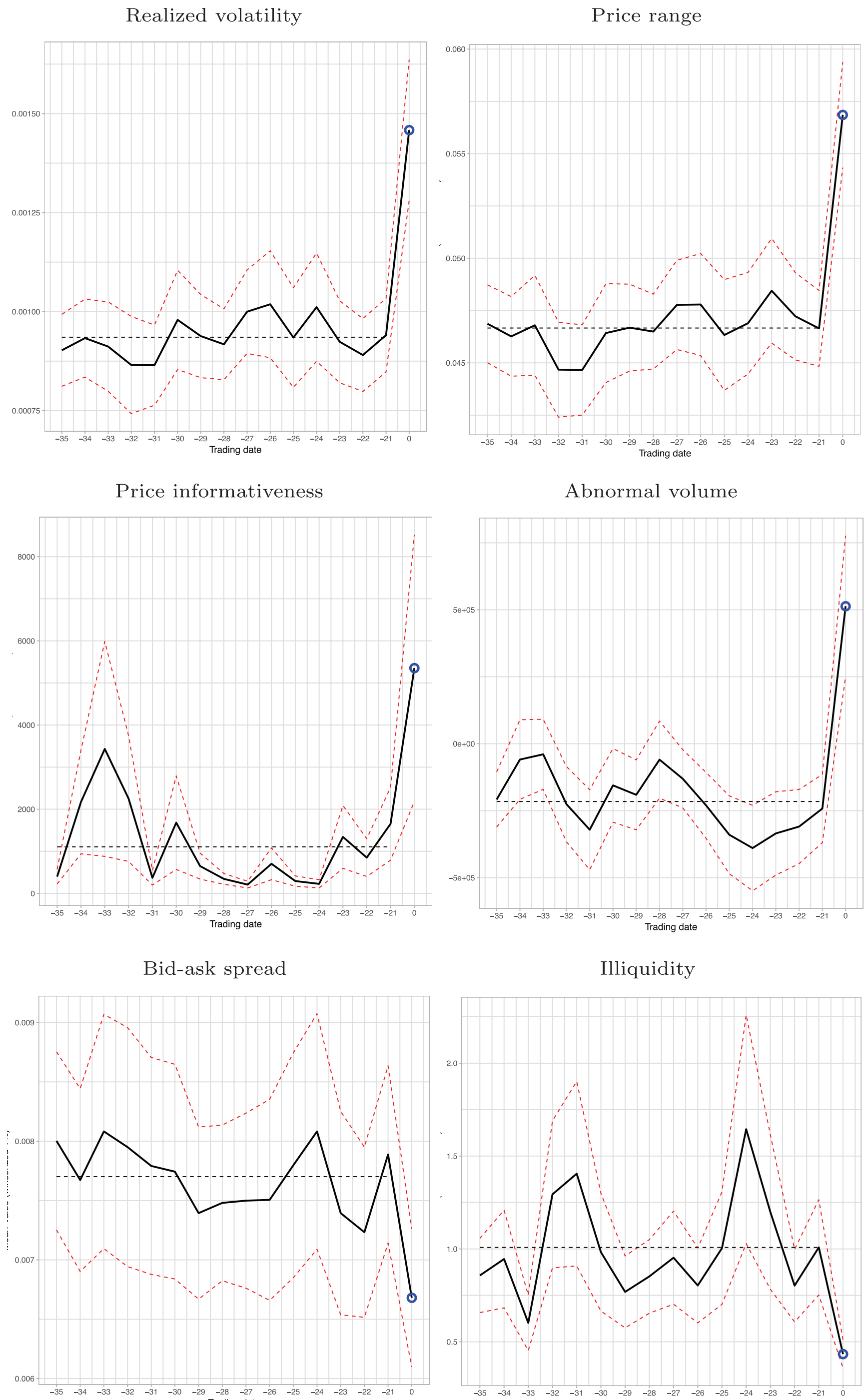


Figure 2
Stock-based AIPs: Daily mean values around informed trading
The figure presents the average values (aggregated across all trades) of stock-based AIPs, along with their two-standard-error bounds (red-dotted line). Dates [−35, −21] correspond to the control window period [$T_{\text{first}}-35$, $T_{\text{first}}-21$], as described in Section 3. 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].

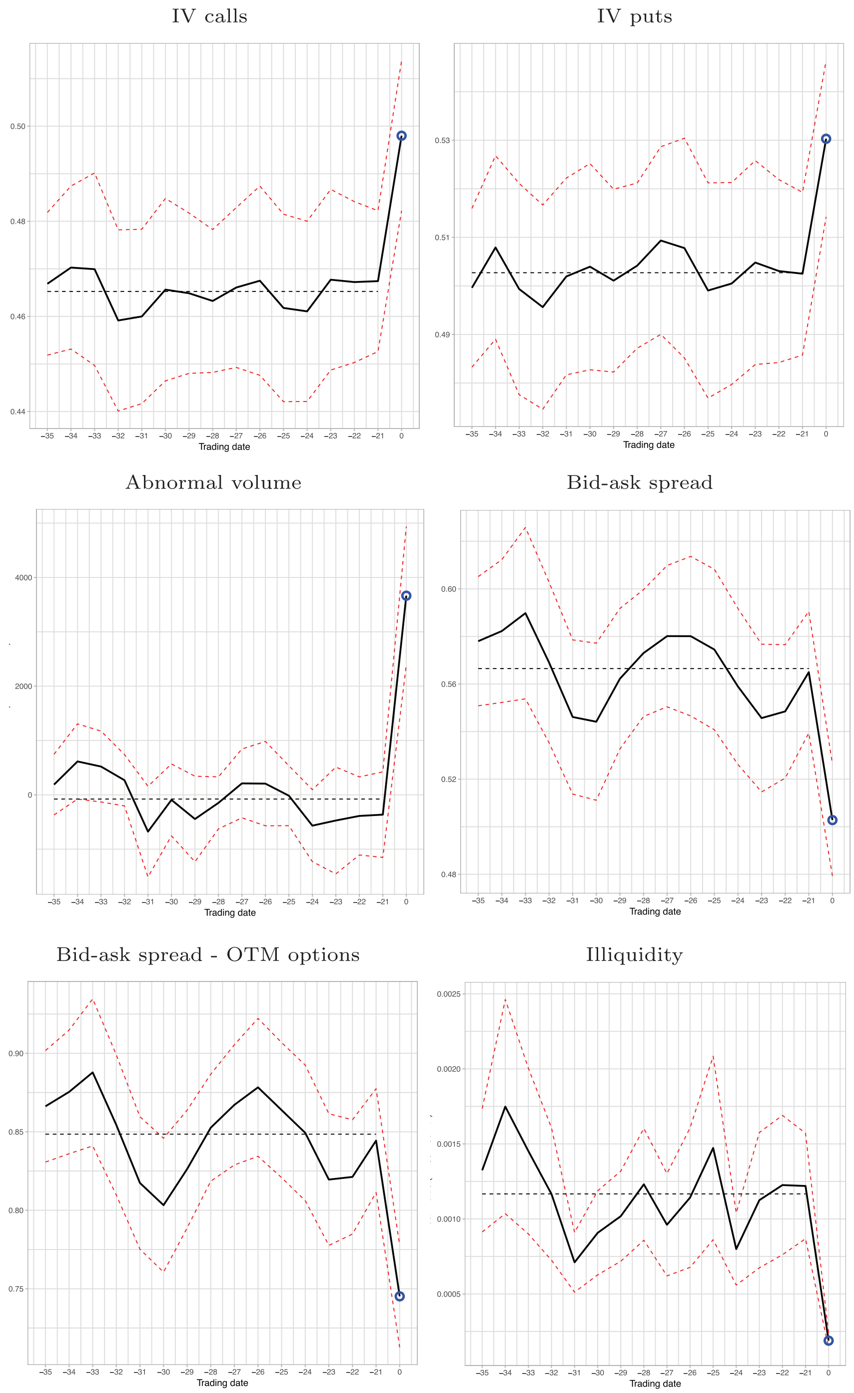


Figure 3
Option-based AIPs: Daily mean values around informed trading
The figure presents the average values (aggregated across all trades) of option-based AIPs, along with their two-standard-error bounds (red-dotted line). Dates [-35, -21] correspond to the control window period [T_{first}-35, T_{first}-21], as described in Section 3. 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].

In Figures 2 and 3, we display the average values of several stock- and option-based AIPs 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 AIPs 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 AIPs's response: volatility and abnormal volume measures increase, whereas, perhaps surprisingly, illiquidity measures decrease. Second, we do not observe any clear abnormalities 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 AIPs value in the control window can approximate the corresponding unconditional mean. We want to stress that the behavior of AIPs 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 AIPs, 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.

3.2 Results from a regression design

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

$$AIP_{it} = c \times Controls_{\tau_i} + d_i + e_t + \Delta \times InfoTrade_{it} + \mathrm{e}_{it}, \tag{2}$$

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 predetermined 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 $\tau_i := T_{i,\text{first}} - 35$.[16] 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 trends differentially affect a given AIPs on days with $InfoTrade$=1 and days in the control window. Nonetheless, to account for the possibility that

[16] Since the informed trades are 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, in unreported results, we obtain essentially the same Δ estimates using a specification with no controls.

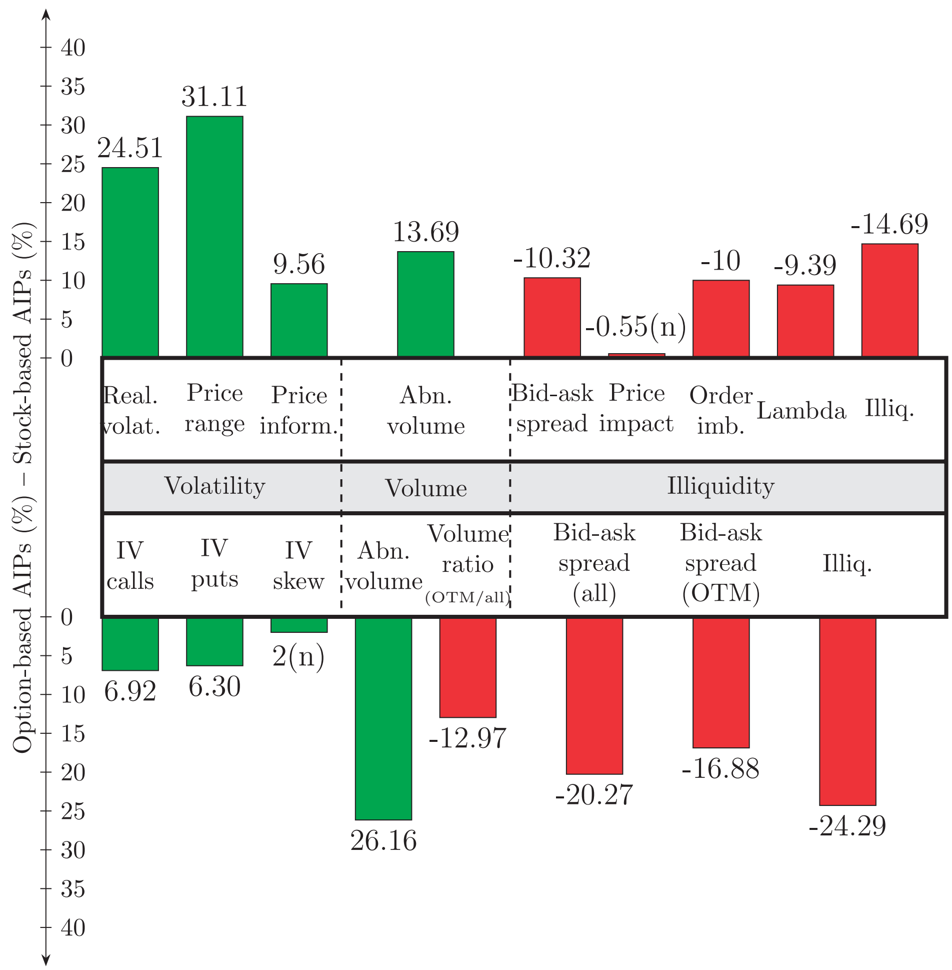


Figure 4
Main empirical design: Results summary
For each information AIPs, the numbers displayed in the bars correspond to the percentage ratio between the estimated coefficient on days with precisely identified informed trading and the corresponding full-sample standard deviation. Green (red) columns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

AIPs vary generically over time, we include time fixed effects e. To avoid effects from extreme outliers, we winsorize all AIPs at 1%. We cluster standard errors at the firm level to account for serial correlation in residuals.

Table 5 presents the results from estimating the regression model (2) for stock-based AIPs. To provide some perspective on economic magnitudes, given the different units in which AIPs are measured, the top panel of Figure 4 shows the values of estimated coefficients relative to the corresponding AIPs sample standard deviation as a percentage. First, we confirm that AIPs 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. *Realized variance* and *Price range* increase by 24.51% and 31.11%, respectively, and *Price Informativeness* also displays abnormally higher values. Third,

Table 5
Stock-based AIPs: Estimation results

	Volatility			Volume	Illiquidity				
Based on	Realized volatility	Price range	Price inform.	Abn. S. volume	Bid-ask spread	Price impact	Order imb.	Lambda	S. illiq.
InfoTrade	0.698***	1.143***	794.961**	401.611***	−0.100***	−0.111	−0.015***	−0.033***	−0.418***
	(0.118)	(0.131)	(393.525)	(109.563)	(0.015)	(0.473)	(0.005)	(0.007)	(0.071)
ln(SIZE)	−0.994**	−1.836***	−922.171*	1,090.575	−0.272***	−4.899***	−0.072***	−0.071***	0.097
	(0.477)	(0.445)	(491.921)	(851.666)	(0.052)	(1.187)	(0.013)	(0.017)	(0.071)
ln(VOL)	0.563*	0.857***	102.267	−544.517	−0.281***	−1.563	−0.005	−0.061***	−0.352***
	(0.335)	(0.324)	(554.952)	(424.200)	(0.059)	(1.600)	(0.013)	(0.013)	(0.106)
TURNOVER	8.138	21.052	−27,764.956	66,719.204**	29.338***	221.508**	1.216	5.379***	26.484***
	(16.908)	(23.174)	(40,708.387)	(31,074.977)	(3.467)	(92.927)	(0.755)	(0.812)	(6.147)
PRC	0.114***	0.104***	20.204	−78.222*	−0.007**	−0.007	−0.001	0.001	−0.004
	(0.042)	(0.028)	(43.457)	(43.460)	(0.003)	(0.067)	(0.001)	(0.001)	(0.004)
Constant	0.380	0.243	−308.887	11.063	−0.305***	−2.405***	−0.046***	−0.042***	0.066
	(0.247)	(0.156)	(350.466)	(158.765)	(0.024)	(0.581)	(0.006)	(0.005)	(0.045)
#obs	10,319	10,358	10,249	10,318	10,241	10,191	10,191	10,176	10,283

The dependent variables are daily stock-based AIPs at the firm level over the period 1995–2015. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Abnormal s. volume is 13.69% higher and, consistent with the univariate time-series observation, all illiquidity measures display lower values. For example, *Bid-ask 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 5-minute price impact measure.

Table 6 presents the results for option-based AIPs. The bottom panel of Figure 4 displays changes in AIPs values relative to their respective standard deviations. As in the case of stocks, most option-based AIPs 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.[17] 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 *Bid-ask spread* (all) and *O. illiq* suggest that illiquidity is lower, on average, when informed traders trade.

We have conducted a number of additional robustness tests that largely corroborate our results. Here, we discuss two of them. First, we control for the possibility that $InfoTrade=1$ could be correlated with other salient news, which would then explain the patterns in our data. While theoretically possible, practically, it seems less likely as it would go against the mere notion of private information. Nonetheless, we try one such confounding predictor, the forecast revisions by analysts, and estimate our baseline model conditioning on the indicator variable *Revision* equal to 1 on days when analyst forecasts change. The results, presented in Table C3 of the Online Appendix, indicate that our effects are unlikely to be explained by such spurious correlation. Second, it is possible that $InfoTrade=1$ does not uniquely define informed trading dates and other dates surrounding our event date are equally important. To address such autocorrelation, which could be either due to informed trading or due to investors' betting on scheduled corporate events, we intentionally skip 20 trading dates before the first instance of insider trading. We find no abnormal behavior in AIPs in such preperiod. As a research inquiry, we also allow for counterfactual informed trading dates by shifting $InfoTrade=1$ by 1 to 4 trading dates. Table C4 of the Online Appendix, which presents the results, shows that the conditional response of the AIPs visibly weaken as we move farther away from the $InfoTrade=1$ date, suggesting a lower likelihood of

[17] Our findings on volatility are based on the rich cross-section of option contracts with different maturities and moneyness. We aggregate these results by weighing observations with respect to the corresponding open interest. To establish the robustness of our results, we have also considered alternative specifications in which the 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.

Table 6
Option-based AIPs: Estimation results

	Volatility			Volume		Illiquidity		
Based on	IV calls	IV puts	IV skew	Abn. o. volume	Vol. ratio (OTM/all)	Bid-ask spr. (all)	Bid-ask spr. (OTM)	O. illiq.
InfoTrade	0.021***	0.016**	0.005*	3,298.912***	−4.089*	−0.065***	−0.074***	−0.130***
	(0.006)	(0.007)	(0.003)	(613.481)	(2.137)	(0.015)	(0.021)	(0.026)
ln(SIZE)	0.137	−0.132*	0.012	3,667.738	−51.804	−0.337	−0.700**	0.100
	(0.121)	(0.077)	(0.032)	(3,759.968)	(34.754)	(0.235)	(0.306)	(0.073)
ln(VOL)	−0.178***	−0.106***	−0.088**	22,365.206***	42.639***	0.196**	0.338***	0.103***
	(0.054)	(0.036)	(0.036)	(7,974.873)	(13.432)	(0.082)	(0.108)	(0.032)
TURNOVER	8.004*	7.217**	7.251**	−1679102.965***	−6,064.726***	−4.435	−10.343	−10.590***
	(4.504)	(3.007)	(2.976)	(650,906.813)	(1,217.406)	(5.862)	(8.034)	(2.871)
PRC	−0.000	0.005***	−0.000	−65.908	0.564	0.005	0.012**	−0.003*
	(0.002)	(0.001)	(0.001)	(78.940)	(0.666)	(0.004)	(0.006)	(0.002)
Constant	−0.049	0.102***	0.032	−13,991.952***	17.589	0.127	0.291*	−0.070*
	(0.060)	(0.038)	(0.032)	(5,242.858)	(18.088)	(0.122)	(0.158)	(0.041)
#obs	2,583	2,571	2,332	2,588	2,588	2,533	2,533	2,486

The dependent variables are daily option-based AIPs at the firm level over the period 1995–2015. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

informed trading for dates adjacent to those for which their presence is precisely identified.[18]

Overall, our results indicate that AIPs display noticeably 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 AIPs types: volatility and volume AIPs display higher values while illiquidity AIPs display lower values.

3.3 Sensitivity to trading intensity

The evidence in Section 3.2 suggests that AIPs 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 AIPs. 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. Next, we estimate the regression model (2) for all AIPs conditional on low- and high-intensity trades. Table 7 displays the results.

Even though the qualitative results for each subsample are not different from those of the unconditional sample, the associated t-statistics are, generally, higher for high-intensity cases in both stock and option markets. In fact, some AIPs 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 AIPs value. We further test the statistical differences across both subsamples by interacting *InfoTrade* with *Intensity*—an indicator variable equal one for high-intensity trades. The untabulated results indicate that high-intensity trades are statistically significantly different from low-intensity trades, especially for measures of illiquidity.

4. Evaluating Potential Selection Bias

One of the main empirical challenges in evaluating the impact of informed trading on AIPs is the unobservability of investors' information. The estimation of Δ in Equation (1) using trades from SEC investigations has a distinct advantage as regulators can document the use of advance knowledge regarding a corporate event, thus offering a unique opportunity of accessing individual traders' information sets. Therefore, one can (1) eliminate uncertainty about whether traders had private information about firm fundamentals, (2) provide

[18] We note that shifting the treatment date forward would be more problematic for two reasons. First, many informed trades occur on the day or one day before public events. Second, it is difficult to rule out the possibility that other traders learn about the private signal from the trades of insiders.

Table 7
Conditioning on trade intensity

	Volatility			Volume	Illiquidity				
Based on	Realized volatility	Price range	Price inform.	Abn. s. volume	Bid-ask spread	Price impact	Order imb.	Lambda	S. illiq.
A. Stock-based AIPs: 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
B. Stock-based AIPs: 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)	Bid-ask spr. (all)	Bid-ask spr. (OTM)	O. illiq.
C. Option-based AIPs: 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
D. Option-based AIPs: 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

The dependent variables are AIPs over the period 1995–2015. This table presents results for low- and high-intensity trades, as defined in Section 3. Panels A and B report the results for stock-based AIPs and Panels C and D for option-based AIPs. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

precise informed trading dates, (3) identify in which market the trade took place, and (4) connect each trade with the specifics of the information set, 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, instead, on assumptions whether a group of traders possess private information or the idea that informed trading is more likely around specific corporate announcements. Under such circumstances, one can misclassify pairs of asset–date (i,t) for which informed traders are present (spurious case in Figure 5).

Despite the aforementioned attractive characteristics, 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. Given that insider traders do not publicly announce their trades, from the SEC litigation files, for a given company and period, one computes $\mathbb{E}[AIP|I=1,\mathbb{I}_{\text{det}}=1]$, where $\mathbb{I}_{\text{det}} \in \{0,1\}$ denotes the detection outcome. By the definition of conditional

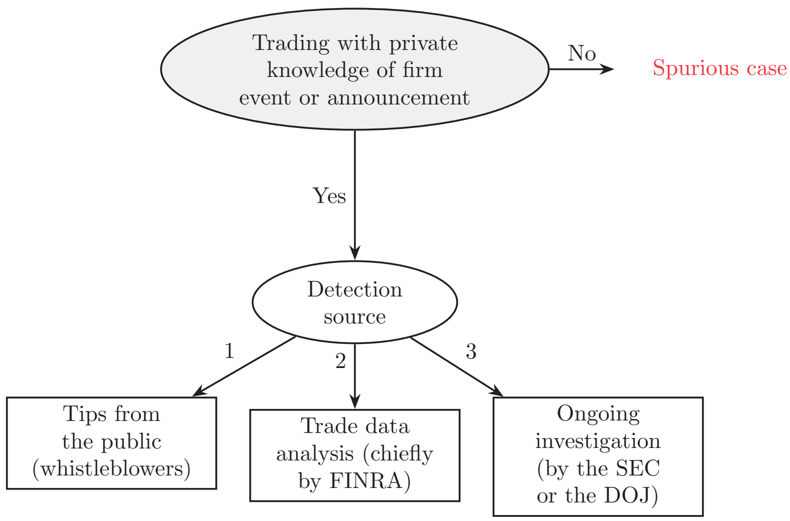


Figure 5
Institutional setting and potential selection sources

expectation, these can be expressed as $\frac{\mathbb{E}[AIP \times \mathbb{I}_{\text{det}}|I=1]}{\mathbb{E}(\mathbb{I}_{\text{det}}|I=1)}$. One does not face any identification challenge if the AIPs and detection are unrelated, that is, $\mathbb{E}[AIP \times \mathbb{I}_{\text{det}}|I=1] = \mathbb{E}[AIP|I=1]\mathbb{E}[\mathbb{I}_{\text{det}}|I=1]$. Otherwise, the detection process could bias estimates of Δ. An extreme case would be if the regulator only detected insider traders based on the public behavior of the AIPs considered here. We regard such extreme detection scenario as unlikely.[19] However, to better understand the plausibility of a detection bias, one must consider how insider traders get detected. As Figure 5 illustrates, the agencies in charge of insider trading prosecution, the SEC and the DOJ, usually detect insider trading activity from three major sources: whistleblower tips from the public, trade data analyses, or as part of an ongoing investigation.[20] We briefly describe each of the primary sources in Section 4.1 and, based on their specific institutional characteristics, we design and conduct empirical tests that evaluate the potential incidence of selection bias.

[19] For example, the results in Section 3 show that many stock-based AIPs 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 regulator 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.

[20] Additional sources of detection exist. Based on a Freedom of Information Act request to the SEC, we learned that during our sample period, 1.41% of the investigations originated in "issuer disclosure" and 1.5% from "other federal or local agencies." The SEC did not provide additional details about either category. In addition, insider trading investigations can be triggered by information contained in media articles.

4.1 Detecting insider trades

4.1.1 External tips. External tips include many situations in which an individual, such as an angry spouse or a business partner, contacts the SEC and blows the whistle on a violation of insider trading regulations. Given direct knowledge of traders' actions, this source of detection is, arguably, the least likely to be correlated with movements in the AIPs. Tipping from market participants, such as exchanges or brokers-dealers, instead, could be related to some dimensions of the trading process but is generally reported to FINRA.[21] Although the SEC does not provide formal statistics on the origination source of insider trading investigations, prior evidence suggests that a significant fraction of investigations originate 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).

4.1.2 Trade data analyses. Regulating agencies can monitor market activity and look for violations of securities laws and regulations. During our sample period, the SEC did not have a formal analytical framework to continuously monitor trades and the value of AIPs of the type we consider. Trade data analysis in equity markets is conducted by SROs and chiefly by FINRA.[22] Within FINRA, the Office of Fraud Detection and Market Intelligence (OFDMI) looks for suspicious trade activity using a monitoring software called Securities Observation News Analysis and Regulation, SONAR. When a major announcement comes out, FINRA's personnel can check whether SONAR picked up any unusual movements for the companies involved. The OFDMI reports to have sent 303 insider trading referrals to the SEC in 2014.[23] During our sample period, insider trading in option markets was surveilled by the Options Regulatory Surveillance Authority (ORSA).

4.1.3 Existing investigation. Insider trade detection could also originate as a consequence of an ongoing or past investigation. For example, a whistleblower could trigger an investigation regarding insider trading in firm A through which,

---

[21] For example, an option dealer, taking on the opposite side of the trade, may contact FINRA after experiencing losses to a potential insider trader just before a merger. Even when such tips may be affected by trades, they may rely on access to traders' identities as well as traders' profits and losses, a source of information that is nonpublic and not necessarily reflected in AIPs. At least since 2011, whistleblowers have an economic incentive to report to the SEC to receive a monetary reward (see Section 4.2). FINRA also has a whistleblower hotline. However, FINRA only has the authority over broker-dealers and their employees, so some of these calls are likely to be referred to the SEC.

[22] We were not able to find information about any such continuous monitoring program on the SEC Web site. The lack thereof is supported by interviews we conducted with SEC officials. We thank an anonymous referee for providing helpful background on the role of SROs and FINRA.

[23] See FINRA (2015).

subsequently, the SEC learns about insider trading in firms B and C. Thus, insider trades in B and C do not originate in routine trade monitoring but may involve access to specific brokerage accounts, phone conversations, wiretaps, and/or text messages, like in the notorious investigation of Raj Rajaratnam and the Galleon Fund. Moreover, the initial trigger of the investigation need not regard insider trading but other types of crimes, for example, those investigated by the Federal Bureau of Investigations and later referred to the SEC. As mentioned above, unlike FINRA, the SEC did not have continuous access to trade data during our sample period. An additional possibility, however, is for the SEC to detect suspicious links among repeat offenders, individuals, and business entities by connecting pieces of information using "bluesheet" data that the agency gathered in various investigations. The SEC's ability to initiate investigations using this trader-connections approach is limited because it collects such information on an ad hoc basis.[24]

4.2 Whistleblowers: Evidence from the SEC WRP

Regulators have traditionally relied on information provided by whistleblowers to strengthen the enforcement of illegal insider trading laws. Because whistleblowers have some personal or business connection with the involved traders, it is less likely that the detection of these traders is based on the behavior of market-level AIPs such as liquidity levels. In this section, we exploit a change in the regulatory environment affecting insider trading whistleblowers. 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 (1) derived from the independent knowledge or analysis of a whistleblower, (2) not known to the SEC from any other sources, and (3) 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 actions and, thus, such cases are significantly less prone to detection selection concerns based on AIPs values. Hence, if selection bias drives our results, we would expect AIPs 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 102 different cases, of which 55 were investigated through the program and 47 have no precise source of investigation. Table D5 of the Online Appendix summarizes various trading characteristics for the two types of cases, which are fairly similar along most

---

[24] In 2010, the SEC created the Market Abuse Unit as a platform able to better study information flows in financial markets and in 2012 adopted Rule 613 of Regulation NMS, which requires the SROs to collaborate in the creation and maintenance of a consolidated audit trail (CAT). A CAT would potentially enhance the ability of the SEC to have a more proactive role in the detection of insider trading. Multiple delays, however, prevented this system from being implemented by the end of our sample period.

Table 8
SEC Whistleblower reward program cases: Estimation results

A. Stock-based AIPs

	Volatility			Volume	Illiquidity				
Based on	Realized volatility	Price range	Price inform.	Abn. s. volume	Bid-ask 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

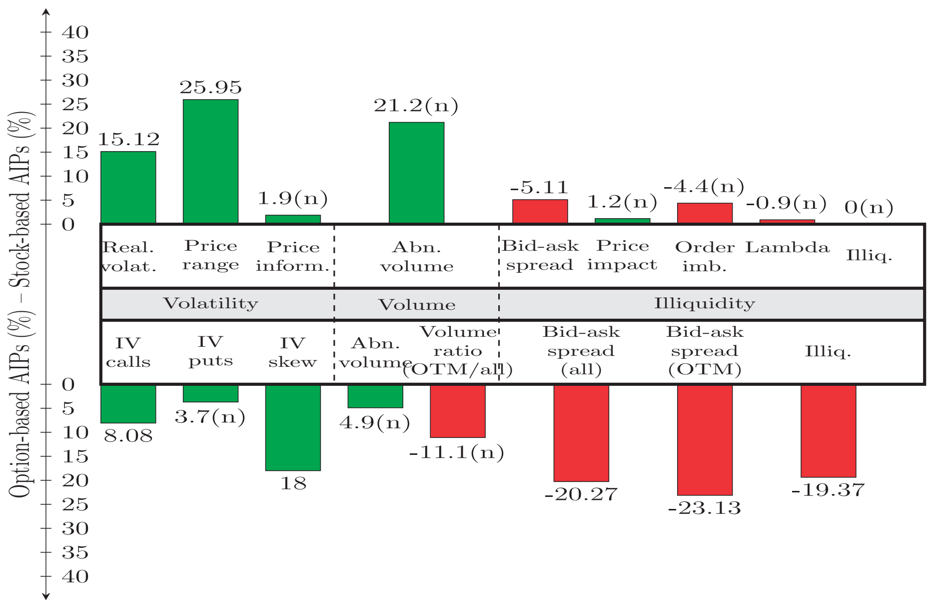
B. Option-based AIPs

Based on	Volatility			Volume		Illiquidity		
	IV calls	IV puts	IV skew	Abn. o. volume	Vol. ratio (OTM/all)	Bid-ask spr. (all)	Bid-ask 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

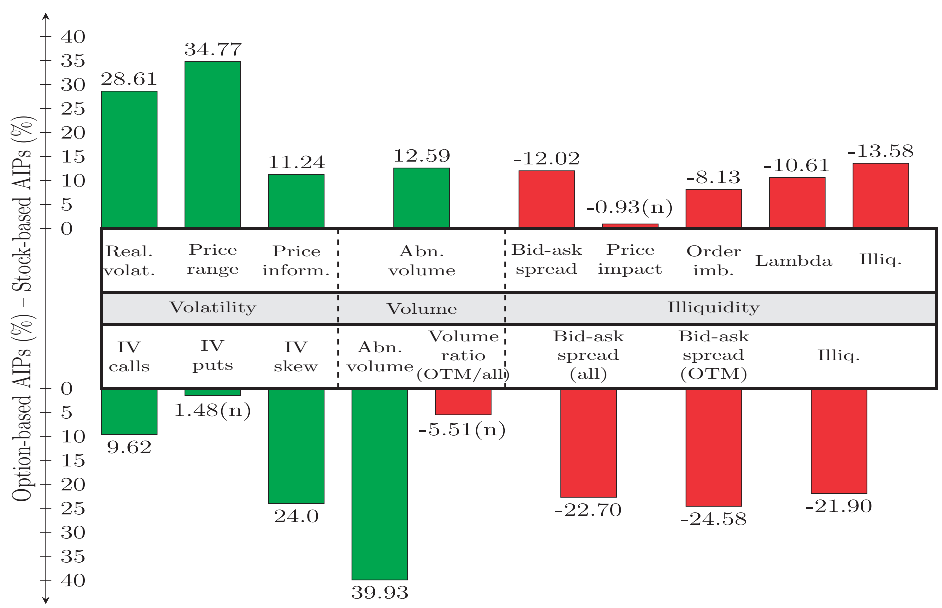
This table presents results for the subsample of SEC WRP cases. The dependent variables are AIPs. Panel A reports results for stock-based AIPs, and Panel B the results for option-based AIPs. The dependent variables are daily stock-based AIPs at the firm level over the period 1995–2015. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

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 and greater profits per trade.

Table 8 reports the results from estimating the regression model in Equation (2) for the WRP sample and panel A of Figure 6 displays the value of the *InfoTrade* coefficient relative to the AIPs's corresponding standard deviation. The patterns for volatility, volume, and illiquidity essentially remain 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 impact on *IV* is stronger for call options, which further increases *IV skewness*. Regarding illiquidity, all measures in option markets, bid-ask spreads and *O. illiq*, remain negative and statistically significant. *Bid-ask spread* is also negative and statistically significant in stock markets. *Lambda* and *Order imbalance* are negative but statistically insignificant. The illiquidity patterns are thus relatively stronger in option-based AIPs. *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*-statistic is smaller. For options, the coefficient is positive, but its magnitude is smaller.



WRP sample



No-FINRA sample

Figure 6
Results summary: Conditioning on the origin of the investigation
For each AIPs, the numbers on the bars correspond to the percentage ratio between the estimated Δ coefficient on the corresponding full-sample standard deviation. Green (red) columns correspond to positive (negative) effects on the AIPs value. Estimates that are not statistically significant are indicated by "(n)."

4.3 Trade data monitoring: Evidence from SRO involvement

Based on the description of the continuous trade analysis process above, we hypothesize that the likelihood of nonzero correlation between AIPs and the detection of insider trading is greater for those investigations in which FINRA and/or ORSA were involved. On the other hand, if these SROs were not involved, it is more likely that the investigation source is different than trade analysis, for example, information can be obtained through a whistleblower or a referral from a Federal agency. Under this hypothesis, and if a selection bias related to such correlation is responsible for our findings, one should observe different AIPs empirical patterns when SROs are involved in the investigation. In particular, if no SRO assistance is found, we would expect that the effects are unlikely biased due to sample selection.

To test this hypothesis, we search for evidence of FINRA and ORSA involvement in all available SEC online press releases and litigation files related to insider trading.[25] For the period 2004–2015 during which the agencies were active, we find that, of 278 investigations, the SEC acknowledges SRO assistance in 117 cases. We use this information to partition our sample to cases with SRO involvement, either FINRA or ORSA (the “FINRA" subsample), and the rest (the “no-FINRA" subsample). Table D5 of the Online Appendix displays summary statistics for each subsample. The subsamples are similar regarding strategic dimensions, such as trades per trader, trades per firm and timing. The no-FINRA investigations involve, on average, larger firms and are associated with larger traders’ profits.

Table 9 presents the baseline test regression results for the FINRA/no-FINRA subsamples. From the perspective of any potential biases, the sample of interest is one for cases without FINRA involvement (panels B and D), because it is for that subsample that we would expect the magnitudes to diminish if the results were purely driven by the detection ability of various investigators. Both for stock- and for option-based AIPs, the estimated *InfoTrade* coefficients yield the same qualitative patterns as those from the full sample analysis of Section 3. Further, panel B of Figure 6 shows the estimated coefficients relative to the corresponding AIPs sample standard deviation as a percentage. From a quantitative perspective, the conditional value increase in volatility and volume, and the decrease in value of illiquidity are also of similar magnitudes.

4.4 Ongoing investigations: Evidence from the investigation size

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 based on abnormal market AIPs. Intuitively, for a generic case with, say, 10 firms, it is unlikely that detection in *each* firm was based on independent publicly observed AIPs movements. Rather, even when the investigation originated from screening one firm’s AIPs values, it is likely that trades in the remaining firms were uncovered as part of subsequent research into the trades of defendants and their social network. Therefore, if detection bias is at work, we should expect

[25] Ahern (2018) considers a similar approach

Table 9
Conditioning on FINRA

	Volatility			Volume	Illiquidity				
Based on	Realized volatility	Price range	Price inform.	Abn. s. volume	Bid-ask spread	Price impact	Order imb.	Lambda	S. illiq.
A. Stock-based AIPs: FINRA									
InfoTrade	0.186	0.465**	227.615	439.712**	−0.062**	−0.098	−0.019*	−0.028*	−0.522***
	(0.198)	(0.210)	(0.000)	(222.728)	(0.029)	(0.000)	(0.010)	(0.016)	(0.121)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	2,647	2,647	2,635	2,646	2,642	2,620	2,620	2,616	2,624
B. Stock-based AIPs: No FINRA									
InfoTrade	0.844***	1.340***	971.590**	380.065***	−0.113***	−0.161	−0.013**	−0.035***	−0.391***
	(0.139)	(0.156)	(488.147)	(123.234)	(0.018)	(0.533)	(0.006)	(0.008)	(0.085)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	7,671	7,710	7,613	7,671	7,598	7,570	7,570	7,559	7,658

	Volatility			Volume		Illiquidity		
	IV calls	IV puts	IV skew	Abn. o. volume	Vol. ratio (OTM/all)	Bid-ask spr. (all)	Bid-ask spr. (OTM)	O. illiq.
C. Option-based AIPs: FINRA								
InfoTrade	0.016	0.029**	−0.003	2,518.678***	−6.245*	−0.051**	−0.032	−0.121***
	(0.012)	(0.013)	(0.004)	(620.541)	(3.334)	(0.023)	(0.033)	(0.042)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	999	996	884	999	999	945	945	941
D. Option-based AIPs: No FINRA								
InfoTrade	0.025***	0.004	0.012***	3,839.171***	−2.369	−0.084***	−0.118***	−0.138***
	(0.006)	(0.007)	(0.004)	(997.598)	(2.832)	(0.019)	(0.026)	(0.034)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	1,584	1,575	1,448	1,589	1,589	1,588	1,588	1,545

This table presents results for firms with small and large stock market capitalization. The sample is the SEC WRP cases. The dependent variables are AIPs. Panels A and B report the results for stock-based AIPs, and Panels C and D the results for option-based AIPs. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 4 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

the correlation with AIPs to be stronger for single-company cases. Based on this idea, we classify the SEC investigations according to the number of firms involved, n. About 80% of investigations involve only one or two companies (n_{low}). The remaining 20% involve a greater number of firms, n_{high}, with up to $n=25$ in the sample. Because of the size disparities, each subsample contains a similar number of trades. We then conduct the baseline test in Equation (2) for each subsample. Table 10 reports the results. We can see that both for stock- and option-based AIPs, the *InfoTrade* coefficients display the same patterns found in the full sample. In particular, volatility measures increase in value and illiquidity measures decrease in value. For stocks, abnormal volume is higher in the case of small investigations, as one could expect based on the detection hypothesis. However, this is not the case for stocks, the economic and statistical significance for the coefficient of abnormal volume is larger in the case of investigations involving many firms.

Table 10
Conditioning on the investigation size

Based on	Volatility: Realized volatility	Volatility: Price range	Volatility: Price inform.	Volume: Abn. s. volume	Illiquidity: Bid-ask spread	Illiquidity: Price impact	Illiquidity: Order imb.	Illiquidity: Lambda	Illiquidity: S. illiq.
A. Stock-based AIPs: Small Investigations									
InfoTrade	0.647**	1.075***	834.568	203.158	− 0.085*	0.141	− 0.011	− 0.007	− 0.593***
	(0.256)	(0.305)	(660.962)	(199.599)	(0.048)	(0.706)	(0.011)	(0.015)	(0.192)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	5,170	5,199	5,125	5,161	5,129	5,090	5,090	5,076	5,139
B. Stock-based AIPs: Large investigations									
InfoTrade	0.747***	1.227***	764.512	606.505**	− 0.104***	− 0.386	− 0.018**	− 0.061***	− 0.229**
	(0.200)	(0.249)	(625.546)	(265.330)	(0.028)	(0.869)	(0.008)	(0.015)	(0.100)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	5,149	5,159	5,124	5,157	5,112	5,101	5,101	5,100	5,144

	Volatility: IV calls	Volatility: IV puts	Volatility: IV skew	Volume: Abn. o. volume	Volume: Vol. ratio (OTM/all)	Illiquidity: Bid-ask spr. (all)	Illiquidity: Bid-ask spr. (OTM)	Illiquidity: O. illiq.
C. Option-based AIPs: Small investigations								
InfoTrade	0.029*	0.020	0.005	4,922.315**	−4.576	−0.055	−0.058	−0.049***
	(0.015)	(0.014)	(0.013)	(2,178.236)	(2.933)	(0.034)	(0.036)	(0.016)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	1,332	1,324	1,235	1,334	1,334	1,314	1,314	1,298
D. Option-based AIPs: Large investigations								
InfoTrade	0.012	0.010	0.006	1,598.167	−3.655	−0.083*	−0.102	−0.212**
	(0.015)	(0.017)	(0.005)	(961.795)	(4.160)	(0.045)	(0.062)	(0.085)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#obs	1,251	1,247	1,097	1,254	1,254	1,219	1,219	1,188

This table presents results for firms with small and large stock market capitalization. The sample is the SEC WRP cases. The dependent variables are AIPs. Panels A and B report the results for stock-based AIPs, and Panels C and D the results for option-based AIPs. $InfoTrade$ is an indicator variable equal to 1 for asset-day pairs with informed trading. Section 4 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

5. Informed Traders' Strategic Timing and Illiquidity Measures

One of our key findings is the negative conditional response of illiquidity in both stock and option markets. This section explores mechanisms that could rationalize this counterintuitive connection. We focus first on the possibility that informed traders strategically time their trades. In particular, if trading costs are high due to temporary market illiquidity, an informed investor might want to time the execution of his trades to minimize such costs and, thus, one could observe low illiquidity levels when informed investors trade. This possibility, the *timing channel*, has been considered in the seminal work of Admati and Pfleiderer (1988) and more recently by Collin-Dufresne and For (2015, 2016) for the case of SEC 13D filers. The question remains whether such a pattern results from the strategic behavior of traders or is driven by other unobservable factors. At the outset, we evaluate whether uninformed volume is abnormally high on days when *InfoTrade* equals 1. Next, we take advantage of a unique feature of our data, the fact that we can observe the date when traders *receive*

Table 11
Informed, uninformed, and abnormal volume on *InfoTrade=1* days

Volume metric	Stocks				Options			
	Full	WRP	no-FINRA	Large inv.	Full	WRP	no-FINRA	Large inv.
A. Informed and S. uninformed shares of total volume, by sample								
Informed vol.	10.76	8.81	12.83	9.60	34.27	26.19	29.05	31.35
Uninformed vol. proxy	89.24	91.19	87.17	90.40	65.73	73.81	70.95	68.65
Total vol.	100	100	100	100	100	100	100	100
B: Informed volume and abnormal volume by sample								
	Full	WRP	no-FINRA	Large inv.	Full	WRP	no-FINRA	Large inv.
Informed vol.	215.97	343.93	256.14	255.13	790.91	1,073.07	912.77	676.23
Abnormal vol.	447.16	594.7	459.58	748.66	5,123.46	10,226	6,378.91	5,966.74
Informed/Abnormal (%)	48.30	57.83	55.73	34.08	15.44	10.49	14.31	11.33

This table reports summary statistics for various measures of trading volume on $InfoTrade=1$ days. *Informed volume* is given by the trades of insider traders. *Uninformed volume proxy* is given by total volume net of the trades by insider traders. *Abnormal volume* refers to trading volume in excess of predicted volume, as described in Section A of the Online Appendix. Full refers to the full sample results, WRP to the SEC Whistleblower Reward sample, no-FINRA to the sample without FINRA's involvement, Large inv. to the sample with a large number of firms being investigated. Section 4 describes these samples. Panel A presents the results on the share of informed/uninformed volume in total volume. Panel B shows the number of shares (for stocks) and contracts (for options) traded within each category. Table D6 of the Online Appendix gives the sample period and additional sample characteristics.

private information about firm fundamentals. Finally, we evaluate the possibility that informed traders use limit orders and study the response of illiquidity AIPs by firm size.

5.1 An assessment of informed and abnormal volume

To gain a better perspective on why trade volume is abnormally high on *InfoTrade*=1 days, one must evaluate the connection between the trades of the informed and those of other market participants. For that, we compute an additional measure of volume, *Uninformed Volume Proxy* (UVP), given by the difference between the total volume on *InfoTrade*=1 and the informed volume identified in the SEC insider trading investigations. UVP is an imperfect measure of uninformed volume, of course, because it is not feasible to access every trader's information set. To provide further detail on their relative importance, we compute the average share of informed volume and UVP, relative to total trading volume, on *InfoTrade*=1 days. Panel A of Table 11 presents the results. For stocks, informed volume and UVP are approximately 10.76% and 89.24% of the total volume; for options, 34.27% and 65.73%, respectively.

Next, we compare informed volume in both stock and option markets to *Abnormal volume* on *InfoTrade*=1 days. Panel B of Table 11 shows that even though for stocks a significant proportion of the abnormal volume can be attributed to informed trades, slightly more than one half can be explained by UVP. The fact that abnormal volume is far from being fully explained by informed trades is consistent with the timing channel: Insiders are more likely to trade when the uninformed volume is abnormally high. For options, informed traders represent a significantly lower proportion of abnormal volume, 15.4%. Arguably, if the timing mechanism operates, one would expect that the incentive to strategically time trades is stronger for single name stock options, which are significantly more illiquid than stocks. The fact that UVP explains a greater proportion of abnormal volume for options, therefore, is consistent with the hypothesis that informed traders strategically wait for days with a relatively greater option volume. We assess the statistical significance of the UVP response within the baseline regression specification. The results, presented in Table E6 of the Online Appendix, confirm that the increase in UVP on *InfoTrade*=1 is economically and statistically significant in both stock and option markets.

An alternative explanation to the timing channel is that informed traders are detected based on high market volume. If this were the case, a selection mechanism would sample high-UVP days more often. To evaluate this possibility, we analyze the robustness of the volume responses across the WRP, no-FINRA, and Large Investigation subsamples (see Section 4), which we argue are less likely subjected to detection bias. Panels A and B of Table 11 show that the economic magnitudes of key variables are remarkably consistent across samples. For stocks, for example, informed trading accounts for a proportion between 34.08% and 57.83% of *Abnormal Volume*, leading to similar conclusions about the increase in UVP on *InfoTrade*=1 days. For options, informed trades account for a share of *Abnormal volume* that ranges from 10.49% to 14.31%. The only noticeable difference with the full sample is that the proportion of informed trades is smaller in the WRP sample, as one would expect if the WRP allows for the identification of relatively more subtle trades. We note, however, that this relation is not observed for the share of informed stock volume in *Abnormal volume*, which is higher in the WRP sample than in the full sample. Overall, we observe consistent qualitative patterns which supports the view that strategic informed trade timing is a response to varying uninformed volume.

An additional alternative possible explanation of the relation between UVP and *Abnormal volume* is that the presence of informed traders causes uninformed traders to trade more. Although we cannot disregard this possibility entirely, arguably, the relatively low median fraction of informed trades and the fact that informed traders unequivocally and systematically trade in the same direction makes this alternative explanation less likely. However, because we do not observe the precise intraday times when informed trades are executed, it is difficult to precisely assert whether informed trades lead some uninformed trades. Therefore, to further explore the strategic timing channel, we next analyze a plausibly exogenous source of heterogeneity and short and long-lived private information.

Table 12
Illiquidity AIPs: Tests of strategic timing

	Stock				Option	
Market	Bid-ask spread	Order imbalance	Lambda	S. illiq.	Bid-ask spread (OTM)	O. illiq.
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
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
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		

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 AIPs. Panel A reports the results for short horizons and all assets, and Panels B and C reports 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. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

5.2 Trading horizon effects

The timing of informed trading is limited by the private information horizon. Based on this economic intuition, we partition our sample according to whether the trades in the investigation have 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 3 days. We argue that, traders do not select when they receive the private tip and that, within a short horizon, it would be relatively difficult for a trader to time his trades well.

Table 12 shows the estimated values of Δ for the illiquidity that display abnormal behavior in the full sample analysis of Section 3. Panels A to C display the results for all stocks, large caps, and small caps, respectively. We observe that the only AIPs 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 *Bid-ask spread* for a sample of OTM options written on large caps, and for stock-based *Bid-ask spread* and *Lambda* in the case of small caps, albeit the t-statistics are small. In most cases, the value of the Δ estimates for 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. The fact that some AIPs remain negative even for relatively short horizons suggests that traders still maintain some ability to time trades within a 3-day period and/or that additional factors, such as the use of limit orders, are at play.

6. Use of Limit Orders by Informed Traders

The effect on illiquidity measures can be affected by a second strategic dimension, the use of limit orders. That the use of limit orders by informed traders adds to market liquidity could help explain 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 order types, with 62 limit orders (73%) and 23 market orders.[26]

We exploit the subsample with identified order types to test the hypothesis that trades executed using limit orders are associated with higher same-day liquidity. Table 13 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 *Bid-ask 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 some perspective on their correlation with illiquidity, but establishing causality is more difficult. Unlike the test in Section 5.2, 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.

In a similar spirit to previous tests, we estimate the empirical model in Equation (2) for the subsample of firms with a market capitalization below and above the median value in the sample. The results, presented in panels D and E of Table 13, 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

---

26 Note that the quoted use of limit orders refers to nonmarketable limit orders by which the informed trader achieves execution on the passive side of the trade. Putting this proportion in perspective requires some out-of-sample group of traders to act as a benchmark. Kelley and Tetlock (2013) study retail investors' order placement using a proprietary sample from a large U.S. online brokerage firm. These authors report 178 million trade executions with market orders and 47 million with limit orders, resulting in less than 25% limit order use. The higher percentage in our sample is consistent with the fact that it contains both retail investors and professionals such as portfolio managers. Intuitively, the proportion of limit order use should be higher for the latter.

Table 13
Stock illiquidity AIPs: Tests on order types and market capitalization

	Bid-ask spread	Order imbalance	Lambda	S. illiq.
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
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
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
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
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

The dependent variables are AIPs. 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 1 for asset-day pairs with informed trading. Section 3 defines $InfoTrade$ and control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms. *p <.1; **p <.05; ***p <.01.

of *Bid-ask spread* and *Order Imbalance* for small stocks suggest 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.

We also assess the importance of market capitalization within the subset of whistleblower cases, separately for stocks and options. The disadvantage of using these subsamples is a relatively small sample size. Table E7 of the Online Appendix presents the results. Our results indicate that the illiquidity effects are more pronounced for a sample of small stocks, especially within option-based AIPs. However, we take this evidence as only supportive of our hypothesis because of weak statistical power.

Overall, the results in this section are consistent with the view that informed traders are likely to use limit orders and do so relatively more often for assets with high liquidity costs (that is, small caps). The strategic timing of informed orders analyzed in Section 5 and the limit order channels are, of course, not mutually exclusive. 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.

7. Market Makers: Learning and Event Uncertainty

To shed more light on the finding that the average bid-ask spread is lower on days with privately informed traders, in this section we focus on an additional potential explanation related to the role of market makers. In contrast to the canonical sequential trade model of Glosten and Milgrom (1985), our identification setting contemplates the possibility that no private information accrues on some trading days. Therefore, rational market makers are concerned not only with the direction of private information but also with its presence or absence, like in the seminal work of Easley and O'Hara (1992, EOH92). We first study the implications of such information structure for the average bid-ask spread on days with and without privately informed traders. Then, we provide evidence consistent with the learning mechanism.

7.1 Learning under event uncertainty and the conditional distribution of the bid-ask spread

Our empirical results show that, on average, the bid-ask spread is lower on days with *InfoTrade*=1. We conjecture that, when the market makers update beliefs on the asset value under event uncertainty, the order flow on *InfoTrade*=1 days could lead to a rapid resolution of uncertainty and generate narrow bid-ask spreads vis-a-vis days with no informed traders. To assess the plausibility of this conjecture, we simulate artificial market sessions with the same market participants and information structure used in EOH92. We calibrate the model's parameters to several market characteristics. We explore the sensitivity of the implications to the PIN value, a quantity which is not directly observable. We relate this pivotal parameter to the data by drawing on the distinction between large-cap and small-cap companies. Intuitively, small-cap stocks are considered those with a greater probability of informed trading, which allows us to evaluate the model's predictions by comparing the results for large and small caps.

Here, we briefly summarize the setting. Time is discrete, and there is a single asset with uncertain future value V with support $\{\underline{V}, \overline{V}\}$. An information event is the realization of a random signal about V: with probability $\alpha > 0$, an informed trader receives a binary signal $\Psi \in \{L, H\}$ about the value of V. If no signal arrives, the realization is $\psi = 0$. Let $\mathbb{E}[V|\psi = L] = \underline{V}$, $\mathbb{E}[V|\psi = H] = \overline{V}$, and $\mathbb{P}(\psi = L) = \delta > 0$. Being uninformed about the realization of ψ, a single risk-neutral market maker acts competitively and sets bid and ask quotes for one unit of the asset. Let μ be the proportion of informed trades, a quantity that drives the probability that an informed trader is randomly selected to trade in a given period. If selected, the informed trader submits an order with probability one. Liquidity traders are selected with probability $1-\mu$, and, if selected they submit a buy or a sell market order with equal probability p.[27] At time $t=1$, prior beliefs $\rho_{0,1}:=\mathbb{P}(\psi=0)=1-\alpha$, $\rho_{L,1}:=\mathbb{P}(\psi=L)=\alpha\delta$, and $\rho_{H,1}:=\mathbb{P}(\psi=H)=\alpha(1-\delta)$ permit the determination of bid and ask quotes given by $\mathbb{E}[V|S]$ and $\mathbb{E}[V|B]$, respectively, where the conditional expectations are consistent with Bases rule. In each period $t>1$, the market maker updates the quotes in response to the arriving order $Q\in\{B,S,N\}$, that is, a buy market order, a sell market order, or no order. For example, at a generic time t, the bid quote must equal $\mathbb{P}\left[\psi=L|Q^{t-1},S\right]\underline{V}+\mathbb{P}\left[\psi=H|Q^{t-1},S\right]\overline{V}+\mathbb{P}\left[\psi=0|Q^{t-1},S\right]\mathbb{E}[V]$, where $Q^{t-1}=\{Q_1,...,Q_{t-1}\}$. Over any given trading day, the average value of the bid-ask spread depends on two factors: the starting bid-ask spread level and the rate of convergence of beliefs to their full information values. For example, everything else being constant, when the probability of uninformed orders is high, the initial bid-ask spread is low but so is the speed of convergence given that the informed trader can hide his trades more effectively. Conversely, when the probability of an informed order is high, the initial bid-ask spread is high and so is the speed of belief convergence.

In this environment, Easley and O'Hara show that the strong law of large numbers implies that Bayesian posteriors exponentially converge to their full information values. For example, when no information event takes place, the lack of abnormal volume implies that $\rho_{0,t}=\mathbb{P}\left(\psi=0|Q^t\right)\to 1$ a.s. and the quotes converge to $\mathbb{E}[V]$. Thus, the bid-ask spread narrows over time. To better understand our empirical results, however, we must analyze the *finite sample* properties of the daily average of the bid-ask spread conditional on ψ, that is

$$\Delta_{\text{spread}}=\underbrace{\frac{1}{N}\left[\sum_{t=1:N}(\text{ask}_t-\text{bid}_t)|\psi\in\{L,H\}\right]}_{\text{days with InfoTrade=1}}-\underbrace{\frac{1}{N}\left[\sum_{t=1:N}(\text{ask}_t-\text{bid}_t)\right]}_{\text{days with InfoTrade=0}}. \quad (3)$$

where N represents the number of intraday time periods. Because the exact finite sample distribution of (3) is not available, we simulate market sessions representing 1 day and calibrate parameters as follows. A period represents 1 minute, so a trading session of 6.5 hours, like in the U.S. stock market, has 390 periods. Information received at the "opening" of day t is revealed at the "closing" of the same day. Thus, prior beliefs are the same at the beginning of each day. The asset has an expected value of $\mathbb{E}[V]=\$15$, akin to that in our sample, and $\underline{V}$ is calibrated using the empirical distribution of *Strength* for negative news (see Table 4), so that $\underline{V}=(1-.1846)\times\$15=\$12.21$. We set $\delta=0.5$ and thus $\overline{V}$ is determined by $\frac{\$15-\delta\underline{V}}{1-\delta}=\$17.22$. The probability p is calibrated so that the mean number of liquidity trades is 232.2, matching the value for within-sample small caps for a generic year, 2007 (see Table B1 in

---

[27] In EOH92, the conditional probability of a sell liquidity order is $\gamma\epsilon^S$, where γ is the proportion of liquidity sellers and ϵ^S is the probability that the liquidity trader submits an sell order after checking the quote. Therefore, the notation here can be seen as $p=\gamma\epsilon^S=(1-\gamma)\epsilon^B$, like in proposition 4 of their paper.

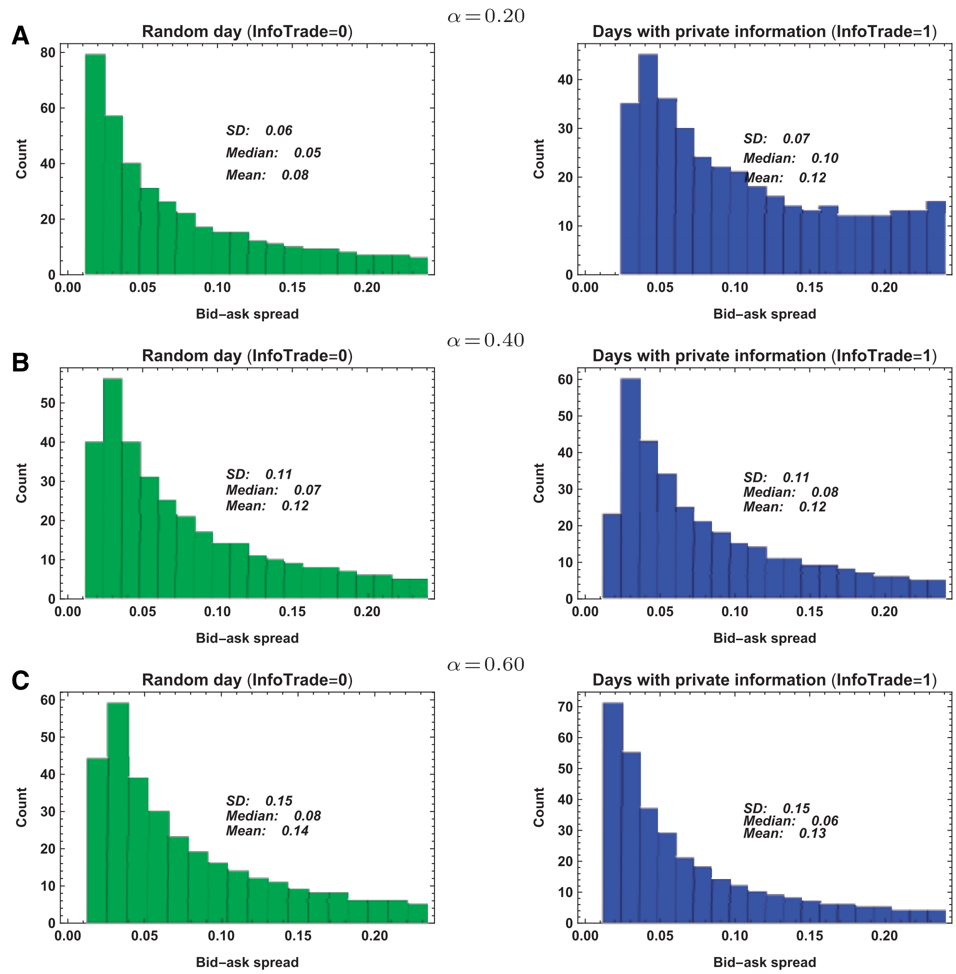


Figure 7
Simulated distribution of the bid-ask spread with event uncertainty
This figure presents the simulated distribution of the bid–ask spread over a trading date using the algorithm and parameter values described in Section 7. In each panel, for a given probability of informed trading, α, the left-side histogram shows the unconditional distribution of the bid-ask spread, that is, the distribution for a random day. The right-side histogram shows the conditional distribution of the bid-ask spread, that is, the distribution for days with informed trading.

the Online Appendix), resulting in $p=0.297$. Given p, we set $\mu=0.147$ so that the proportion of informed trades on days with private information is 22.5%, matching the average between the shares of informed trades for stocks and options given in Table 11. We consider three different values for the probability of an information event, $\alpha \in \{0.2, 0.4, 0.6\}$, and simulate 20,000 trading days in each case.

Figure 7 shows the resultant daily distributions of the bid-ask spread. On the left, each panel shows the distribution for a random day, with a probability of informed trading given by α, similar to $InfoTrade=0$ days. On the right, each panel shows the distribution for days with informed traders, similar to $InfoTrade=1$ days. Panel A shows that when α is relatively low, the average bid-ask spread is higher on days with informed traders, implying that

$\Delta_{\text{spread}} > 0$. For an intermediate value, $\alpha = 0.4$, panel B shows that the average and median values are approximately equal. For a relatively large value of α, however, panel C shows that the average value of the bid-ask spread is *lower* on days with private information, consistent with our initial conjecture. Therefore, everything else being constant, one could observe $\Delta_{\text{spread}} < 0$ for sufficiently large values of α. To gain further intuition, Figure 8 illustrates the expected evolution of the of the bid-ask spread on days with and without private information, that is, $\psi \in \{L, H\}$ and $\psi = 0$, respectively. When $\alpha = 0.2$, the observation of high volume offers the market maker a strong signal of the presence of informed traders. The bid-ask spread remains at a relatively high level initially then decreases as information provided by order flow imbalance aggregates, but at a relatively low rate, as panel A illustrates. On the other hand, with $\alpha = 0.6$, the simulation results displayed in panel C confirm that the bid-ask spread converges to zero more rapidly on days with private information. Given that high volume is anticipated with a high α prior belief, the market maker puts a relatively higher weight on the order imbalance signal in this case, and quotes move to their full information values more rapidly (thus $\Delta_{\text{spread}} < 0$). For the intermediate value $\alpha = 0.4$, panel B shows that the speeds of convergence of the bid-ask spread on random days and days with informed traders are approximately the same.

We contrast the implications of EOH92 for the finite sample distribution of the bid-ask spread with those for the trade price variance. Figure 9 shows the conditional distribution of the price variance using the same parameter values. We can observe that in contrast to the bid-ask spread, all values of α imply that the price variance is higher on days with private information. Indeed, as Easley and O'Hara argue (EOH92, p. 598), as trades are positively correlated in the model, times of low variances tend to be grouped and occur in periods with little trade. Therefore, variances are positively correlated with volume. Their prediction and the simulation results here are consistent with the empirical findings in this paper regarding the positive conditional response of *Realized variance* to the presence of informed traders.

The theoretical analysis in this section suggests that rational learning in the presence of event uncertainty can play a role in justifying the observed negative conditional response of the bid-ask spread. The fact that such a negative response depends on the PIN value suggests that the learning channel we consider is unlikely to be the single one at play. At the same time, we find that a negative conditional response is more likely for high PIN values. This theoretical prediction is consistent with the empirical finding that small caps display a stronger negative response, as panels D and E of Table 13 indicate.

We note that in the canonical event uncertainty model that we have considered, the informed trader only submits market orders in a probabilistic fashion. Such characterization of the informed trader is unlike that considered in Section 5, where the informed trader timing channel and a limit order channel were considered. Therefore, the mechanism considered in this section is not

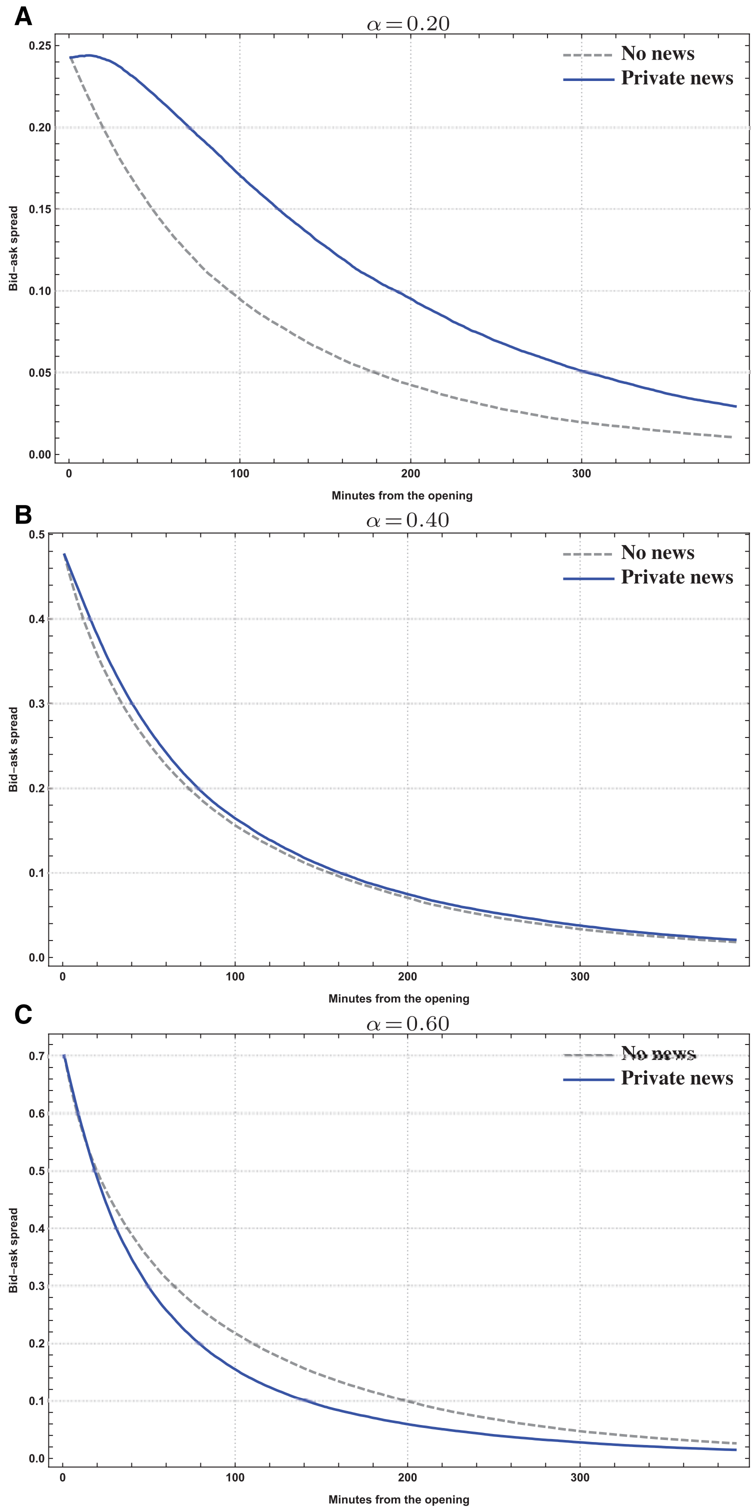


Figure 8
Simulated time series of the bid-ask spread with event uncertainty
This figure presents simulation results for the average value of the bid-ask spread over a trading session (390 minutes) using the algorithm and parameter values described in Section 7. In each panel, for a given probability of informed trading, α, the dashed line (solid line) shows the time series of the average bid-ask spread for days with no informed trading (with informed trading).

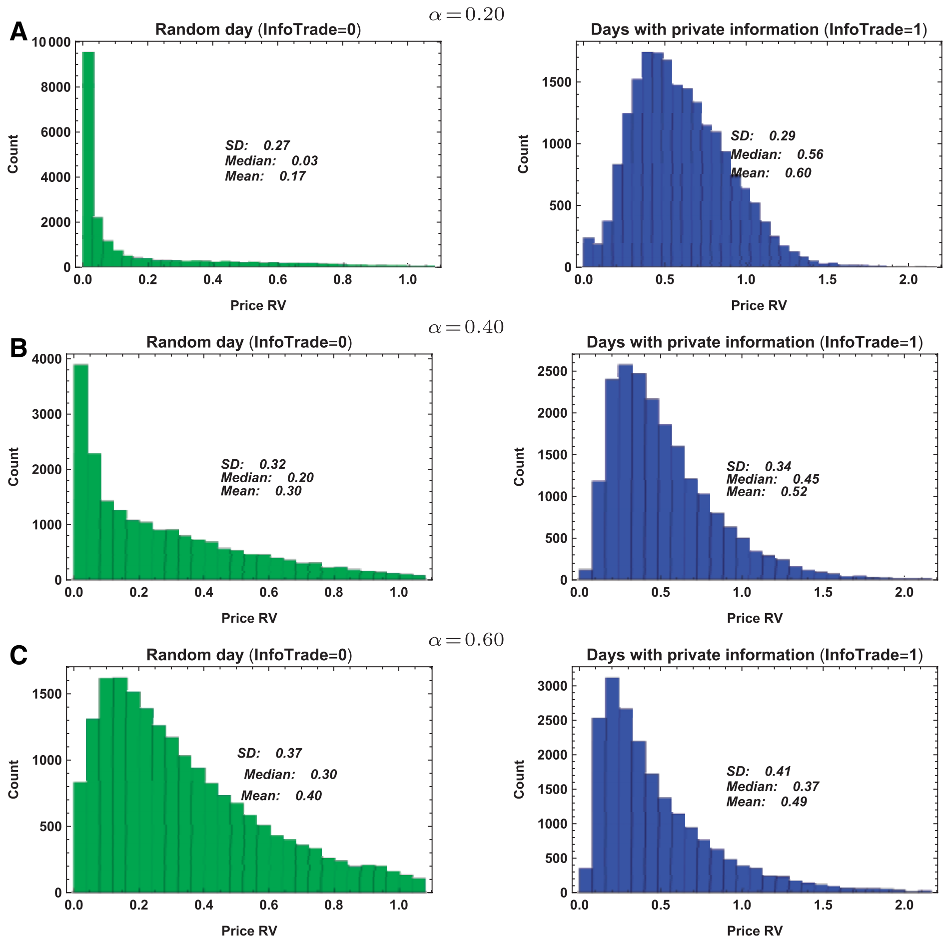


Figure 9
Simulated distribution of the realized variance with event uncertainty
This figure presents the simulated distribution of the realized price variance over a trading date using the algorithm and parameter values described in Section 7. In each panel, for a given probability of informed trading, α, the left-side histogram shows the unconditional distribution of the realized variance, that is, the distribution for a given random day. The right-side histogram shows the conditional distribution of the bid-ask spread, that is, the distribution for days with informed trading.

implied by the results in Section 5. Of course, the strategic trading channel and the market making channel are not mutually exclusive and it would be interesting to study their interaction further in future theoretical work.

7.2 Do prices respond to informed trading? Learning and price convergence

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 *InfoTrade*=1 days. Panel A of Table 14 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

Table 14
Returns and Learning

A. 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

B. Convergence

	Small caps			Large caps		
	Realized volatility	Price range	Bid-ask spread	Realized volatility	Price range	Bid-ask spread
InfoTrade	0.567**	0.895***	−0.148***	−0.119	−0.267	0.013
	(0.233)	(0.332)	(0.041)	(0.107)	(0.178)	(0.021)
Convergence	0.001	−0.257	0.024	−0.152	−0.462**	0.045*
	(0.267)	(0.364)	(0.052)	(0.121)	(0.211)	(0.023)
Interaction	0.645*	1.045**	−0.026	0.399***	1.118***	−0.038
	(0.382)	(0.471)	(0.055)	(0.150)	(0.270)	(0.028)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#obs	6,007	6,026	6,000	4,208	4,208	4,207

This table presents results related to learning patterns in the data. Panel A reports the returns accrued by stock investors on days when informed investors trade for positive and negative news. Market is a value-wighted portfolio of all stocks in CRSP. Panel B compares cases in which the stock return accrued on informed trading day coincides with the direction of news to which insider trading applies to those in which such convergence does not happen. We consider separately cases of small-cap stocks (Columns 1–3) and large-cap stocks (Columns 4–6). *InfoTrade* is an indicator variable equal to 1 for asset-day pairs with informed trading. *Convergence* is an indicator variable equal to 1 if the direction of the return on textitInfoTrade equal 1 day coincides with the sign of the news to which a given trade relates. *Interaction* is a product of *InfoTrade* and *Convergence*. Section 3 defines $InfoTrade$ and the control variables. All regressions include firm and time fixed effects. Standard errors (in parentheses) are clustered around firms.*p <.1;**p <.05;***p <.01.

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 respond to the actions of the informed traders and, on average, change in the direction of private information.

To further evaluate the role of learning, we compare the conditional behavior of the bid-ask spread and price variance for cases in which stock prices seem to be converging to the value of the private signal versus for those that do not. We define *Convergence* as a binary variable that equals 1 if the sign of the stock return on the day when insiders trade—defined as an open-to-open return—coincides with the direction of the news they trade on, and zero otherwise. Our variable of interest is the interaction term between *InfoTrade* and *Convergence*. Intuitively, during convergence days, learning appears to be occurring. Panel B of Table 14 displays the estimation results, separately for small and large market cap stocks. We include both *Realized variance* and *Price range* as measures of price variance. The interaction coefficient indicates, first, that price variance is indeed higher when learning appears to be taking place, both for small and large market caps. Second, for the bid-ask spread, the interaction coefficients are negative both for small and large caps, although they are not statistically significant. Therefore, when learning takes place, we find that price variance increases further and the bid-ask spread decreases further. Finally, we evaluate the robustness of our results by considering different measures of stock returns and extrapolating our results to a sample of option-based AIPs. Throughout all the specifications, the results, reported in Tables F8 to F10 of the Online Appendix, are consistent with our original interpretation. Taken together, the empirical evidence suggests that the trade-based learning process reinforces the conditional response of AIPs on days when informed traders trade.

8. Discussion of Potential Institutional Biases

In Section 4, we discussed and analyzed the possibility of selection bias regarding the detection of informed traders. In principle, even if the detection were uncorrelated with the value of a given AIPs, the institutional processes could impound a bias if the behavior of such AIPs were correlated with the decision to advance a litigation against a seemingly guilty trader. Although theoretically possible, we argue that the decision to litigate is primarily driven by factors that are in principle unrelated to AIPs. 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 AIPs. Second, factors related to the social consequences, for example, the number of affected firms, or the individual characteristics of the trader, such as recurrent insider trading or the fact that the individual committed other securities-related violations. If repeated offenses were indeed considered, the institutional process could increase the representation of investigations involving multiple firms, but not necessarily the estimates of Δ. Section 4.4 shows that the patterns describing the conditional behavior of AIPs do not depend on the number of firms in the investigation.

Third, the decision to litigate 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 overrepresent cases for which the value of information is high. However, for any signal about the 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 1992; Back and Baruch 2004). Therefore, if this type of selection exists, we could be oversampling less prudent traders, that is, those who could have traded more aggressively to impound information into prices. Under such circumstances, this type of selection would arguably work against the considered AIPs displaying abnormal behavior on days with $InfoTrade = 1$.

To obtain a better perspective on whether the value of information affects our baseline results, we turn to regression analysis. Formally, we define a variable

Strength that, like in Table 4, 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 primary variable of interest. If *Strength* has a monotonic effect on the value of AIPs, this should be captured by the interaction term. The estimation results in Tables F11 and Table F12 of the Online Appendix 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 tiny amounts, like in the theoretical model of DeMarzo, Fishman, and Hagerty (1998). Although this type of bias would be difficult to rule out, we would not expect very small trades to influence market aggregates in the first place. Moreover, the fact that our sample contains many small retail traders suggests that the SEC does not follow a strict litigation rule based on large trade thresholds.

9. Discussion of Implications

A large literature examines the predictions of asymmetric information theories based on the statistical power of publicly available AIPs.[28] Our finding that many of these AIPs 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 AIPs perform and offers new insights for future investigation. In this section, we provide a brief discussion of implications for several strands of related literature.

9.1 Implications for empirical analyses of asymmetric information

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

9.1.2 Volume and signed order flow. We find that AIPs contained in abnormal volumes are positively correlated with the presence of informed

[28] 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, Goldstein, and Jiang (2007), Ferreira and Laux (2007), and Roll, Schwartz, and Subrahmanyam (2009), among many others. Examples also can be found in the accounting literature, for example, Frankel and Li (2004).

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 AIPs, this result calls for more emphasis on volume. This fact seems to be more relevant in the current market environment, given the disruption coming from 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 flow. 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.

9.1.3 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 linking higher levels of illiquidity with greater adverse selection risk. As noted in Section 1, this fact could affect the economic interpretation of results that rely on such approach. Overall, we conclude that inference regarding adverse-selection risk based on illiquidity measures requires a structural approach that explicitly models timing options, as well as the learning process of uninformed participants.

9.1.4 Volatility. The footprints of information are not only reflected in conditional responses of volume and illiquidity but also in those of volatility measures. Thus, our results advocate a broader use of volatility-based AIPs 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 et al. (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.

9.1.5 Combining AIPs. 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 AIPs. Bharath, Pasquariello, and Wu (2009) develop an index of asymmetric information by combining seven AIPs using principal component analysis.[29] This approach is plausible because it does not rely on a

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

single AIPs to draw economic conclusions. However, the choice of AIPs in such exercises—how many, which 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 firms' 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).

9.2 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 only consider trade volume as AIPs and trigger an investigation if volume $> \overline{v}$, for a given threshold $\overline{v}$. Generally, a regulator could consider a vector of proxies, $\underline{\text{AIP}}$, and design a rule that triggers an investigation if $\mathbb{P}(I=1, G=1|\text{AIP}) > \overline{P}$. 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. Such rule will affect the equilibrium behavior of informed traders, such as in Kacperczyk and Pagnotta (2018). We hope that our results provide guidance for future endeavors in this spirit.

9.3 Implications for theories of informed trading

9.3.1 Multimarket approach. We find direct evidence that precisely identified informed traders actively include stocks *and* options in their trading strategies, providing support to the analyses in Back (1993), Biais and Hillion (1994), and Easley, O'Hara, and Srinivas (1998), among others. This fact highlights the importance of modeling the information transmission process from a multiasset 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.

9.3.2 Information structure and Trading Strategies. The structure of the PIN model has been enriched and extended by Easley et al. (2008), Odders-White and Ready (2008), and Duarte and Young (2019), 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, like 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 and that less reliance on signed order flow in is desirable. Related to the latter conclusion is the fact that we document that informed traders use both market *and* limit orders, consistent with the predictions in Pagnotta (2010) and Roşu (2016), among others. Intuitively, informed limit orders weaken the connection between the aggressive side of a trade and the use of private signals.

10. Concluding Remarks

In this paper, we have provided a comprehensive evaluation of AIPs in stock and option markets using trades that are *unequivocally* based on nonpublic information about firm fundamentals. We believe that our findings can inform future empirical and theoretical research on adverse selection in financial markets. We conclude discussing some potential limitations and opportunities for future work.

Our conclusions about the conditional behavior of AIPs are robust to several tests of potential selection bias in the informed trade detection method. Despite these robustness results and the fact that our sample contains a significant cross-sectional heterogeneity in traders background and institutional roles, we do not claim that informed traders in the investigations represent every type of informed trader. For example, in our sample, an important proportion receive a private signal about a future corporate announcement, chiefly an acquisition plan of a target firm, or a surprise in earnings announcements. The behavior of such trader is likely influenced by a high degree of confidence in the validity of the signal and by a relatively short information horizon. Therefore, everything else being constant, traders in the sample may behave relatively more aggressively than, say, a portfolio manager that generates his signal through fundamental research and expects the discrepancies between the market price and the fundamental value to dissipate over several years. The latter could behave more cautiously and have a more moderate impact on the behavior of AIPs. On the other hand, despite information precision and the length of information horizons, insider traders may trade relatively cautiously if they anticipate the risk of enforcement actions. For example, they could rely relatively less on out-of-the-money options relative to an investor who does not internalize such risk. In principle, the net effect of these factors could be that insider traders are not significantly more or less aggressive than other informed traders. It is, however, challenging to assert precisely the net effect of these factors in a nonlaboratory setting. Future work could expand the stylized experimental setting in Bloomfield, O'Hara, and Saar (2005) to address this issue.

Our empirical design relies on the canonical information structure of Easley and O'Hara (1992) in which private information arrives randomly and privately informed traders are not found every trading day. An exciting avenue for future research is the study of more sophisticated and possibly more realistic information structures. 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 in Banerjee and Breon-Drish (2017) suggest that inference is also sensitive to the introduction of an information acquisition decision. More work in this area will help to better understand the interaction between informed investors, market makers, and other market participants.

References

Admati, A. R., and P. Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies* 1:3–40.

Ahern, K. R. 2018. Do proxies for informed trading measure informed trading? Evidence from illegal insider trades. Working Paper.

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* 69:2279–337.

Back, K. 1992. Insider trading in continuous time. *Review of Financial Studies* 5:387–409.

———. 1993. Asymmetric information and options. *Review of Financial Studies* 6:435–72.

Back, K., and S. Baruch. 2004. Information in securities markets: Kyle meets Glosten and Milgrom. *Econometrica* 72:433–65.

Back, K., and K. Crotty. 2015. The informational role of stock and bond volume. *Review of Financial Studies* 28:1381–427.

Back, K., K. Crotty, and T. Li. 2018. Identifying information asymmetry in securities markets. *Review of Financial Studies* 31:2277–325.

Banerjee, S., and B. Breon-Drish. 2017. Dynamic information acquisition and strategic trading. Working Paper.

Baruch, S., M. Panayides, and K. Venkataraman. 2017. Informed trading and price discovery before corporate events. *Journal of Financial Economics* 125:561–88.

Bharath, S. T., P. Pasquariello, and G. Wu. 2009. Does asymmetric information drive capital structure decisions. *Review of Financial Studies* 22:3211–43.

Biais, B., L. Glosten, and C. Spatt. 2005. Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets* 8:217–64.

Biais, B., and P. Hillion. 1994. Insider and liquidity trading in stock and options markets. *Review of Financial Studies* 7:743–80.

Black, F. 1975. Fact and fantasy in the use of options. *Financial Analysts Journal* 31:36–41.

Bloomfield, R. J., M. O'Hara, and G. Saar. 2005. The 'make or take' decision in an electronic market: Evidence on the evolution of liquidity. *Journal of Financial Economics* 75:165–99.

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* 28:2723–69.

Chae, J. 2005. Trading volume, information asymmetry, and timing information. *Journal of Finance* 60:413–42.

Chakravarty, S., H. Gulen, and S. Mayhew. 2004. Informed trading in stock and option markets. *Journal of Finance* 59:1235–57.

Chan, K., Y. P. Chung, and W.-M. Fong. 2002. The informational role of stock and option volume. *Review of Financial Studies* 15:1049–75.

Chen, Q., I. Goldstein, and W. Jiang. 2007. Price informativeness and investment sensitivity to stock price. *Review of Financial Studies* 20:619–50.

Cohen, L., C. Malloy, and L. Pomorski. 2012. Decoding inside information. *Journal of Finance* 67:1009–43.

Collin-Dufresne, P., and V. Fos. 2015. Do prices reveal the presence of informed trading? *Journal of Finance* 70:1555–82.

———. 2016. Insider trading, stochastic liquidity and equilibrium prices. *Econometrica* 84:1441–75.

Cremers, M., and D. Weinbaum. 2010. Deviations from put–call parity and stock return predictability. *Journal of Financial and Quantitative Analysis* 45:335–67.

DeMarzo, P. M., M. J. Fishman, and K. M. Hagerty. 1998. The optimal enforcement of insider trading regulations. *Journal of Political Economy* 106:602–32.

Del Guercio, D., E. R. Odders-White, and M. J. Ready. 2017. The deterrent effect of the Securities and Exchange Commission's enforcement intensity on illegal insider trading: Evidence from run-up before news events. *Journal of Law and Economics* 60:269–307.

Duarte, J., and L. Young. 2009. Why is PIN priced? *Journal of Financial Economics* 91:119–38.

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* 120:269–85.

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? *Journal of Finance* 57:2185–221.

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* 53:431–65.

Ferreira, M. A., and P. A. Laux. 2007. Corporate governance, idiosyncratic risk, and information flow. *Journal of Finance* 62:951–89.

FINRA. 2015. Catching the bad guys: Inside FINRA's fraud surveillance and insider trading groups. September 8, 2015. https://www.finra.org/investors/catching-bad-guys-inside-finras-fraud-surveillance-and-insider-trading-groups.

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* 37:229–59.

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* 92:153–81.

Grossman, S. 1976. On the efficiency of competitive stock markets where traders have diverse information. *Journal of Finance* 31:573–85.

Hasbrouck, J. 2009. Trading costs and returns for U.S. Equities: Estimating effective costs from daily data. *Journal of Finance* 64:1445–77.

Hendershott, T., D. Livdan, and N. Schurhoff. 2015. Are institutions informed about news? *Journal of Financial Economics* 117:249–87.

Holden, C. W., and S. E. Jacobsen. 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance* 69:1747–85.

Johnson, T. L., and E. C. So. 2017. A simple multimarket measure of information asymmetry. *Management Science* 1909:1–26.

Kacperczyk, M., and E. S. Pagnotta. 2018. Becker meets Kyle: Inside insider trading. Working Paper.

Kacperczyk, M., S. van Nieuwerburgh, and L. Veldkamp. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84:571–626.

Kelley, E. K., and P. C. Tetlock. 2013. How wise are crowds? Insights from retail orders and stock returns. *Journal of Finance* 68:1229–65.

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* 53:1315–35.

Meulbroek, L. K. 1992. An empirical analysis of illegal insider trading. *Journal of Finance* 47:1661–99.

Odders-White, E. R., and M. J. Ready. 2008. The probability and magnitude of information events. *Journal of Financial Economics* 87:227–48.

Pagnotta, E. S. 2010. Information and liquidity trading at optimal frequencies. Working Paper.

Roll, R., E. Schwartz, and A. Subrahmanyam. 2009. Options trading activity and firm valuation. *Journal of Financial Economics* 94:345–60.

Roşu, I. 2016. Liquidity and information in order driven markets. Working Paper.

Wang, J. 1993. A model of intertemporal asset prices under asymmetric information. *Review of Economic Studies* 60:249–82.

Wang, Y., and M. Yang. 2017. Insider trading when there may not be an insider. Working Paper.

Online Appendix to "Chasing Private Information"

Marcin Kacperczyk* **Emiliano S. Pagnotta†**

A Supplement to Section 2: The Construction of AIPs

In this appendix, we provide formal definitions of the AIPs and discuss their empirical implementation. For clarity, we group AIPs into volatility, volume, and illiquidity, and based on whether they originate in stock or option markets. Table B1 provides summary statistics for the number of trades for the sample period.

A.1 AIPs in Stock Markets

Stock-based AIPs 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 AIPs at the 1% level to mitigate the influence of outliers. In addition to dollar-weighted averages, we also compute intraday stock-based AIPs 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.

Price Range. We define the daily price range as

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

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

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

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:

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

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 AIPs 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 AIPs are winsorized at the 1% level to mitigate the influence of

[2] Two related AIPss 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.

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 date 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 Supplement to Section 3: Additional Summary Statistics

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
Traders Professional Background and Corporate Positions

This table reports the numbers of insider traders in our sample distributed according to their job category. In **Panel A**, we break down the sample by the corporate function; **in Panel B**, we break down the sample according to the types of finance jobs; in Panel C, we summarize the data according to non-finance jobs. The sample period is 1995–2015.

Category	Number
Panel A: Corporate Function	
Employee/Low-level management	105
Mid-level management	84
Vice president	67
CEO	38
Director	32
President	32
CFO	24
Other	28
Panel B: Job in Finance	
Portfolio manager/Hedge fund	59
Broker or dealer	30
Analyst	21
Trader	16
Financial Advisor	7
Other	16
Panel C: Non-finance Jobs	
Business owner/self employed	52
Lawyer/Attorney	42
Medical doctor / dentist	24
Accountant	14
Sales	13
Engineer /IT	12
Real estate broker	11
Other	26

C Supplement to Section 4

Table C3
Controlling for Analysts' Forecast Revisions

The dependent variables are AIPs. **Panels A and B** report the results for stock-based and option-based AIPs. $InfoTrade$ 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. *Revision* measures the change on a given day relative to the previous consensus value of earnings. Earnings are one-year forecasts. 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 AIPs									
InfoTrade	0.698***	1.140***	794.432**	401.027***	-0.100***	-0.106	-0.015***	-0.032***	-0.419***
	(0.118)	(0.131)	(393.864)	(109.655)	(0.015)	(0.473)	(0.005)	(0.007)	(0.071)
Revision	-0.000*	-0.000***	-0.037***	-0.029***	0.000***	0.000**	0.000***	0.000***	-0.000***
	(0.000)	(0.000)	(0.013)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,318	10,357	10,248	10,317	10,240	10,190	10,190	10,175	10,282

	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 B: Option-based AIPs								
InfoTrade	0.021***	0.016**	0.006*	3,300.957***	-4.095*	-0.065***	-0.074***	-0.130***
	(0.006)	(0.007)	(0.003)	(613.655)	(2.138)	(0.015)	(0.021)	(0.026)
Revision	-0.000**	-0.000***	0.000***	2.956*	-0.009***	-0.000*	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(1.671)	(0.002)	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,583	2,571	2,332	2,588	2,588	2,533	2,533	2,486

Table C4
Counterfactual Insider Trading Dates: Stock-based AIPs

This table presents results for regressions with alternative (counterfactual) info trade dates. The dependent variables are stock-based AIPs. **Panel A** lags info trade by one trading day, **Panel B** by two trading days, **Panel C** by three trading days, and **Panel D** by four trading days. $InfoTrade$ 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: 1-day Lag					
InfoTrade$_{-1}$	0.241**	0.368***	561.713	35.511	-0.050***	-0.878	-0.009*	-0.011	0.015
	(0.099)	(0.120)	(386.059)	(107.840)	(0.018)	(0.548)	(0.005)	(0.008)	(0.109)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	9,083	9,118	9,008	9,093	9,001	8,944	8,944	8,934	9,042
				Panel B: 2-day Lag					
InfoTrade$_{-2}$	0.660***	1.243***	-147.520	2,299.898***	-0.103***	-0.772	-0.013**	-0.059***	-0.283***
	(0.134)	(0.158)	(263.986)	(228.565)	(0.023)	(0.532)	(0.005)	(0.009)	(0.096)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	9,129	9,164	9,055	9,138	9,050	8,996	8,996	8,982	9,085
				Panel C: 3-day Lag					
InfoTrade$_{-3}$	0.049	0.106	-483.283**	-26.566	-0.039**	0.381	-0.014**	-0.016*	0.017
	(0.097)	(0.122)	(209.210)	(108.555)	(0.019)	(0.591)	(0.006)	(0.009)	(0.091)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	8,920	8,953	8,843	8,928	8,837	8,778	8,778	8,764	8,871
				Panel D: 4-day Lag					
InfoTrade$_{-4}$	0.047	-0.078	132.733	-158.876	-0.050**	-0.285	-0.006	-0.013	-0.217**
	(0.099)	(0.123)	(361.998)	(100.741)	(0.022)	(0.636)	(0.006)	(0.009)	(0.103)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	8,886	8,918	8,811	8,894	8,804	8,751	8,751	8,738	8,843

Table C4 (Continued)
Counterfactual Insider Trading Dates: Option-based AIPs

This table presents results for regressions with alternative (countrfactual) info trade dates. The dependent variables are option-based AIPs. **Panel A** lags info trade by one trading day, **Panel B** by two trading days, **Panel C** by three trading days, and **Panel D** by four trading days. $InfoTrade$ 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.
Panel A: 1-day Lag								
InfoTrade_{-1}	0.017***	0.017**	0.005	1,684.230**	2.421	-0.043**	-0.059**	-0.074**
	(0.005)	(0.007)	(0.004)	(726.922)	(2.762)	(0.018)	(0.025)	(0.036)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,820	1,812	1,655	1,825	1,825	1,776	1,776	1,718
Panel B: 2-day Lag								
InfoTrade_{-2}	0.007	0.016*	0.001	2,847.940***	-3.294	-0.021	0.039	-0.073*
	(0.008)	(0.009)	(0.004)	(1,099.086)	(2.550)	(0.020)	(0.032)	(0.039)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,879	1,870	1,711	1,758	1,886	1,833	1,833	1,773
Panel C: 3-day Lag								
InfoTrade_{-3}	0.017***	0.003	0.015***	1,601.316*	2.317	-0.034*	-0.042	-0.002
	(0.006)	(0.007)	(0.004)	(900.227)	(2.920)	(0.019)	(0.027)	(0.059)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,748	1,741	1,591	1,754	1,754	1,705	1,705	1,650
Panel D: 4-day Lag								
InfoTrade_{-4}	0.012**	0.008	0.007*	2,230.736**	1.295	-0.040**	-0.047	-0.064*
	(0.005)	(0.008)	(0.004)	(1,010.487)	(2.850)	(0.019)	(0.029)	(0.036)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,737	1,731	1,581	1,743	1,743	1,694	1,694	1,640

D Supplement to Section 5

Table D5
Summary Statistics on Detection Source Subsamples

This table reports summary statistics for the three groups of cases investigated in our sample. WRP/Non-WRP refers to cases investigated using Whistleblower Reward Program over the period 2011-2015. FINRA/Non-FINRA refers to cases investigated by FINRA and those where FINRA is not a source of investigation. Small/Large Investig. refers to cases involving a small number of companies (<=2) and a large number of companies (>2). All groups are matched using the same time periods.

Characteristic/Sample	WRP	Non-WRP	FINRA	Non-FINRA	Small Investig.	Large Investig.
Sample Period	2011–2015	2011-2015	2004–2015	2004–2015	1995-2015	1995-2015
Number of investigations	55	47	117	161	324	104
Average Market Cap (in $bn)	13.29	4.89	4.94	9.62	4.81	14.98
Turnover (in % per day)	1.9	1.65	1.47	1.77	1.56	1.65
# days from private tip to a trade	17.16	20.01	23.21	15.22	18.03	11.35
# days from a trade to information disclosure	13.95	14.85	18.3	17.27	25.69	12.52
# days from the first to the last trade	21.72	16.32	13.65	22.00	34.71	25.82
Trades per firm	26.46	34.30	19.06	23.32	13.42	16.61
Trades per trader	21.05	19.8	20.17	17.01	4.65	17.00
Reported profits ($ millions)	1.58	0.31	0.31	1.44	1.10	1.13

E Supplement to Section 6

Table E6
Abnormal Volume Net of Informed Trading

This table presents results for abnormal volume (Columns 1 and 4), seemingly uninformed volume defined as normal volume net of informed trades (Columns 2 and 5), and total volume (Columns 3 and 6) for stock-based and option-based AIPs, respectively. Whenever a total trade size is available, but not the distribution of quantities across dates, the left hand size variable is computed under the following alternative assumptions. In calculating net volume, we linearly interpolate trade size over the trading horizon. $InfoTrade$ 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.

	Abnormal Volume (1)	S. Uninformed Volume (2)	Total Volume (3)	Abnormal Volume (4)	S. Uninformed Volume (5)	Total Volume (6)
	Stock-based AIPs			Option-based AIPs		
Panel A: Full Sample						
InfoTrade	395.226***	320.285***	567.074***	3,550.558***	2,552.382***	1,622.495***
	(109.656)	(43.721)	(107.971)	(707.034)	(416.646)	(316.110)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,317	10,251	10,358	2,588	2,432	2,596
Panel B: WRP Sample						
InfoTrade	717.336*	755.427***	960.770***	3,222.437*	3,119.018***	2,278.556**
	(418.341)	(197.896)	(29.348)	(1,793.005)	(987.722)	(1,021.369)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	1,650	1,638	1,651	392	356	364
Panel C: no-FINRA Sample						
InfoTrade	374.422***	358.074***	618.780***	4,374.279***	2,137.532***	1,476.780***
	(123.643)	(69.252)	(12.775)	(1,156.143)	(498.705)	(382.297)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	7,674	7,630	7,713	1,589	1,483	1,540
Panel D: Large Investigations Sample						
InfoTrade	635.460***	357.638***	604.800***	1,911.568***	1,781.835***	1,080.072**
	(187.243)	(107.108)	(13.727)	(725.220)	(464.460)	(428.285)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	5,157	5,132	5,159	1,254	1,205	1,230

Table E7
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 AIPs. **Panels A and B** report the results for stock-based AIPs and **Panels C and D** the results for option-based AIPs. $InfoTrade$ 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 AIPs: 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 AIPs: 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 AIPs: 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 AIPs: 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

F Supplement to Sections 8 and 9

Table F8
Conditioning on Price Convergence: Stock-based AIPs

This table presents results for cases conditioning on whether returns match the sign of the private information tip. The dependent variables are stock-based AIPs. **Panels A, B, and C** report the results under three different definitions of conditioning returns: open to open, close to close, and open to close. $InfoTrade$ 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: Open-to-open Return									
InfoTrade	0.398***	0.529**	824.084	167.117	-0.122***	1.026	-0.009	-0.029**	-0.429***
	(0.145)	(0.215)	(654.790)	(177.787)	(0.028)	(0.963)	(0.009)	(0.014)	(0.120)
Convergence	0.003	-0.251	634.933	-186.522	-0.038	0.276	0.006	0.001	-0.097
	(0.164)	(0.216)	(676.790)	(207.106)	(0.029)	(0.976)	(0.009)	(0.014)	(0.143)
Interaction	0.497**	0.962***	-8.470	354.977	0.028	-1.840	-0.008	-0.007	-0.001
	(0.243)	(0.295)	(878.342)	(232.113)	(0.035)	(1.139)	(0.012)	(0.017)	(0.151)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,215	10,234	10,157	10,194	10,207	10,157	10,157	10,142	10,162
Panel B: Close-to-close Return									
InfoTrade	0.408***	0.406**	443.713	96.612	-0.048*	0.670	-0.003	-0.012	-0.257***
	(0.137)	(0.185)	(537.556)	(109.737)	(0.025)	(0.741)	(0.008)	(0.011)	(0.080)
Convergence	0.062	-0.261	45.078	-206.495	0.117***	0.545	0.032***	0.025**	0.347***
	(0.154)	(0.201)	(566.012)	(174.107)	(0.028)	(0.882)	(0.008)	(0.012)	(0.113)
Interaction	0.527**	1.252***	641.669	508.752**	-0.072**	-1.295	-0.014	-0.031**	-0.225*
	(0.237)	(0.279)	(834.351)	(201.733)	(0.034)	(0.997)	(0.011)	(0.015)	(0.129)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,249	10,268	10,191	10,228	10,241	10,191	10,191	10,176	10,196
Panel C: Open-to-close Return									
InfoTrade	0.360**	0.478**	652.428	78.959	-0.127***	1.031	-0.012	-0.028*	-0.483***
	(0.148)	(0.213)	(694.231)	(153.154)	(0.028)	(0.986)	(0.009)	(0.015)	(0.104)
Convergence	-0.024	-0.291	711.211	-130.464	-0.032	0.456	0.008	0.001	-0.039
	(0.167)	(0.222)	(597.610)	(206.379)	(0.030)	(0.995)	(0.010)	(0.014)	(0.134)
Interaction	0.538**	1.013***	254.995	497.461**	0.036	-1.770	-0.003	-0.008	0.096
	(0.248)	(0.294)	(957.417)	(219.406)	(0.034)	(1.151)	(0.011)	(0.018)	(0.138)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,215	10,234	10,157	10,194	10,207	10,157	10,157	10,142	10,162

Table F9
Conditioning on Price Convergence: Option-Based AIPs

This table presents results for cases conditioning on whether returns match the sign of the private information tip. The dependent variables are option-based AIPs. **Panels A, B, and C** report the results under three different definitions of conditioning returns: open to open, close to close, and open to close. $InfoTrade$ 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.
Panel A: Open-to-open Return								
InfoTrade	0.032***	0.010	-0.002	1,269.034	6.670	-0.122***	-0.176***	-0.192***
	(0.008)	(0.011)	(0.006)	(820.959)	(4.427)	(0.030)	(0.039)	(0.049)
Convergence	0.014	-0.000	-0.019**	-1,798.473*	9.203**	-0.084**	-0.113**	-0.079
	(0.011)	(0.014)	(0.008)	(1,024.217)	(4.638)	(0.035)	(0.048)	(0.058)
Interaction	-0.015	0.010	0.010	3,028.096**	-16.116***	0.080**	0.147***	0.087
	(0.013)	(0.015)	(0.008)	(1,303.836)	(5.162)	(0.037)	(0.050)	(0.063)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,582	2,570	2,331	2,587	2,587	2,532	2,532	2,485
Panel B: Close-to-close Return								
InfoTrade	0.044***	0.015*	0.022***	1,164.072**	1.397	-0.060**	-0.075**	-0.248***
	(0.008)	(0.008)	(0.006)	(578.311)	(3.298)	(0.026)	(0.034)	(0.050)
Conv	0.033***	-0.006	0.019***	-1,725.365*	6.188	0.016	0.055	-0.158***
	(0.011)	(0.012)	(0.007)	(952.237)	(3.895)	(0.032)	(0.045)	(0.054)
Interaction	-0.038***	0.000	-0.028***	3,747.508***	-9.361**	-0.007	0.010	0.196***
	(0.014)	(0.014)	(0.008)	(1,258.617)	(4.512)	(0.034)	(0.048)	(0.064)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,583	2,571	2,332	2,588	2,588	2,533	2,533	2,486
Panel C: Open-to-close Return								
InfoTrade	0.035***	0.014	-0.003	1,432.550*	6.076	-0.126***	-0.177***	-0.196***
	(0.008)	(0.011)	(0.006)	(848.322)	(4.289)	(0.031)	(0.040)	(0.050)
Conv	0.011	-0.005	-0.017**	-2,121.003**	10.697**	-0.080**	-0.116**	-0.076
	(0.011)	(0.014)	(0.007)	(1,023.758)	(4.921)	(0.034)	(0.047)	(0.057)
Interaction	-0.021	0.002	0.011	2,683.782**	-14.782***	0.086**	0.148***	0.094
	(0.013)	(0.015)	(0.008)	(1,363.693)	(4.917)	(0.039)	(0.051)	(0.064)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,582	2,570	2,331	2,587	2,587	2,532	2,532	2,485

Table F10
Conditioning on Price Convergence and Market Cap: Option-Based AIPs

This table presents results for cases conditioning on whether returns match the sign of the private information tip. The dependent variables are AIPs. **Panel A** reports the results for large-cap option-based AIPs under alternative conditioning returns. **Panel B** presents results for small-cap option-based AIPs. $InfoTrade$ 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.
Panel A: Open-to-close Return: Large Caps								
InfoTrade	0.029***	0.016	-0.009	1,840.765*	6.370	-0.124***	-0.202***	-0.182***
	(0.010)	(0.014)	(0.008)	(1,016.563)	(4.558)	(0.031)	(0.046)	(0.061)
Conf	0.007	-0.013	-0.012	-1,729.688	10.141*	-0.072**	-0.120**	-0.084
	(0.010)	(0.015)	(0.008)	(1,207.051)	(5.347)	(0.033)	(0.051)	(0.064)
Interaction	-0.016	0.009	0.009	2,362.060	-13.729***	0.072*	0.157***	0.084
	(0.012)	(0.016)	(0.009)	(1,588.445)	(4.977)	(0.038)	(0.054)	(0.070)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	2,127	2,123	1,914	2,132	2,132	2,077	2,077	2,070
Panel B: Open-to-close Return: Small Caps								
InfoTrade	0.047***	0.003	0.012	542.403	5.950	-0.106	-0.056	-0.265***
	(0.011)	(0.015)	(0.009)	(1,617.479)	(11.452)	(0.087)	(0.072)	(0.079)
Conv	-0.027	0.012	-0.074***	-1,188.061	20.406	0.010	0.046	-0.031
	(0.031)	(0.030)	(0.024)	(1,595.310)	(14.314)	(0.101)	(0.108)	(0.180)
Interaction	0.040	-0.003	0.065**	-305.298	-25.395	0.031	-0.040	0.172
	(0.048)	(0.042)	(0.031)	(1,812.681)	(17.753)	(0.116)	(0.140)	(0.236)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	455	447	417	455	455	455	455	415

Table F11
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 AIPs. **Panel A** reports the results for stock-based AIPs, **Panel B** the results for option-based AIPs. $InfoTrade$ 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 AIPs

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 AIPs

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 F12
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 AIPs. **Panel A** reports the results for stock-based AIPs, **Panel B** the results for option-based AIPs. $InfoTrade$ 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 AIPs

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 AIPs

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

G Additional Tests

Table G13
Conditioning on Volume Percentage: Stock-based AIPs

The dependent variables are stock-based AIPs. VolPerc is an indicator variable equal to one if stock volume exceeds a 1% (**Panel A**), 5% (**Panel B**), and 10% (**Panel C**) cutoff, and it is not the date of a material public announcement. $InfoTrade$ 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		Illiquidity				
	Realized Volatility	Price Range	Price Inform.	Quoted Spread	Price Impact	Order Imb.	Lambda
			Panel A: 1% cutoff				
InfoTrade	0.645***	1.060***	808.801**	-0.098***	0.004	-0.015***	-0.029***
	(0.118)	(0.129)	(398.841)	(0.016)	(0.475)	(0.005)	(0.007)
VolPerc	6.917***	9.186***	-374.471	-0.041	1.262	-0.009	-0.044*
	(1.304)	(1.100)	(549.344)	(0.056)	(1.693)	(0.018)	(0.023)
Interaction	-2.208	-2.188	-476.981	-0.137	-7.598***	0.001	-0.195***
	(1.747)	(1.636)	(859.769)	(0.104)	(2.938)	(0.031)	(0.074)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,319	10,358	10,249	10,241	10,191	10,191	10,176
			Panel B: 5% cutoff				
InfoTrade	0.378***	0.706***	516.704	-0.087***	-0.329	-0.010*	-0.022***
	(0.098)	(0.099)	(377.390)	(0.015)	(0.476)	(0.005)	(0.007)
VolPerc	3.156***	5.033***	141.308	-0.045*	1.051	-0.016**	-0.032**
	(0.354)	(0.330)	(603.730)	(0.024)	(0.898)	(0.008)	(0.013)
Interaction	1.315*	1.323*	2,944.229	-0.123	1.598	-0.041**	-0.093**
	(0.744)	(0.781)	(2,663.155)	(0.080)	(2.560)	(0.019)	(0.037)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,319	10,358	10,249	10,241	10,191	10,191	10,176
			Panel C: 10% cutoff				
InfoTrade	0.323***	0.555***	697.914*	-0.078***	-0.054	-0.010*	-0.021***
	(0.095)	(0.095)	(401.523)	(0.015)	(0.500)	(0.005)	(0.007)
VolPerc	2.224***	3.626***	466.563	-0.038**	1.246**	-0.015***	-0.036***
	(0.196)	(0.190)	(445.410)	(0.016)	(0.631)	(0.005)	(0.008)
Interaction	1.001**	1.526***	326.673	-0.121**	-1.077	-0.018	-0.048**
	(0.465)	(0.491)	(1,464.785)	(0.051)	(1.610)	(0.015)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs	10,319	10,358	10,249	10,241	10,191	10,191	10,176

Table G13 (Continued)
Conditioning on Volume Percentage: Option-based AIPs

The dependent variables are option-based AIPs. VolPerc is an indicator variable equal to one if stock volume exceeds a 1% (**Panel A**), 5% (**Panel B**), and 10% (**Panel C**) cutoff, and it is not the date of a material public announcement. $InfoTrade$ 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			Illiquidity	
	IV Calls	IV Puts	IV Skew	Quoted Spr. (all)	Quoted Spr. (otm)
Panel A: 1% cutoff					
InfoTrade	0.018***	0.014**	0.004	-0.061***	-0.074***
	(0.007)	(0.007)	(0.003)	(0.015)	(0.021)
VolPerc	0.020	0.009	0.005	0.021	0.048
	(0.013)	(0.014)	(0.009)	(0.039)	(0.046)
Interaction	0.084**	0.053	0.057	-0.155	-0.032
	(0.035)	(0.081)	(0.037)	(0.158)	(0.091)
Controls	Yes	Yes	Yes	Yes	Yes
#Obs	2,583	2,571	2,332	2,533	2,533
Panel B: 5% cutoff					
InfoTrade	0.017***	0.014**	0.004	-0.051***	-0.072***
	(0.006)	(0.006)	(0.003)	(0.014)	(0.020)
VolPerc	0.011*	0.007	-0.003	0.026	0.019
	(0.007)	(0.008)	(0.005)	(0.036)	(0.044)
Interaction	0.015	0.005	0.015	-0.104**	-0.028
	(0.019)	(0.021)	(0.009)	(0.051)	(0.066)
Controls	Yes	Yes	Yes	Yes	Yes
#Obs	2,583	2,571	2,332	2,533	2,533
Panel C: 10% cutoff					
InfoTrade	0.012*	0.006	0.003	-0.051***	-0.062***
	(0.006)	(0.006)	(0.004)	(0.015)	(0.021)
VolPerc	0.019***	0.021***	-0.003	-0.010	0.001
	(0.006)	(0.008)	(0.004)	(0.021)	(0.028)
Interaction	0.021	0.021	0.010	-0.042	-0.046
	(0.013)	(0.016)	(0.007)	(0.035)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes
#Obs	2,583	2,571	2,332	2,533	2,533

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.