Nontechnical summary

One of the roles of investors in the economy is to do research on companies and make a judgement on what they are worth. This research activity benefits the economy widely, since better information helps people decide where to direct their limited resources. Investors who do this research need to get paid if they are to continue the job, and a typical way that investors make money from their research is by buying and selling the stock of the companies that they are researching. For example, they can buy the stock of a company that they think is undervalued. However, analysis from academics is showing that this trading might have become costlier than it used to be. If so, the cost could discourage investors from doing much research in the first place, which could mean the economy has less information about where to direct its resources.

The reason for the increase in cost is said to be the rise of high-speed, capital-light trading firms, sometimes grouped as "algorithmic traders." These firms are said to use machine learning and other technologies to infer when investors who have information are trading. The firms can then trade "the same way" as informed investors, buying when they are buying or selling when they are selling, and (in a sense) stealing some of their profits. Theory written by academics give warrant for the claim (Yang and Zhu 2019; Baldauf and Mollner 2020); and empirical work by academics shows that some proxies of high-speed trading are correlated with proxies of bad pricing (Weller 2018; Lee and Watts 2018).

In our paper, we try to fill a gap in the empirical work. It has not been possible to look at a long-term sample of long-term investors and ask after their trading profits or ask about how they interact with high-speed traders. These are both central to the concerns raised in the theoretical work. To ask these research questions, we empirically identify a group of long-term informed investors and a group of high-speed traders in a long sample of equity trading. The data sample comes from the Toronto Stock Exchange, the leading exchange for equity prices in Canada. We identify the trader groups using a set of anonymous account codes in the data, and we study the market over an 11-year sample, 2005–2015, which is the period of the rise of algorithmic trading in Canada.

What we find stands in contrast to some of the empirical findings. The trading volumes and profits of the informed investors are roughly flat through the sample, and our metrics of price efficiency are also flat. This is despite a slow growth (though later, a fall) in the presence of high-speed trading. We find the informed investors are patient and sophisticated in how they trade, trading passively and over long time periods, which helps disguise their presence in markets from a detection algorithm. They also time their trading to coincide with periods of high liquidity in financial markets. They seem so good at their trading that they adversely select the high-speed traders: the informed buy from the high-speed traders before positive earnings surprises, and the informed sell to the high-speed traders before negative earnings surprises. In general, rather than being the "prey" in financial markets, informed traders appear to be the "apex predators."

Some of these findings have policy relevance. It is unclear that efforts to protect informed investors from high-speed traders are really needed. In fact, informed investors seem to be so good at their trading that relatively little of what they know makes it into prices. For example, their price impacts are near zero day-over-day, and the impacts are *negative* intraday. Rather, regulators might rather wish to increase competition in the financial sector among informed investors; other ideas for getting more information into prices are to enable greater trading by insiders or to increase the frequency of earnings reports.

Trading on long-term information*

Corey Garriott[†] Ryan Riordan[‡]

April 12, 2020

Abstract

Predatory trading discourages informed investors from gathering information and trading on it. However, using 11 years of equity trading data, we do not find evidence that informed investors are being discouraged. They have roughly constant volumes and profits through the sample. They are sophisticated, trading patiently over weeks and timing their trading to achieve negative price impacts, leaving price efficiency unchanged. We identify shorter-term traders and, in contrast to theory, find that they supply liquidity by trading in the opposite direction of the informed. Inefficient prices may be the result of informed investors' sophisticated trading and not of predatory short-term trading.

JEL classification: G20, G14, L10

Keywords: Strategic trading, informed trading, price efficiency

*We are grateful to Jason Allen, Jean-Philippe Dion, Jean-Sébastien Fontaine, Slava Fos, Terrence Hendershott, Michael Mueller, Andriy Shkilko, Adrian Walton, Gunther Wuyts, Bart Zhou Yueshen and Marius Zoican, to seminar participants at Wilfrid Laurier University, and to conference participants at the Western Finance Association 2020, HEC Montréal, FiRe Graz, SAFE 2019, and the 15th Central Bank Conference on Market Microstructure, for advice and suggestions. This paper represents the opinions of the authors and does not reflect the views of the Bank of Canada.

[†]Bank of Canada, Financial Markets Department. E-mail: cgarriott@bankofcanada.ca.

[‡]Smith School of Business, Queen's University. E-mail: ryan.riordan@queensu.ca.

Informed investors research the value of investments and trade on what they discover. They buy if asset prices are too low and sell if asset prices are too high, guiding prices to better reflect value. Market participants that do not perform fundamental research rely on the research of informed investors as they must take the price of an asset as given. The same prices, in turn, are used by investors and even by company managers to allocate capital. When prices reflect poorly the value of investments, it can lead to costly managerial decisions and can even diminish economic growth (Dessaint, Foucault, Frésard, and Matray 2019; van Binsbergen and Opp 2017). Thus, anything that makes informed investing more difficult and costly, or that otherwise inhibits the revelation of information in prices, can have far-reaching effects in the economy.

Recent theoretical work has suggested a new source of costs for informed investors (Yang and Zhu 2019; Baldauf and Mollner 2020). Computerized short-term traders could be using machine learning to detect the trading of an informed investor and could use the knowledge to trade ahead of it, thereby "stealing the information rent" (Stiglitz 2014). If so, short-term traders might be discouraging informed investors. Since prices are used to allocate capital, this might affect macroeconomic outcomes. Empirical work, in some cases, confirms the theory, showing a troubling decrease in measures of price efficiency correlated with proxies of short-term trading (Weller 2018; Lee and Watts 2018). Still, the literature has not inquired directly after the health of the informed investors themselves, nor does it examine how they interact with shorter-term traders. In this paper, we fill the gap, using a dataset with a time span and granularity that enables a study of investors with long-lived information.

Using 11 years of trading data from the Toronto Stock Exchange (TSX), the primary stock exchange in Canada, we identify groups of short-term traders and informed investors empirically. For short-term traders, guided by the relevant theory (Yang and Zhu 2019; Baldauf and Mollner 2020), we look for traders who are likely informed about short-term price movements. We use the criterion that their buying and selling anticipate the short-term (five-second) forward return. In contrast, to identify informed investors, we look for investors who repeatedly build large positions the right way before companies release their quarterly reports. These reports can potentially reveal what informed investors know and hence should motivate them to finish building any position before the revelation. To look for informed investors, we use the criterion that the investor builds large positions in the direction of the unexpected component of the report, does so materially, and does so repeatedly for the sampled stocks and for the 11-year period.

After identifying the trader types, we study their trading activity and their profits. Our first finding is that the informed investors we identify are alive and well throughout the sample—and that the same cannot be said about the short-term traders. The trading and profitability of the informed investors is roughly constant since the financial crisis. While we cannot observe how much information the informed investors acquire, their trading volume and profitability should be positively correlated with this activity. The cumulative profits of the informed investors during periods before large earnings surprises is essentially a straight line from zero in 2005 to three billion CAD in 2015. Their trading volume hovers around five billion shares traded per quarter and does not trend after the crisis. This suggests that changes in market structure (O'Hara and Ye 2011), trading technologies (Hendershott, Jones, and Menkveld 2011), and data sources (Dugast and Foucault 2018) have not diminished informed investors' ability to acquire and profitably trade on information. This result stands in contrast to the concerns in recent theoretical and empirical papers.

If the informed investors are indeed "healthy" through the sample, are there changes in price efficiency? To test the hypothesis, we look at two measures of the informational efficiency of prices: the jump ratio of Weller (2018); and a simplification of the jump ratio that we call relative return. Both measures compare the size of a stock return before an earnings report to the return after the report, attributing a large return after the report to low price informedness. In our results, neither measure when graphed shows any change to the informedness of prices over the sample. In a panel regression, we show that stocks with higher levels of short-term trading are no more or less likely to have efficient prices with respect to earnings-related information. Using the lagged price as an instrument for short-term traders, as in Weller (2018), we still find no relationship between differences in short-term trading across stocks and price informedness. Surprisingly, even when the informed traders trade more, the measures do not display greater information in prices. Thus far, our results suggest that long-term, informed investors continue to profit from private information and that they even can avoid detection.

So, how do informed investors avoid detection, given the growing presence of computerized trading? We explain this using theory on informed trading. The first part of our explanation derives from the Kyle (1985) tradition of models, in which informed traders build their position using multiple small trades to hide their presence. This way, the "signal" of the informed trades is lost in the "noise" of the uninformed trades. The second part is motivated by Kaniel and Liu (2006) and Boulatov and George (2013), who predict that patient informed traders will minimize execution costs by trading using mostly limit orders. The last part of our explanation comes from Collin-Dufresne and Fos (2016), who extend the insight of Kyle to predict that informed investors not only split up their trading over time but also time their trading to coincide with uninformed trading. Specifically, the informed investors time their trading with the volatility of uninformed trading, minimizing price impact and avoiding detection, thus maximizing the information rent.

We give results consistent with these theoretical explanations. First, we show that the informed investors we identify are spreading their trading out widely over the weeks prior to the earnings report, consistent with the Kyle (1985) tradition. This result is in contrast with theory that assumes informed investors are impatient and trade quickly or trade via one main order (Yang and Zhu 2019; Baldauf and Mollner 2020), which would make their trading easier to detect (van Kervel and Menkveld 2018; Korajczyk and Murphy 2019; Hirschey 2020). In order for short-term traders to profit from the information of earnings-informed investors, they would have to be patient and have access to considerable capital, something that is inconsistent with our understanding of how they operate. Making detection harder, a majority of the orders that informed investors use are limit orders, consistent with Kaniel and Liu (2006) and Boulatov and George (2013). As would be implied by the patient use of limit orders, we find the informed investors achieve low price impacts—in fact, the impacts are *negative* at intraday frequencies. Last, by fitting the data to the linear prediction from Collin-Dufresne and Fos (2016), we show the informed investors skillfully time their trade with the volatility of uninformed trading and do so successfully at both intraday and interday frequencies. In fact, the informed time their trading with the uninformed to a greater degree than short-term traders. The informed investors seem to be so good at hiding their trading intentions that, using a VAR, we can show they adversely select the short-term traders, who sell to the informed investors before positive earnings surprises and who buy from the informed investors before negative earnings surprises.

Our paper is related to a set of empirical papers that study informed trading in individual securities. Meulbroek (1992), Cornell and Sirri (1992), Chakravarty and McConnell (1999), Chakravarty (2001), Kacperczyk and Pagnotta (2018) and Shkilko (2019) use insider-trading disclosures as a laboratory to study informed trading. Collin-Dufresne and Fos (2016) use Schedule 13D filings to measure informed trading. Koudijs (2015, 2016) uses a historical setting to measure informed trading in British stocks trading in Amsterdam. Finally, Bushee and Goodman (2007) uses accounting information and institutional trading data to identify potentially informed investors. These studies generally find a response of long-term stock prices to informed trading, often because they study cases of informed traders who have only days or even hours to trade. Our paper enriches this evidence with granular data at a longer-term horizon. We can explain results that lower price impacts occur on days with greater informed activity (Kacperczyk and Pagnotta 2018; Collin-Dufresne and Fos 2015), since the earnings-informed investors achieve negative price impacts.

Our results provide an alternative interpretation of papers that find a negative relationship between shorter-term trading and price efficiency. Lee and Watts (2018) and Weller (2018) suggest that increases in algorithmic trading lead to less information in prices. They argue that informed investors acquire less information because their investment decisions are exploited by algorithms. We can reinterpret the lower efficiency of prices as a result of the sophisticated trading of informed investors themselves, who may be able to trade *better* in the presence of short-term traders, who provide liquidity opportunities. With respect to the theoretical literature, such as Yang and Zhu (2019) and Baldauf and Mollner (2020), we understand these models as studying a case distinct from ours and closer to the empirical papers studying informed investors who must trade quickly, as in a fire sale, or who can trade only once. In our case, the informed investors trade over weeks and weeks. These results call into question the common belief that private-information production is important for price efficiency through the price discovery accomplished by informed trading. Rather, in a world where the informed can disguise their presence, the importance of protecting the privately informed trader is unclear. Ironically, greater "predatory trading" might even be helpful, as it would amplify the price impacts of the informed.

I. Data and methodology

We are grateful to the Toronto Stock Exchange for providing us on-site access to their data. The TSX is Canada's primary equity exchange. It is the best venue to use to study long-term trading in Canada because it has deepest market and is the price leader for trade in Canadian stocks.

The sample contains records of all trades and all top-of-book updates (best bid, best ask, size at bid, size at ask) during 2005–2015 for 307 stocks on the TSX and on Alpha, a TSX-owned trading venue that uses taker-maker pricing. The 307 stocks were chosen for their membership on the TSX Composite stock index in 2006 and in 2014. The trade records have fields for ticker, price, quantity, date, millisecond or microsecond timestamp, and anonymous trader identifiers. The top-of-book records have fields for ticker, price, quantity, date, and millisecond or microsecond timestamp. We augment the ticker field by merging it with

CUSIP codes, which enables us to unite data from companies that change their ticker during the sample period.

The trader identifier codes we use were created by TSX exchange members, who assign the codes to individual approved traders. Approved traders are human beings who manage trading flows for a company and its clients. An approved trader may have one or more trader identifier codes assigned, and the codes can be used to aggregate the orders of any number of the trader's clients, with the important exception of clients who have been given direct market access by the exchange member. Canadian regulation requires clients who have direct market access to have their own unique trader identifiers, though they may still have more than one. Otherwise, the reason and rationale for the assignment of an identifier code varies. For example, codes are used to group or separate orders by nature of handling (manual vs. algorithmic), by trading strategy, by business channel or customer type, or even by client. Anecdotally, the larger the client, the more likely an exchange member might assign it a unique identifier.

We merge the TSX data with data from Bloomberg consisting of quarterly earnings reports for all companies in the sample. These reports have fields of ticker, date, dividends per share, and the median analyst expectation of earnings per share. Earnings per share conveys important fundamental information about a company as it is the quarterly profit. Stock analysts attempt to predict the earnings, and the difference between the median analyst expectation and the realized earnings per share is computed as the earnings surprise. Earnings announcements are the most important form of a firms' communication with capital markets and are associated with large permanent changes in stock prices (Kothari 2001).

TABLE I HERE

The result is an 11-year long sample of intraday trading, intraday pricing, and earnings information. A long sample is advantageous for our study because we will use the earnings reports to identify trader types, and earnings reports come out only once a quarter. The long sample also affords us the ability to document changes in trading and profitability over time. Since we will use only the top half of earnings reports by the magnitude of the surprise, the sample can provide a maximum of $307 * 11 * 2 = 6,754$ material earnings surprises useful to identify the informedness of the accounts. In fact, we have only $N = 4,322$ material earnings surprises, since many stocks enter or exit the sample partway. Table I gives quarterly averaged metrics on the stocks, accounts, and events in the TSX sample during four epochs: 2005–2006, 2007–2009, 2010–2012, and 2013–2015. In summary, the stocks in the sample are liquid, which is expected of stocks in the main index. The bid-ask spread ranges from 33 to 18 basis points, and mean price volatility ranges from 31 to 52. As is typical on modern electronic markets, the median trade size is small, approaching 100 shares in the last epoch. Dividing the average number of trades by the number of minutes in a quarter, there are about five trades per minute per stock.

TABLE II HERE

Table II explores earnings announcements, split by their direction (positive vs. negative) and by their materiality (material vs. immaterial). *Material surprises* are earnings surprises for a stock that are ranked in the top half of positive surprises by the magnitude of the surprise or in the bottom half of negative surprises, excluding zero-magnitude surprises. Consistent with the definition of materiality, the median surprise in earnings per share for positive material surprises is much greater than that of positive immaterial surprises: 0.07 versus 0.02. Positive material surprises also have a larger interquartile range, 0.04 to 0.15 versus 0.01 to 0.04. Negative surprises are similar. There is evidence of post-earnings price drift in the sample; see Martineau (2019) for analysis. The seven-day return is larger than the two-day return for material surprises.

We confirm that the earnings surprises are "surprising," since the identification regime depends on correlating investor behaviour with releases of novel information. In Table III, viewable in the appendix, we regress the materiality of earnings surprises on the lagged size of the surprise and the lagged materiality, using logit with and without fixed effects. We also regress the surprise size (both signed and unsigned) on lagged size and lagged materiality (both signed and unsigned). The earnings surprises are not substantially predictable or autocorrelated, with pseudo-R^2 and R^2 coefficients ranging from 0.001 to 0.014. In the case with the highest R^2, 0.014, the absolute value of size is mildly predictable using the lagged absolute value of size. In other words, large surprises (good or bad) have a mild tendency to be followed by large surprises (good or bad). But when returning the sign to the surprise, the signed surprise is unpredictable, with an R^2 of 0.001. This makes profiting from the information in previous information releases difficult.

FIGURE 1 HERE

Last, Figure 1 shows weekly average returns over all stocks during the eight weeks before, week of, and four weeks after the material earnings surprises. It shows the returns for positive and negative surprises separately, to check for symmetry. The surprises in the data are again "surprising," in that prices only weakly anticipate the news. Particularly for negative surprises, the eight-week cumulative return is almost flat before the earnings announcement. Even for positive surprises, where the weekly returns during the second and third week before the announcement are on the order of the post-period, the average return during the week before the surprise is negative.

The figure presages our later results. Prices do not appear to anticipate the earnings surprises: 80% of the price discovery happens the week of the earnings report and the week after. For the positive surprises, the prices drift in the opposite direction of the surprise on the weeks nearest to the reports. If there are traders in the sample who are privately informed about the stocks, they do their best to get their trades in before the earnings reports, which risks revealing part of their private information. This figure shows that informed traders are apparently able to trade without moving the price substantially in the direction of the surprise. Earnings surprises are some of the most regular events on stock markets, so if anything should coordinate the revelation of information, these should. The lack of a strong price drift in the direction of the earnings surprise poses questions about how much price discovery is happening in financial markets.

A. Identification of informedness types

We flag the accounts by stock and by quarter as *earnings-informed investors* and as *five-second traders*. Accounts are flagged by stock and quarter, so the same account can be flagged as different types for different stocks or different quarters. The same account can be flagged as both for the same stock and quarter; this happens for only 4% of volumes.

The first flag we assign is the five-second trader flag. This flag is intended to capture traders who make money by being among the first to react to new public information; we have in mind strategies such as statistical arbitrage, trading on news releases, trading against temporary deviations in the price, and trading on short-term momentum. For us, five-second trading would include aggressive algorithmic trading but is not limited to this capital-light, proprietary business. We identify five-second traders by regressing the five-second log return on an account's unexplained, lagged, five-second net trading flow,

$$r_{s,t} = \beta \widetilde{flow}_{s,i,t-1} + \epsilon_{s,i,t}, \tag{1}$$

fit once a quarter, where $r_{s,t}$ is the log return for stock s over the five-second increment t, $flow_{s,i,t-1}$ is the shares bought less shares sold by account i in stock s in the previous increment, and $\widetilde{flow}_{s,i,t}$ is the residual after removing the predicted $\widehat{flow}$ given by the fitted model,

$$flow_{s,i,t} = \theta \sum_{s \in -s} flow_{s,i,t} + \sum_{\tau=1}^{3} \delta_\tau flow_{s,i,t-\tau} + \sum_{\tau=1}^{3} \gamma_\tau r_{s,i,t-\tau}, \tag{2}$$

where $-s$ is the set of all stocks other than stock s. The model identifies a predictable component of the trading flow using the account's contemporaneous trading in other stocks, the account's lagged trading in the same stock, and lagged returns, so as to identify users whose trades are surprises with respect to market conditions. Accounts whose trading flows are easy to explain in this manner are not likely trading on short-term information, so we subtract the predictable trading flows. Finally, accounts with a positive estimate of β and a t-statistic of at least two are flagged as short-term traders for the stock and quarter.

The class of short-term traders is unlikely to contain market makers or traders otherwise using a passive strategy akin to market making. Our identification requires the return to move in favour of the short-term trader, whereas for market makers the return moves against the trader because of adverse selection. In other words, this is not a sample of intraday market makers who are passively quoting bids and asks. We highlight this aspect of the identification because we will later argue that the short-term traders can be interpreted as supplying liquidity. This interpretation is not "by construction" due to a market-making strategy—rather, the identification targets aggressive trading strategies of the sort considered by Yang and Zhu (2019) and van Kervel and Menkveld (2018). To confirm we reach the target, in the summary statistics, we show the short-term traders use aggressive orders around 70% of the time.

The second flag we assign is the earnings-informed flag. This flag is intended to capture traders who have an informationally advantaged view on a company's value that is long in term but that is eroded, at least in part, by news released in a company's earnings reports. The kind of view we have in mind would be acquired through fundamental research on the company or on its industry, for example via private monitoring or by commissioning studies, or through a superior capacity to process or understand information that is already public, for example via a team of analysts. Successive earnings reports must erode such an informational advantage, progressively or entirely, because all profit-relevant outcomes are eventually revealed in earnings reports, ending with the end of the company.

We flag accounts as earnings informed by asking that their unexplained positions accrued before earnings surprises satisfy a *materiality* criterion *repeatedly* through the sample. We start by computing the unexplained positions. These are the sums of the daily residual trading flow accrued between earnings reports. The residual flow is the flow remaining after subtracting the predicted $\widehat{flow}_{s,i,d}$ given by the fitted model,

$$flow_{s,i,d} = \theta \sum_{s \in -s} flow_{s,i,d} + \sum_{\Delta=1}^{3} \delta_\Delta flow_{s,i,d-\Delta} + \delta_4 \sum_{\Delta=4}^{120} flow_{s,i,d-\Delta} + \sum_{\Delta=1}^{3} \gamma_\Delta r_{s,i,t-\Delta} \tag{3}$$

fit once per inter-earnings-report period, where all terms and indices are as in equation 2 and, in addition, d is the day. The model identifies a predictable component of the trading flow using the account's same-day trading in other stocks, the account's lagged trading in the same stock, the account's recently (120-day) acquired inventory in the same stock, and lagged returns, so as to preference users whose trades are surprises with respect to market conditions and own inventory. Accounts whose trading flows are easy to explain in this manner are not likely trading on long-term information, so we subtract the predictable trading flows. This alleviates concerns that our measure simply identifies large investors randomly buying or selling a security in a given quarter.

Next, we sum the unexplained positions between earnings reports. With the summed positions, we count the number of times they satisfy a *materiality criterion*. The criterion is satisfied whenever the account (a) builds a position size that is material (b) before a material earnings surprise. To remind, material surprises are earnings announcements for which the absolute magnitude of the surprise was greater in absolute value than the median for positive and negative surprises. An unexplained position is of *material size* if it is

1. in the direction of the surprise, meaning a long position before positive surprises and a "short" position before negative surprises;
2. greater than 10,000 shares in size; and
3. in the top quartile of accounts by number of shares among accounts that already satisfy conditions 1 and 2.

For the *repeatedly* criterion, we total the number of times an account satisfies materiality. This total can be thought of as an account's "materiality points," and we total up the number of points by account. Accounts that are in the top 10% satisfy the *repeatedly* criterion. Finally, we flag an account as earnings-informed (informed investor) if it satisfies both criteria *and* it is a quarter in which it is trading in the direction of a material surprise or one of the two adjacent quarters.

B. Summary information on account types

FIGURE 2 HERE

Figure 2 shows the history of volumes by informedness type: earnings-informed investor, five-second trader, and uninformed. On average over time, about a third of trading volume is informed in some sense. It also seems natural that the majority of trading would be uninformed, due to a lemons problems that would otherwise prohibit trade. Of the informed trading, relatively little is both earnings-informed investor and short-term trader. Only 4% of volume is flagged as both types of informed. This is reassuring, as it means the two types of trader are relatively distinct.

TABLE IV HERE

Table IV gives summary statistics for the behaviour of the three account groupings, averaged on an account-by-account basis. The behavioural statistics imply differences in trading strategy confirming that the user types are distinct. First, accounts in the two informed groupings trade more than accounts in the uninformed grouping—more than twice as much as the uninformed—and both in volume and in trade count. Of the three types, the five-second-trader accounts trade the most. However, when ranked by median trade size, it is the uninformed who have the largest trade sizes—almost twice as large as the others—and it is the earnings-informed who have the smallest. This implies that the uninformed trade in relatively large, block-sized quantities, whereas short-term traders trade large volumes in relatively small increments, and earnings informed are in the middle. This is consistent with Barclay and Warner (1993), who suggest that informed traders can best hide their trading using medium trade sizes.

Perhaps the most striking difference is in aggressiveness. While five-second traders trade mostly using market orders (69% aggressive), paying to cross the spread and achieve immediacy, the earnings informed trade mostly using limit orders (43% aggressive), passively building up their positions. That the shorter-term traders use aggressive order-types is consistent with the prediction of Yang and Zhu (2019) and Baldauf and Mollner (2020). This implies very different valuations of immediacy. Five-second traders behave as if they need to trade now, while earnings-informed traders behave more patiently, as if they trade when convenient. This is consistent with five-second traders racing to trade on short-lived public information, whereas earnings-informed investors can be more patient as their information is long-lived and private.

Finally, earnings-informed investors tend to trade repeatedly in the same direction as their previous trades. To measure this, we compute *directionality* as an account's end-of-period position divided by the trading volume during the period. Earnings-informed investors use 77% of their volume to construct their end-of-day positions (and net the rest), compared to only 56% for five-second traders. They are also more directional on the average quarter (including quarters with both material and immaterial surprises), using 40% of their volume to construct their end-of-quarter positions (and netting the rest), compared to only 26% for five-second traders. Since the five-second traders on average end a day or a quarter with a significant position, this group is unlikely to be populated solely by capital-light, low-latency algorithmic traders.

II. Informed investing and price discovery

We begin by addressing the concerns motivated by recent theory: that informed traders are discouraged by shorter-term traders, who are therefore hindering price discovery. To address the concerns, we first inquire after the health of the earnings-informed investors and ask whether they continue to make profits despite the presence of the five-second traders. Then, we ask if earnings-informed investors contribute to price discovery in the same way given the presence of the five-second traders. We find, in summary, that earnings-informed investors are alive and well in terms of volumes and profits, yet it is difficult to detect their contribution to price discovery. This is true regardless of the quantity of five-second trading. It seems that earnings-informed traders are able to disguise their trading.

To study profits by trader type, we compute a profit measure on positions built during the earnings cycle. We accumulate the positions of each trader type starting from eight weeks before an earnings report and ending one week after an earnings report (to allow for post-earnings-report drift). We value the positions at the stock price one week after the earnings report, less the prices paid for the position, plus any trading profits or losses made along the way. We call this the profit from the earnings cycle. Figure 3 shows these profits for the three trader types, accumulated from 2005 to 2015. We accumulate the profits *separately* for material earnings surprises and for <u>im</u>material earnings surprises, as earnings-informed trader types should profit more when there is more value to private information. Note that trading profits are zero-sum.

FIGURE 3 HERE

In Panel A, the earnings-informed investors take profits from material earnings surprises at a roughly constant rate, accumulating three billion CAD during 2005–2015. Though this is consistent with a hypothesis of private and long-lived information, part of this result can be by construction. The earnings-informed investors were selected because their trading is positively correlated with earnings surprises that are positively correlated with returns. What is not by construction that they profit at a constant rate through the sample. If informed traders were put under increasing pressure by short-term traders, their success might have been concentrated in the early part of the sample, before the arrival of changes to the market structure.[1] Instead, the roughly constant growth in profits is consistent with no change to trading costs due to shorter-term traders using back-running or latency-arbitrage tactics. If there were greater costs, one should expect some reduction to quantity trading and

[1] To check whether the earnings informed are simply lucky or random, we perform an out-of-sample check of whether their volumes predict the direction of the <u>im</u>material surprises. For these, they trade the "right" way 64% of the time.

therefore profits. Interestingly, it is the five-second traders who experience a deteriorating performance.

In Panel B, the reverse holds. During immaterial earnings surprises, it is the five-second traders who take profits from the market, though only until 2011. They accumulate a maximum of 600 million CAD in profit, whereas the earnings-informed take a loss from immaterial surprises, though the loss ceases after 2011. One explanation is that, absent a source of fundamental adverse selection, it is the five-second traders who have the informational advantage. They are specialized in reacting to changes in public information. Still, after 2011, the five-second traders begin recording small losses. This is consistent with evidence from Tabb (2017) that profits from high-speed trading in equity markets tapered severely in 2010 and 2011. It could be the five-second traders are no longer profiting by trading at a single market and instead make their profits trading between markets, as in Brogaard, Hendershott, and Riordan (2019) and Menkveld (2013).

Next, we study price discovery, starting with simple measures of price impact. In principle, if informed traders are helping the market discover a better price, their decision to buy should raise the price, and their decision to sell should lower the price. This is the process of *tâtonnement*. The price impact of each trade should impound information into prices, trade by trade.

FIGURE 4 HERE

Figure 4 shows average intraday price impacts for the three groups. *Price impact* is the volume-weighted average percent change in the midquote price from the moment of trade to some time increment after a trade: one second, five seconds, one minute, five minutes, 30 minutes, and time to end of day. To be clear, we are not using the order type (market or limit) to compute price impact; instead, we use all trading regardless of order type. Brogaard et al. (2019) show that both aggressive and passive orders are associated with price impact, motivating our use of all orders rather than just aggressive orders. Earnings-informed investors are in blue; short-term traders are in orange, and uninformed are in gray.

Some aspects of the figure obtain by definition. For example, it is by definition that the five-second price impacts of the five-second traders would be positive, since the type was selected by this criterion. The price impact of five-second trading is statistically strong—it can be seen after thirty minutes (9.5 basis points) and through to the end of day (6.5 basis points). Another aspect that obtains mostly by definition is that uninformed traders have negative price impact, as someone has to be trading with the five-second traders.

But, strikingly, earnings-informed investors also have negative price impacts, despite being selected for informedness. To confirm the result, we compute the price impacts of earnings-informed investors for quarters during which there was a material earnings surprise and for quarters during which the surprise was not material. The price impact of earnings-informed investors is higher (less negative) when they are trading before a material surprise, but it is still always negative. This implies they are successful in trading patiently and mostly passively; that is, when the flow of trade is against their limit orders. Their trading strategy must be a patient one, since using limit orders means waiting for others to show trading interest. The trading does not apparently contribute to price discovery, as earnings-informed investors do not move prices in the direction in which they are trading—quite the reverse.

This is a counterintuitive result, since much of the theory on informed trading since Kyle (1985) has assumed that privately informed investors impound their information into prices by impacting the price through trading. Here, we see informed investors using a strategy of patience to avoid price impact, which saves them costs and disguises their trading. The behaviour is consistent with the predictions of Kaniel and Liu (2006), in which investors with long-run information trade patiently with limit orders. Still, unwilling to give up on the intuition of price impact, we try instead a longer-term and more aggregate measure of price discovery that might therefore detect the presence of informed investors. Instead of a trade-by-trade measure, we use daily net buying to identify price impact in the empirical framework of Kyle's lambda, the coefficient of regression in the linear model,

$$r_t = \lambda flow_{i,t} + \epsilon_{i,t}, \tag{4}$$

in which t is the day, r_t is the return from the end-of-day $t-1$ to the end-of-day t, and $flow_t$ is the trading flow (buy volume less sell volume) of the account type i on day t.

FIGURE 5 HERE

Figure 5 shows the quarterly average lambda over the stocks in the sample computed by trader type. As in Figure 4, the short-term traders have the largest influence on the daily return. Their lambda is around one basis point per $100K net flow in the first half of the sample and increases during the sample to around two basis points. As in Figure 4, the earnings informed traders again have a low price impact, hovering around zero through the sample period. Consistent with the above, the uninformed have a negative price impact as well.

Unable to detect the presence of informed investors using measures of price impact—neither the trade-by-trade impact, nor the daily Kyle's Lambda—we move to see whether we can measure a change in the outcome of price discovery, which is price efficiency. We compute two aggregate efficiency measures: the jump ratio of Weller (2018), and a simplification of the jump ratio that we call relative return. The jump ratio is the ratio of two cumulative abnormal return (CAR) measures of a stock in which the returns are computed during time periods around an earnings report. We use the paper's parameterization: The numerator is the CAR from one day before an earnings report to three days after; the denominator is the CAR from 23 days before the earnings report to one day before:

$$jumpRatio_{i,t} = \frac{CAR_{t+3,t-1}}{CAR_{t-23,t-1}}.$$

The index i is for the stock, and the index t is for the day. The idea here is to compare the "normal" price change of a stock to the impact of an earnings report. On average, if informed investors are making prices more efficient, then earnings reports should impact prices less, so the jump ratio should go down. As a sanity check, we also compute a simplification of the jump ratio without the controls and in which we lengthen the lags of the returns and overlap them:

$$relativeReturn_{i,t} = \frac{r_{i,t+7,t-1}}{r_{i,t+7,t-56}}.$$

The indices i and t have the same meaning as in the jump ratio. The "relative return," is simply a ratio of the one-week stock return after the earnings report to the eight-week stock return before the report to one week after:

Figure 6 shows the time-series average of both measures with the confidence intervals.

FIGURE 6 HERE

There appears to be no time series movement in either measure in the sample. Both measures hover around zero during 2005–2015. There are some spikes in the confidence interval during financial events, such as the 2008–2009 financial crisis and the debt crises. Otherwise neither series is much different from a flat line.

To explore the cross-section as in Lee and Watts (2018) and Weller (2018), we run a panel regression on both measures of price efficiency and attempt to explain it using signed and unsigned trading volumes of the trader types, controls, and stock and quarter fixed effects. We try the regression once for all earnings reports and once for only reports with material earnings surprises. For these two regressions, we used signed trading volume as we have trader types and as the jump ratio and relative return are also signed. We run a third regression, again for all reports, but using the IV setup from Weller (2018) in which the unsigned trading volume of the shorter-term traders is instrumented with the two-week lagged, two-week averaged price.

The controls are: the two-week-lagged, two-week averaged natural logs of *market capi-*

talization, the stock price times the number of shares outstanding; *relative bid-ask spread*, the bid-ask spread in basis points; *price*, the stock price; and *volatility*, the rolling standard deviation of the stock return. We exclude quarters with no record of an earnings report. Standard errors are clustered by stock and quarter.

TABLE V HERE

Table V reports the results of the regressions. In no regression does the signed or unsigned trading volume for any trader type exhibit a statistical relationship with the price-efficiency metric. The controls also bear no relationships, except for the control of volatility in the case of the jump ratio. The within-R^2 are poor, ranging from -0.001 to 0.002. One interpretation of the result is that earnings-informed investors are able to disguise their trading so well that these efficiency metrics cannot detect them. Still, it could also be that the measures are poor measures of price efficiency.

In summary, we have identified an account type, the earnings-informed investor, that is consistently profitable throughout the sample, accruing three billion CAD from material earnings surprises during the sample period. Its trading flows are large and are hard to explain except as the result of private information, yet the flows do not appear to contribute to price discovery, nor do they appear to improve price efficiency. This raises a natural question: If the earnings-informed investors are privately informed and profiting from private information, how are they successfully avoiding the hypothesized predatory behaviour of shorter-term traders—and, moreover, how are earnings-informed investors trading so well that price-impact and price-efficiency metrics do not detect their presence?

III. How the earnings-informed investors trade

To understand how earnings-informed traders achieve such low price impacts, we look at the choices they make when trading. We study which stocks they tend to trade, the speed of their trading, and the sorts of orders they use. What we find is that they are hard to predict. They do not concentrate on any obvious category of stock, and they trade at a roughly constant rate during the weeks before earnings announcements. They mostly use limit orders (being around 57% passive), and they do not adjust their order choice much (at least, against the common explanatory variables in market microstructure). The invariance of their rate of trading and of their order choice can be interpreted as a sign of the generality of their trading strategy, and it also means they would be hard to manipulate. We then study whether the earnings-informed react to the trading of five-second traders, and, again, the earnings-informed do not adjust their trading in reaction—much the reverse. It is the five-second traders who are reactive (and we find it is to their detriment). Last, to gain more traction, we switch to a more "structural" approach. We have more success explaining the trading of earnings-informed investors using the framework of Collin-Dufresne and Fos (2016). As the authors predict, the way informed investors trade is by timing their trading during noisy periods of uninformed trading.

A. The earnings-informed investors' choice of stocks

First, we study stock choice. Table VI gives the results of a panel regression of quarterly volumes by trader type on quarterly stock characteristics, including a set of industry dummies. For comparison, we study the stock choices of both the earnings-informed and the five-second traders. The stock characteristics are: *uninformed volume*, the trading volume of uninformed traders in millions of shares; log *market capitalization*, the stock price times the number of shares outstanding; log *relative bid-ask spread*, the bid-ask spread in basis points; and log *volatility*, the rolling standard deviation of the stock return. We also regress on nine industry dummies as listed. All specifications use quarter fixed effects; the second specifications use stock fixed effects. Standard errors are clustered by stock and quarter.

TABLE VI HERE

The industry dummies do a poor job of explaining stock choice for both trader types, neither of whom appear to specialize in particular major industries. Rather, for both of the informed trader types (both long- and short-term informed), the strongest explainer of stock choice is the uninformed trading volume—uninformed volume begets informed volume. For earnings-informed investors, trading volumes are 14% the size of uninformed volumes; for five-second traders, trading volumes are 15% those of the uninformed. This implies, as in Figure 2, that the traders identified as informed (either about earnings reports or about short-term returns) comprise around a third of volumes. In our interpretation, stocks in which there are opportunities to trade without adverse selection are those in which traders who are informed like to trade.

Comparing the two groups, the volumes of five-second traders are easier to explain than those of the earnings-informed investors. After the fixed effects for stocks are applied, the within-R^2 of the five-second traders is 0.23 compared to just 0.08 for the earnings-informed. The five-second traders have a statistically significant preference for high volatility and high market capitalization, even after the application of stock fixed effects. In contrast, after applying fixed effects, the earnings-informed have no statistically significant explainers other than uninformed volume. While we can claim to explain some of the stock choice of five-second traders—they prefer high volatility and large-size stocks, much like algorithmic and high frequency traders (Hendershott et al. 2011; Brogaard, Hendershott, and Riordan 2014)—the fixed effects do almost all of the explaining for the earnings-informed. While this is a negative finding and is precarious to interpret, the lack of a result is consistent with an unobserved true motivator for earnings-informed investing, namely the discovery of long-lasting, fundamental information for certain stocks and in certain quarters.

B. The earnings-informed investors' choice of immediacy

Having studied the stock choice of the earnings-informed, we move to how quickly they trade. To do so, we examine the average weekly trading flows of the three types of trader around the information events in our sample, the earnings surprises. To focus on how earnings-informed investors behave, we focus particularly on the material earnings surprises—to remind, material surprises are earnings announcements for which the absolute magnitude of the surprise was greater in absolute value than the median for positive and negative surprises. If an earnings-informed investor knows a stock is significantly mispriced, they also know the next earnings report is likely to reveal at least part of the mispricing. Thus, it is before the material earnings announcements that earnings-informed investors need to finish trade to secure their informational rents.

FIGURE 7 HERE

Figure 7 shows the weekly average cash flows of the three account types during the eight weeks before, the week of, and the four weeks after the material earnings surprises. *Cash flows* are the sums of the dollar values of all purchases (signed positive) less all sales (signed negative) during the week. Earnings-informed investors are in blue; short-term traders are in orange, and uninformed in gray. The dollar values of the three trading flows from the three accounts must sum to zero, as there is a buyer for every seller. For negative earnings surprises, we multiply the cash flows by -1 before averaging them (so that all surprises are signed "positive" for ease of comparison).

Some aspects of the figure obtain by definition. For example, it is by definition that the earnings-informed investors are net buyers before the material surprises, as they were selected to do this. What is not by definition is that the cash flows of earnings-informed investors have consistent signs and magnitudes for every week around the earnings surprise. On average, they are spreading out their trading roughly evenly: four million CAD a week during the eight preceding weeks. This is consistent with a patient approach to building a position as in the Kyle (1985) tradition of monopolistic informed traders, in which the trader maximizes the information rent by amassing a position evenly over time. Last, the earnings-informed investors taper off their trading after the surprise, confirming that the surprises diminish the value of the information by making public some of the private information.

Strikingly, and also not by design, the cash flows of the five-second traders and of the uninformed traders are always in the opposite directions of the earnings announcements. This means they are both losing money with respect to material earnings surprises. While it is not surprising that the uninformed traders would be trading in the opposite direction, as someone has to be trading with the earnings-informed, it is more interesting that traders who are informed in the short-term sense would be so uninformed in the long-term. Traders who are apparently conducting short-term arbitrage *intra*day are trading, on net, the "wrong" way *inter*day.

C. *The earnings-informed investors' choice of order type*

Third, we study the order choice of earnings-informed investors. Table IV has already shown that the trader types differ by their order choice: Earnings-informed investors were 57% passive (using limit orders), whereas five-second traders were 69% aggressive (using marketable orders). This is consistent with Kaniel and Liu (2006) and O'Hara and Ye (2011), in which long-term investors use passive orders to disguise their trading intentions and to save on transactions costs. In Table VII, we go further and attempt to explain the percentage of volumes traded passively by the earnings-informed.

We try a variety of explanatory variables: *volume share of five-second traders*, the percent of a stock's trading volume in which a five-second trader was a counterparty; *volume share of uninformed traders*, the percent of a stock's trading volume in which an uninformed trader was a counterparty; *signed return*, the daily return multiplied by the sign of the earnings surprise for the quarter (-1 for negative surprises); *trades*, the daily number of trades for the stock; *trading volume*, the daily number of shares traded for the stock; and *illiquidity*, the relative spread implied by the first principal component of stock relative spread, inside depth, five-minute per-trade price impact, and 30-minute Kyle's lambda. The *relative bid-bid spread* is the bid-ask spread divided by the midquote; the *inside depth* is the sum of limit orders available at the bid and ask; the *five-minute per-trade price impact* is the price impact computed the usual way as the aggressive-trade price impact, the difference of the midquote contemporaneous to a trade and the midquote five minutes later, signed by the sign of the aggressive trade (positive for buys, negative for sells); and the *30-minute Kyles Lambda*, the lambda computed for 30-minute returns and for aggressive order flow. The third and fourth specifications use stock and quarter fixed effects; standard errors are clustered by stock and quarter.

TABLE VII HERE

The order choice of the earnings-informed investors is remarkably invariant to the explanatory variables. The sole variable that consistently explains order choice is the volume share of the five-second traders. This is interesting—reading Kaniel and Liu (2006), one might expect the earnings-informed to adjust their order composition where there is higher volume, higher volatility, or greater liquidity. High trading volume should make it more likely for a limit order to fill; high volatility should make it more likely the price would cross any given limit order; and greater liquidity should make it easier to trade passively with immediacy. Yet here, the order choice of earnings-informed investors is basically invariant to these and to the other usual explanatory variables of microstructure, with a scant improvement of 0.01 in R^2 attributable to the six variables other than the volume share of the five-second traders.

Given these results, it makes sense that the earnings-informed investors have negative price impacts. They have a relatively invariant policy of using mostly limit orders. Limit orders tend to be filled when the price is moving against them, so earnings-informed investors will tend to buy when the price is falling and sell when the price is rising, achieving a negative price impact. The one wrinkle is the strongest explanatory variable, the five-second volume share, which increases with the passive order choice at a rate of four percentage points of passive order share per 10 percentage points of five-second volume share.

A natural interpretation would be that the earnings-informed are mildly adjusting their order composition to a trading environment with more five-second trading. The interpretation does not have a clear implication for whether earnings-informed investors are predator or prey. On the one hand, it could be that earnings-informed investors switch to passive orders to hide from back-running traders, thereby accepting less immediacy and less volume. Still, on the other hand, it could be the increased trading activity of specifically five-second trading is beneficial for passive trading, as it creates natural opportunities to have an order filled. In this interpretation, the five-second traders could be construed as liquidity providers. The next subsection tests for both of these two implications.

D. The earnings-informed investors' reaction to five-second traders

To test for whether five-second traders are acting more as back-runners or more as liquidity suppliers, we study the dynamic interaction between the earnings-informed investors and the five-second traders. To do so, we use a vector autoregression to measure the influence of informed traders on the price. We fit the VAR model at a half-hour frequency,

$$\begin{bmatrix} r_t \\ flow_{EI,t} \\ flow_{5s,t} \end{bmatrix} = \sum_{\tau=1}^{3} A_\tau \begin{bmatrix} r_{t-\tau} \\ flow_{EI,t-\tau} \\ flow_{5s,t-\tau} \end{bmatrix} + \begin{bmatrix} c_r \\ c_{EI} \\ c_{5s} \end{bmatrix} + \begin{bmatrix} \epsilon_{r,t} \\ \epsilon_{EI,t} \\ \epsilon_{5s,t} \end{bmatrix}, \tag{5}$$

in which t is the half-hour, τ is a lag, r_t is the log return, $flow_{EI,t}$ is the net buying flow (buy volume less sell volume) of earnings-informed investors, and $flow_{5s,t}$ is net buying flow of five-second traders. It is unnecessary to include the uninformed flow as it is a linear combination of the other two.

We fit the VAR once per stock, save the fit coefficients, and generate the impulse response functions (IRFs). Taking the coefficients, we average them and compute the standard error of the average. We report this result in Table VIII, which is viewable in the appendix. The VAR coefficients are reported for transparency, but it is easier to interpret the VARs using the IRFs. Similar to the above, we average the impulse coefficients and compute the standard error of the averages. We begin the analysis with the response of the return to the two trading flows. Figure 8 gives the averaged IRF of the return and the confidence interval implied by the standard error.

FIGURE 8 HERE

Here, we are finally able to discern a positive price impact from earnings-informed investing. Consistent with the previous results, the influence of earnings-informed investing on the price is subtle, whereas the influence of five-second traders is immediate and strong. In Panel A, the response of the return to an impulse of $100K of earnings-informed investing is statistically insignificant at the first and third lags, becoming more clearly significant after two hours. Even when the response is statistically significant, it is barely so. Still, the VAR does the job of identifying the earnings-informed investors' influence on prices, whereas the intraday price impacts and Kyle's Lambdas did not. The only way we have been able to detect a price response to informed trading has been using an autoregressive framework in which the price impact of a trade can build on itself. The response converges to just under 2 basis points. In Panel B, the response to five-second traders is immediate and clearly significant, converging to just over 4 basis points.

FIGURE 9 HERE

Figure 9 gives the averaged IRF of each type's trading flow to the other's flow. In Panel A, the five-second traders respond with strong significance to an impulse of $100K of earnings-informed trading. Strikingly, they respond by trading in the opposite direction. Since the earnings-informed investors are often trading in the direction of a future information release, the five-second traders, who are trading in the other direction, are supplying liquidity in the wake of informed trade. As we noted in the section on identification, this is not due to a passive market-making strategy, since the identification criteria ask for the return to move in the favour of the short-term traders (whereas, for market makers, it would move against the trader due to adverse selection). The five-second traders' response converges on about

$7,000 of trading in the opposite direction after a $100K impulse. This is not only consistent with a strategy of liquidity supply, but it is inconsistent with a strategy of back-running as in Yang and Zhu (2019), in which the back-runner trades "with the wind" of the informed investor and not against it. The rejection of the hypothesis of back-running is consistent with Chen and Garriott (2019).

In contrast, in Panel B, earnings-informed investors do not react to an impulse of five-second trading. The response is statistically insignificant and is near $0 in size. While the five-second traders react to the trading decisions of the earnings-informed, it appears the earnings-informed act like the trading flows of the five-second traders do not exist. Five-second traders may be inframarginal to the choices of trading magnitude and direction of the earnings informed.

This subsection and the preceding three subsections gave us a number of results that we found unexpected. The five-second traders have strong, immediate, and large-sized price impacts, and price impacts are often taken as a sign of private information. Yet, in this section, we find the traders with the large price impacts are the ones being adversely selected, as they trade against the apparently privately informed investors in the sample. In contrast, the earnings-informed investors influence the price weakly, slowly, and subtly, with negative price impacts, which are often taken as a sign of being uninformed. Far from being afflicted by the shorter-term traders, the earnings-informed investors do not change the magnitude or direction of their trading—they only trade a touch more passively. In our interpretation, this does not look like predatory trading. Rather than being the prey, the earnings-informed investors appear to be the apex predators in this ecology.

E. The earnings-informed investors' choice of timing

Last, we try a more "structural" approach. While it is clear that the earnings-informed investors spread out their trading, as in Kyle (1985), and while it is clear they have a policy of trading mostly passively, as in Kaniel and Liu (2006), we have not produced an explanation for how they trade. To do so, we draw from a recent extension of the workhorse Kyle (1985) model, Collin-Dufresne and Fos (2016), which produces a linear prediction for informed trading as a function of a state variable. We compute the state variable and fit the prediction to our data.

Collin-Dufresne and Fos (2016) allow for uninformed trading to have a stochastic volatility σ_t, in which σ_t may have history-dependent growth and volatility terms. This makes the expected trading volume of uninformed traders vary, and it also makes the volumes partially predictable. The volume displays periodic variation that can be "timed," and there is incentive to time it, because price impact is lower when uninformed traders are active. The outcome is that informed traders use σ_t as the state variable expressing when to optimally time their trade, and informed traders scale up or down their activity with the magnitude of the variable. Guided by theory, we therefore compute a "CDF" factor directly using the uninformed trades, and ask which of the trader types (earnings-informed or five-second) is timing their trade more by the CDF factor. We find earnings-informed investors are timing their trade better by the criterion of the factor.

Our analysis doubles as a more full-fledged test of the CDF model than Collin-Dufresne and Fos (2015), in which the authors use insider trading to verify the model's predictions. Since insiders have peculiar incentives, it is possible those results could apply only narrowly to insiders. We confirm the model predictions using a more general notion of informed trading. We also show there is more work to do in this tradition of models, as the CDF factor does not exhaust the timing behaviour of informed traders. Moreover, in the internet appendix, we show our uninformed traders themselves appear to be timing liquidity. Assuming our "uninformed" traders are indeed uninformed, this enables us to recommend the assumption that they are noise traders to be one day relaxed.

Following CDF, we compute the *CDF factor* as $\sigma(\Delta flow_{uninf,t})$, the standard deviation of the first-difference of the uninformed signed trading flow. To test for timing on multiple time horizons, we compute the factor for all of the increments studied in our paper: half-hour, daily, and weekly increments. For the half-hour increment, we compute the variance of the five-minute signed trading flows within the half-hour; for the daily increment, we compute the variance of half-hour flows within the day; and for the weekly increment, we compute the variance of daily flows within the week. This shows our results are not dependent on the interval of computation and that informed traders are able to time at intraday, intraweek and intraquarter frequencies.

To test whether the CDF factor is the sole explanatory variable of the timing of informed volumes, we follow the spirit of Collin-Dufresne and Fos (2016) and include metrics for illiquidity, return volatility, and two measures of uninformed volume other than the standard deviation of its first difference. Given the spirit of the model, we expect informed traders to trade when markets are more liquid, when prices are more volatile, when uninformed volume is higher, and when uninformed volume is more consistent.

For a metric for illiquidity, we take the first principle component of four liquidity metrics averaged for the same time increments in the same way: the *relative bid-bid spread*, the bid-ask spread divided by the midquote; the *inside depth*, the sum of limit orders available at the bid and ask; the *five-minute per-trade price impact,* the price impact computed the traditional way as the aggressive-trade price impact, the difference of the midquote contemporaneous to a trade and the midquote five minutes later, signed by the sign of the aggressive trade (positive for buys, negative for sells); and a *30-minute Kyle's lambda,* the lambda computed from equation 4 for 30-minute returns and for aggressive order flow. Using the first principle component, we compute *illiquidity,* the relative spread predicted by the principle component.

For the other metrics, we compute the *return volatility*, the standard deviation of five-minute returns; *uninformed volume*, the volume of uninformed trading during the interval; and the *standard deviation of uninformed volume,* the standard deviation of changes in the uninformed volume during the interval. We run the regression once in levels and again in z-scores, as an effect is sometimes easier to interpret in normalized values. This shows that our conclusions are robust to normalizations. We focus on the results in z-scores for interpretive ease.

TABLE IX HERE

In Table IX, earnings-informed volumes correlate strongly with the CDF factor even after controlling for liquidity, volatility, and other measures of uninformed volumes. The CDF factor in z-scores has a loading of 0.42, 0.46 and 0.39 at the intraday, daily and weekly frequencies. Its magnitude is rivaled only by the coefficient on the raw volume of uninformed trading. The theoretically warranted explanation is that earnings-informed investors are trading strategically by timing their trading on the uninformed. Still, the other variables are statistically significant in 21 out of 24 coefficients, which implies that the CDF factor is not the only element in the strategy of earnings-informed investors.

TABLE X HERE

Next, using Table X, we compare the results of Table IX to those of the five-second traders. In both levels and in z-scores, the volumes of five-second traders have a weaker loading on the CDF factor, implying they are less strategic in the sense of strategy developed in the theory. In z-scores, the loadings are 0.29, 0.25, and 0.21 at the intraday, daily and weekly frequencies, all of which are lower than those for the earnings informed.

As with the earnings-informed, it is clear the CDF factor is not the only element in the strategy of five-second traders. In fact, the factor loadings of the five-second traders on the other variables are sometimes larger than those of the earnings-informed. For example, the loadings on return volatility are 0.14, 0.17, and 0.19 for the three intervals, compared to only 0.06, 0.08, and 0.09 for the earnings informed. It is intuitive that a group of trader likely containing statistical and latency arbitrageurs would be more attracted to volatility. Thus, an alternative explanation of our data could be that the CDF factor is not the central element in the strategy of five-second traders.

IV. Conclusion

Motivated by concerns in the theoretical literature that it is becoming harder to trade on private information—and thus that prices may be losing meaning—we study long-term informed investors over a period of 11 years. We address the concerns by finding that a group of informed investors have steady trading volumes and are making profits steadily through the sample. Shorter-term traders trade with them, both intraweek and intraday, which are both inconsistent with predatory trading. Further, and unexpectedly, we also found results contributing to wider questions in finance—questions about how or even whether information is getting into prices. We question whether private information acquisition leads to substantial information revelation via trading.

We identify a group of "earnings-informed investors" who repeatedly and materially anticipate large earnings surprises. This group of traders makes profits at a steady rate through the sample and does not experience worsening trading performance during a period of growing computerized trading. This indicates that the work of informed trading may not be becoming harder at least as measured by trading profits and by price impacts. Accordingly, we did not observe any change in aggregate measures of price efficiency. This poses the question of how this group of traders is trading so well. We find they use a fairly invariant policy of patient, largely passive trading resulting in price impacts that, both on a trade-by-trade basis and on a daily aggregate basis, are zero or negative. This is surprising, given the theory on price discovery that suggests informed investors impound their information into prices by moving prices in the direction of their trades. In contrast, we show negative price impacts of individual trades, a sign of a successful disguise of trading intent.

We also identify a group of "five-second traders," and we look at how they interact with the earnings-informed investors. Contrary to concerns that they are trading with the information of the informed, they do not appear to be moving the price in the direction of the information of the earnings-informed. Rather, they trade in the opposite direction of the information of informed investors both intraday and week-over-week. The five-second traders in our data would be better described as liquidity suppliers than as the back-runners or predators described in theoretical models. Ironically, it might be better for price efficiency if they were better predators, since it would elicit a positive price impact from informed trading.

Our sense is that the fundamentally informed investor, who has been construed as the prey in the ecology of the financial market, is in fact an apex predator. We confirm this sense using a test of the trade timing of the earnings-informed investors, who time their trade strongly with the theoretical timing factor given by Collin-Dufresne and Fos (2016). They trade with this factor more strongly than the five-second traders. This suggests that the more important concern for scholars and regulators may not be how to control the shorter-term trading community but how to bring more private information into prices. This is fruitful ground for research as there could be many ways: more frequent earnings reports, more competition among asset managers, or different restrictions on insider trading. Future work could focus on the welfare implications of private-information acquisition versus public-information revelation. If private information does not enter into prices, or price movements are driven primarily by public information, does this change the informational role of the market?

REFERENCES

Baldauf, Markus, and Joshua Mollner, 2020, High-Frequency Trading and Market Performance, *The Journal of Finance* .

Barclay, Michael J, and Jerold B Warner, 1993, Stealth trading and volatility: Which trades move prices?, *Journal of Financial Economics* 34, 281–305.

Boulatov, Alex, and Thomas J George, 2013, Hidden and displayed liquidity in securities markets with informed liquidity providers, *The Review of Financial Studies* 26, 2096–2137.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-Frequency Trading and Price Discovery, *The Review of Financial Studies* 27, 2267–2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2019, Price discovery without trading: Evidence from limit orders, *The Journal of Finance* 74, 1621–1658.

Bushee, Brian J., and Theodore H. Goodman, 2007, Which Institutional Investors Trade Based on Private Information About Earnings and Returns?, *Journal of Accounting Research* 45, 289–321.

Chakravarty, Sugato, 2001, Stealth-trading: Which traders trades move stock prices?, *Journal of Financial Economics* 61, 289–307.

Chakravarty, Sugato, and John J. McConnell, 1999, Does Insider Trading Really Move Stock Prices?, *Journal of Financial and Quantitative Analysis* 34, 191–209.

Chen, Marie, and Corey Garriott, 2019, High-frequency trading and institutional trading costs, *Journal of Empirical Finance* .

Collin-Dufresne, Pierre, and Vyacheslav Fos, 2015, Do Prices Reveal the Presence of Informed Trading?, *The Journal of Finance* 70, 1555–1582.

Collin-Dufresne, Pierre, and Vyacheslav Fos, 2016, Insider Trading, Stochastic Liquidity, and Equilibrium Prices, *Econometrica* 84, 1441–1475.

Cornell, Bradford, and Erik R. Sirri, 1992, The Reaction of Investors and Stock Prices to Insider Trading, *The Journal of Finance* 47, 1031–1059.

Dessaint, Olivier, Thierry Foucault, Laurent Frésard, and Adrien Matray, 2019, Noisy stock prices and corporate investment, *The Review of Financial Studies* 32, 2625–2672.

Dugast, Jérôme, and Thierry Foucault, 2018, Data abundance and asset price informativeness, *Journal of Financial Economics* 130, 367–391.

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of Finance* 66, 1–33.

Hirschey, Nicholas, 2020, Do high-frequency traders anticipate buying and selling pressure?, *Management Science* .

Kacperczyk, Marcin T., and Emiliano Pagnotta, 2018, Chasing Private Information, SSRN Scholarly Paper ID 2695197, Social Science Research Network, Rochester, NY.

Kaniel, Ron, and Hong Liu, 2006, So What Orders Do Informed Traders Use?, *The Journal of Business* 79, 1867–1913.

Korajczyk, Robert A., and Dermot Murphy, 2019, High-Frequency Market Making to Large Institutional Trades, *The Review of Financial Studies* 32, 1034–1067.

Kothari, SP, 2001, Capital markets research in accounting, *Journal of accounting and economics* 31, 105–231.

Koudijs, Peter, 2015, Those Who Know Most: Insider Trading in Eighteenth-Century Amsterdam, *Journal of Political Economy* 123, 1356–1409.

Koudijs, Peter, 2016, The Boats That Did Not Sail: Asset Price Volatility in a Natural Experiment, *The Journal of Finance* 71, 1185–1226.

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315–1335.

Lee, Charles M. C., and Edward M. Watts, 2018, Tick Size Tolls: Can a Trading Slowdown Improve Price Discovery?, SSRN Scholarly Paper ID 3263778, Social Science Research Network, Rochester, NY.

Martineau, Charles, 2019, The Evolution of Financial Market Efficiency: Evidence from Earnings Announcements, SSRN Scholarly Paper ID 3111607, Social Science Research Network, Rochester, NY.

Menkveld, Albert J, 2013, High frequency trading and the new market makers, *Journal of financial Markets* 16, 712–740.

Meulbroek, Lisa K., 1992, An Empirical Analysis of Illegal Insider Trading, *The Journal of Finance* 47, 1661–1699.

O'Hara, Maureen, and Mao Ye, 2011, Is market fragmentation harming market quality?, *Journal of Financial Economics* 100, 459–474.

Shkilko, Andriy, 2019, Insider trading under the microscope .

Stiglitz, Joseph E, 2014, Tapping the brakes: Are less active markets safer and better for the economy? .

Tabb, Larry, 2017, Speed: Why it matters and what can be done .

van Binsbergen, Jules H, and Christian C Opp, 2017, Real Anomalies, Working Paper 23238, National Bureau of Economic Research.

van Kervel, Vincent, and Albert J. Menkveld, 2018, High-Frequency Trading around Large Institutional Orders, SSRN Scholarly Paper ID 2619686, Social Science Research Network, Rochester, NY.

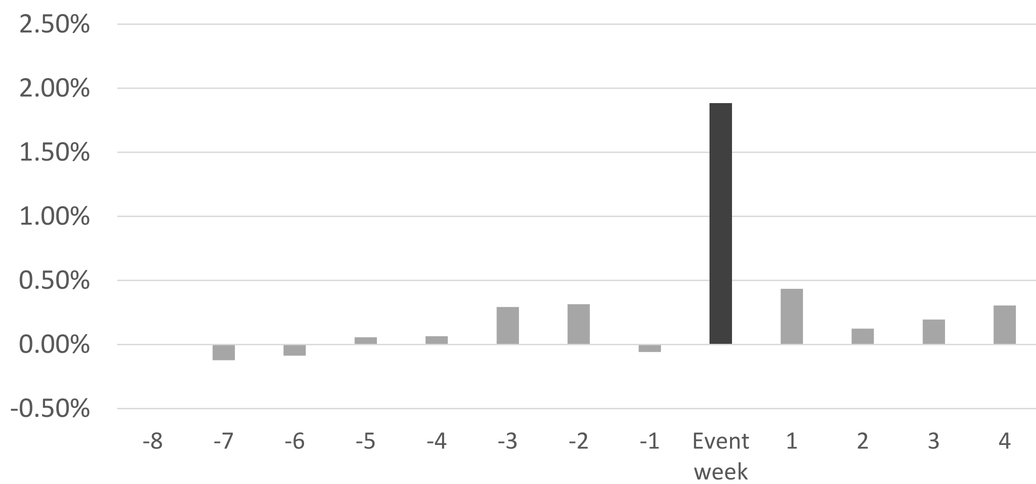
Weller, Brian M., 2018, Does Algorithmic Trading Reduce Information Acquisition?, *The Review of Financial Studies* 31, 2184–2226.

Yang, Liyan, and Haoxiang Zhu, 2019, Back-running: Seeking and hiding fundamental information in order flows, *Review of Financial Studies, Forthcoming* .

Figure 1: Weekly stock returns around material earnings surprises

This figure shows the weekly returns (Friday close price to Friday close price) averaged across stock-weeks in the sample for the eight weeks before, week of, and four weeks after material earnings surprises that are positive and negative. *Positive* and *negative* surprises are classified by whether the reported earnings were greater or lesser than analyst expectations. Earnings surprises are *material* if the absolute magnitude of the surprise was greater in absolute value than the median for positive and negative surprises.

(a) Weekly returns around positive material surprises



(b) Weekly returns around negative material surprises

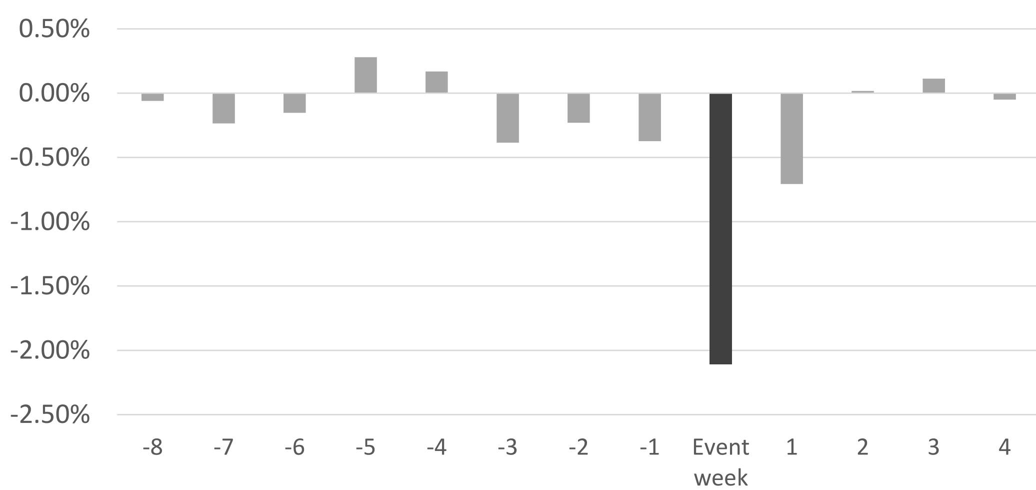


Figure 2: Trading volumes of the three trader groups

This figure shows the quarterly trading volume (in billions of shares) of the three trader groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter.

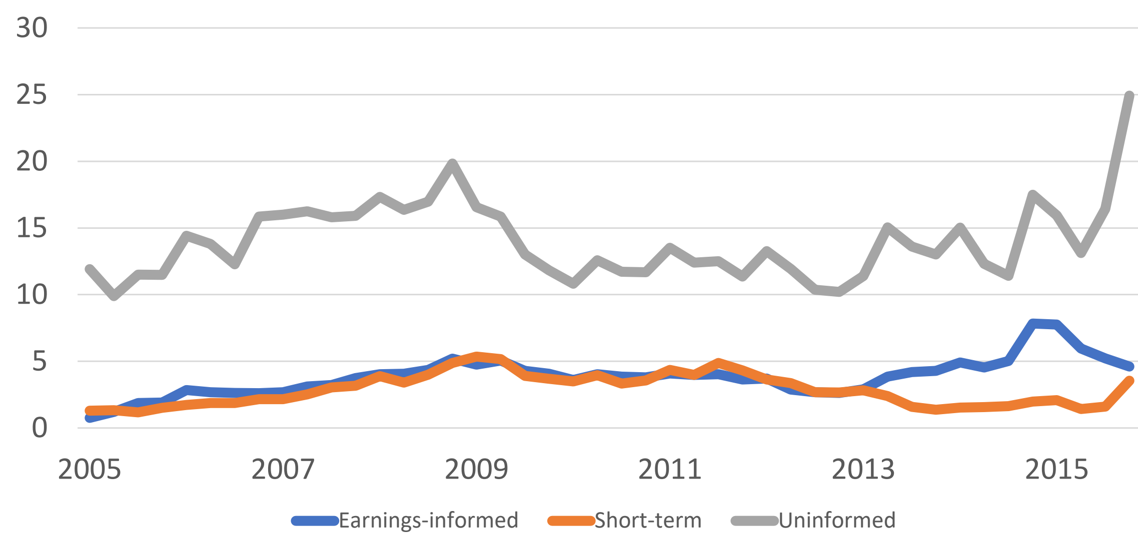
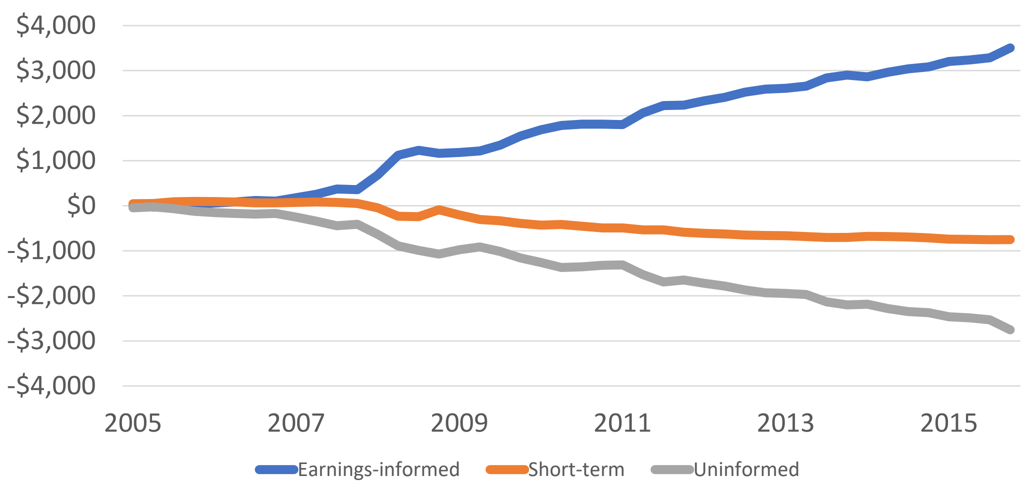


Figure 3: Cumulated profits of the three trader groups

These two figures show the cumulated trading profits (in millions of dollars) of the three trader groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter.

The profits are cumulated separately for material surprises and for immaterial surprises. Earnings surprises are *material* if the absolute magnitude of the surprise was greater in absolute value than the median (for positive and negative surprises separately).

(a) Material surprises



(b) Immaterial surprises

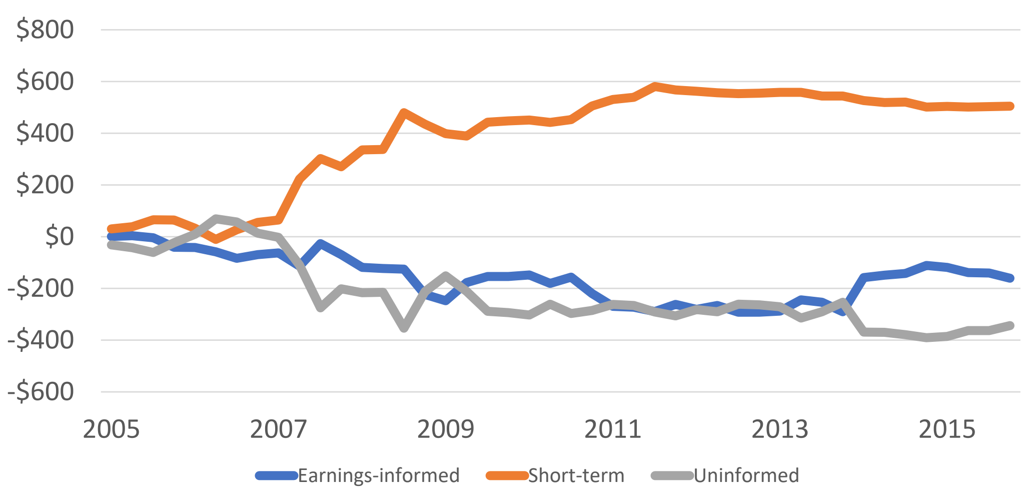


Figure 4: Intraday price impacts, by trader groups

This figure shows the per-trade volume-weighted price impact in basis points averaged over all trades in the sample period of the three trader groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter. Price impacts are the percent changes in the midquote price from the moment of trade to the named increment after a trade.

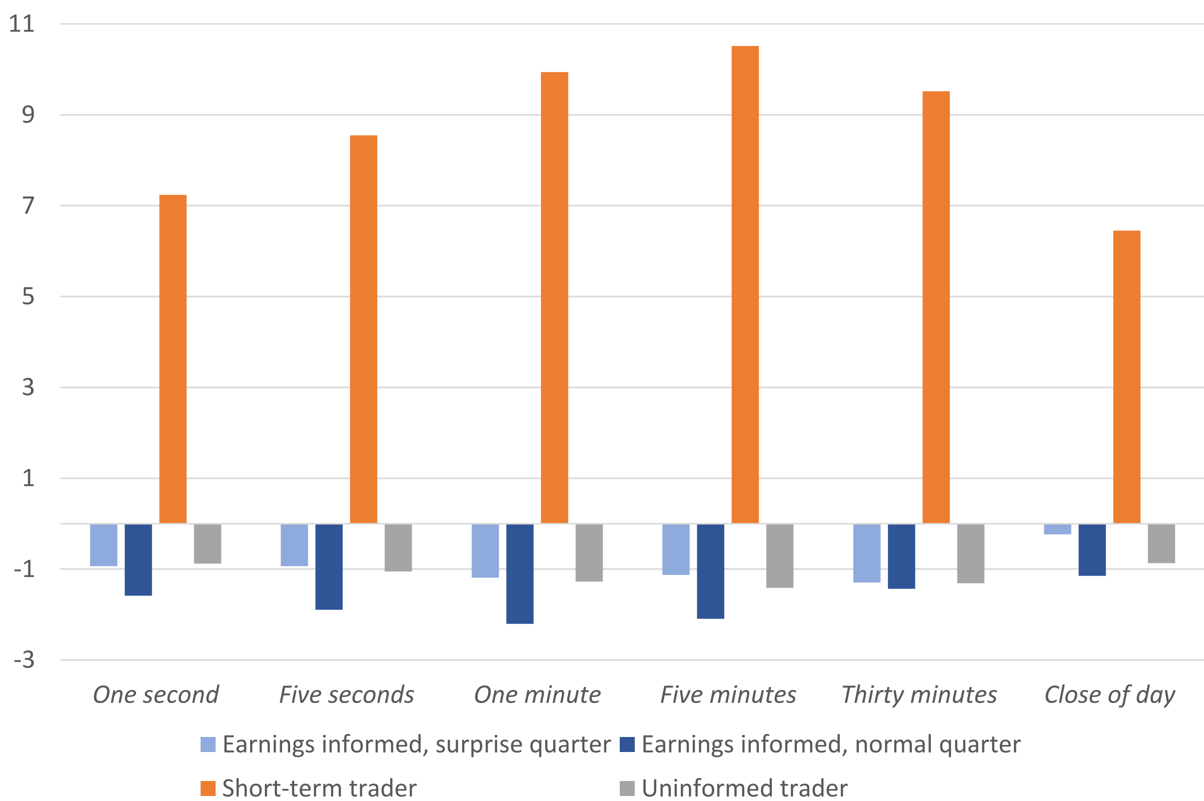


Figure 5: Kyle's lambdas, by trader groups

This figure shows the Kyle's lambda of daily order imbalance on daily returns. Kyle's lambda is a coefficient of regression identified by regressing daily net trading flow (shares bought less shares sold), here in units of \$100K, on the daily midquote return in basis points,

$$r_t = \lambda flow_t + \epsilon_t,$$

where t indexes the day. Each observation in the regression is one stock-day in sample period for each of the three trader groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter.

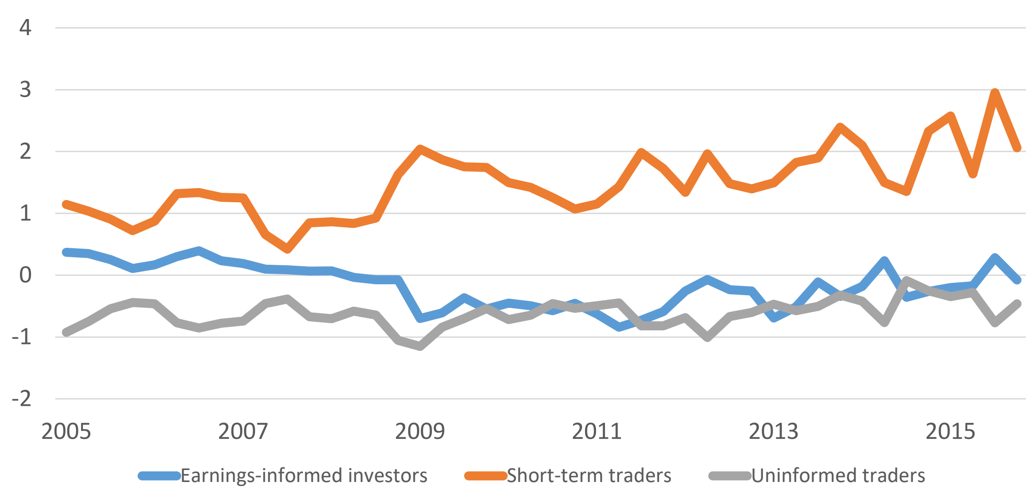


Figure 6: Measures of the informativeness of earnings reports

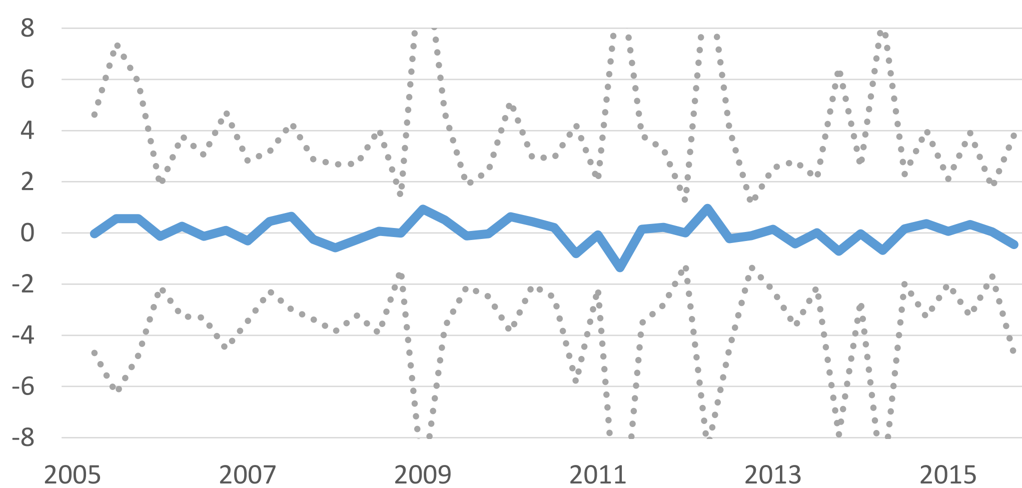
These two figures show the average value and confidence interval of two measures of the information content of earnings reports: the relative return and the jump ratio. The measures are averaged over the stocks in the sample on the days of earnings reports. The first measure, relative return, is a simplification of the jump ratio. It compares the one-week stock return since an earnings report to its return from eight weeks ago,

$$relativeReturn_{i,t} = \frac{r_{i,t+7,t-1}}{r_{i,t+7,t-56}}.$$

The second measure, the jump ratio, is the ratio of the cumulative abnormal returns according to the parameters defining the ratio in Weller (2018),

$$jumpRatio_{i,t} = \frac{CAR_{t+3,t-1}}{CAR_{t-23,t-1}}.$$

(a) Relative return



(b) Jump ratio

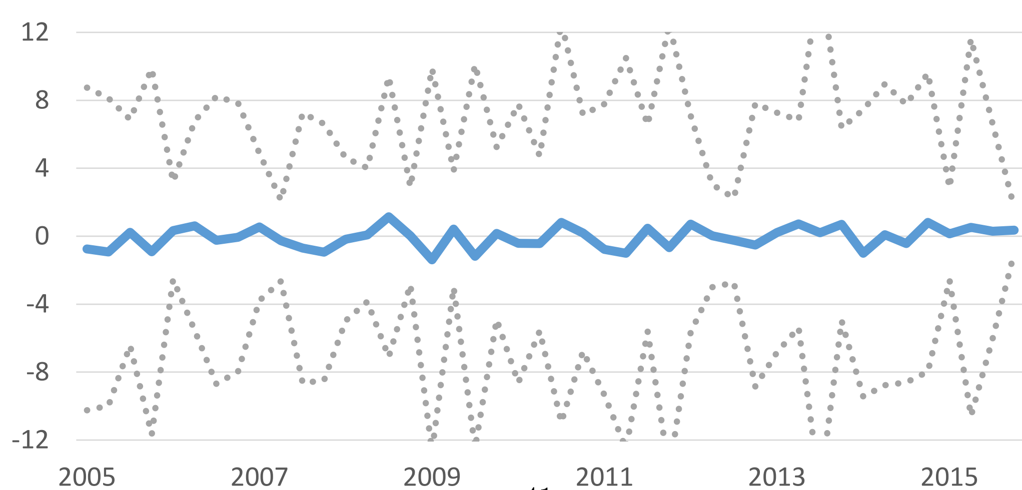


Figure 7: Weekly cash flows for the three types of trader around material surprises

This figure shows the weekly cash flows (cash spent less cash received) averaged across stock-weeks in the sample for the eight weeks before, week of, and four weeks after material earnings surprises for the three trader groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter. Earnings surprises are *material* if the absolute magnitude of the surprise was greater in absolute value than the median for positive and negative surprises.

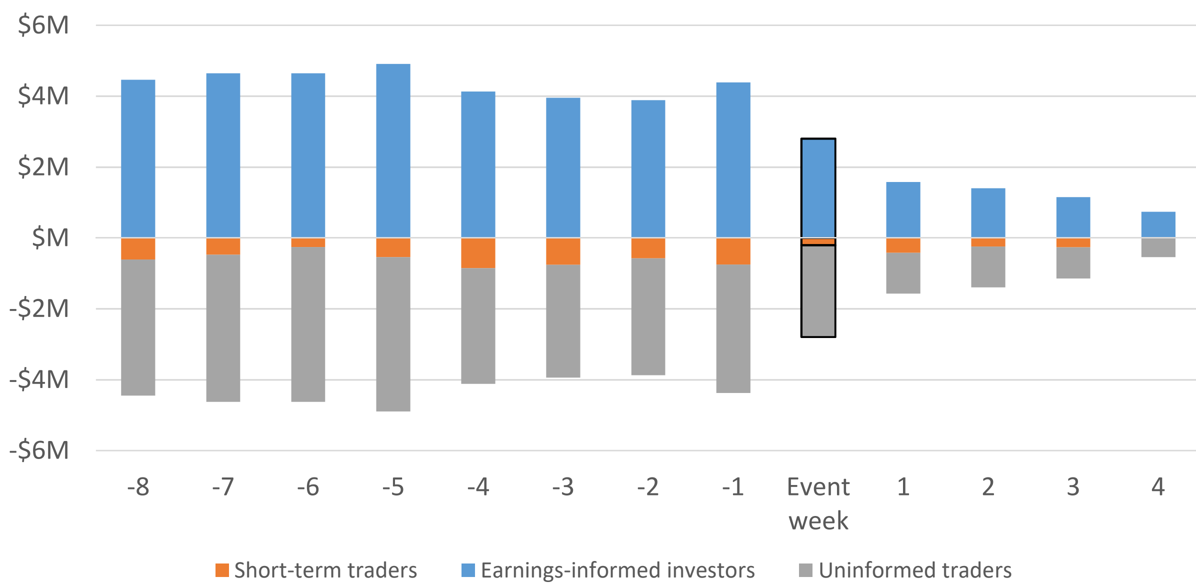
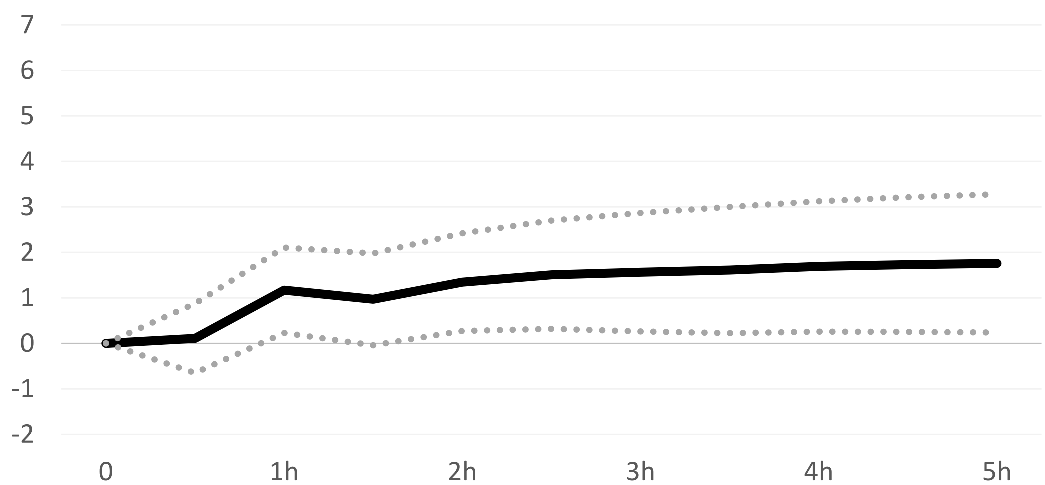


Figure 8: Sample average impulse response of the return to informed trading

This figure shows the mean cumulative impulse response and 95% confidence interval of the stock return (in basis points) to a \$100K impulse in the order imbalance of two trader types, earnings-informed and five-second-informed. The impulse responses are generated from 30-minute VAR models with three lags, fit once per stock, and fit on a vector of three variables: earnings-informed order imbalance, five-second-informed order imbalance, and the return in basis points. The impulse reponses are averaged over all stocks in the sample, and the confidence interval is taken from the standard error of the mean. *earnings-informed investors* are traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *five-second traders* are traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter.

(a) Return response to an impulse of earnings-informed investing



(b) Return response to an impulse of five-second trading

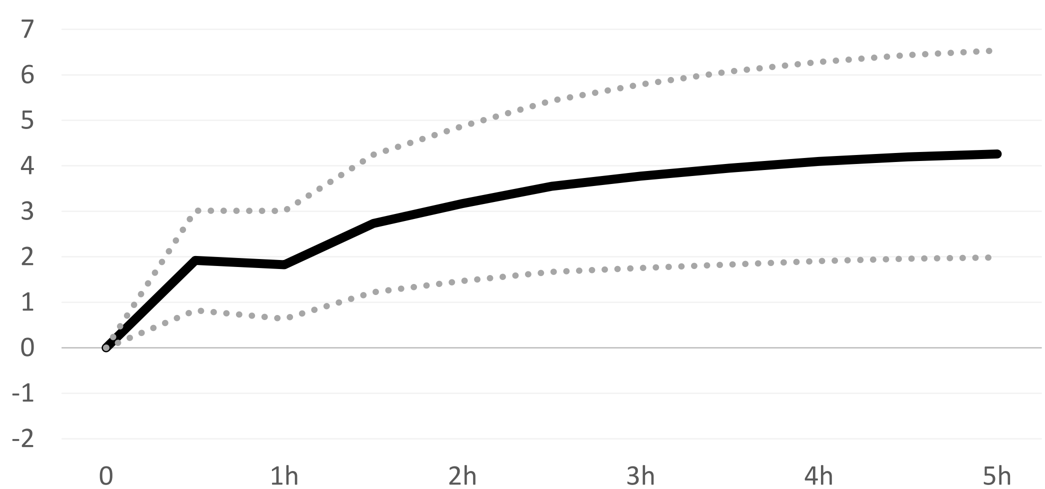
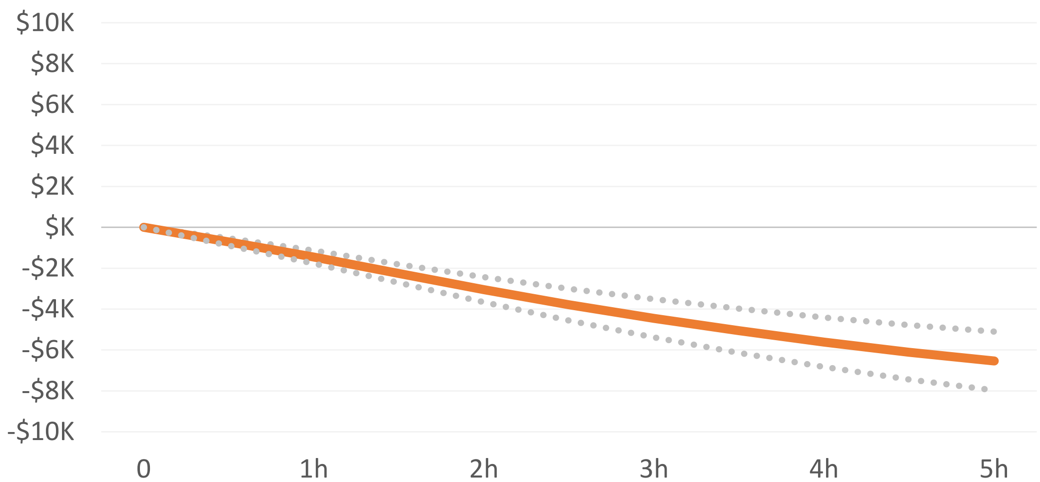


Figure 9: Sample average impulse responses of the two informed trader types to each other

This figure shows the mean cumulative impulse response and 95% confidence interval of the order imbalance (in dollars) of the informed trader groups to a \$100K impulse in the order imbalance of the other trader group. The impulse responses are generated from 30-minute VAR models with three lags, fit once for each stock, and fit on a vector of three variables: earnings-informed order imbalance, five-second-informed order imbalance, and the return in basis points. The impulse reponses are averaged over all stocks in the sample, and the confidence interval is taken from the standard error of the mean. The trader groups are the earnings-informed and five-second-informed. *earnings-informed investors* are traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *five-second traders* are traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; *uninformed traders*, traders who meet neither of these criteria during the quarter.

(a) Response of five-second traders to an impulse of earnings-informed investing



(b) Response of earnings-informed to an impulse of five-second trading

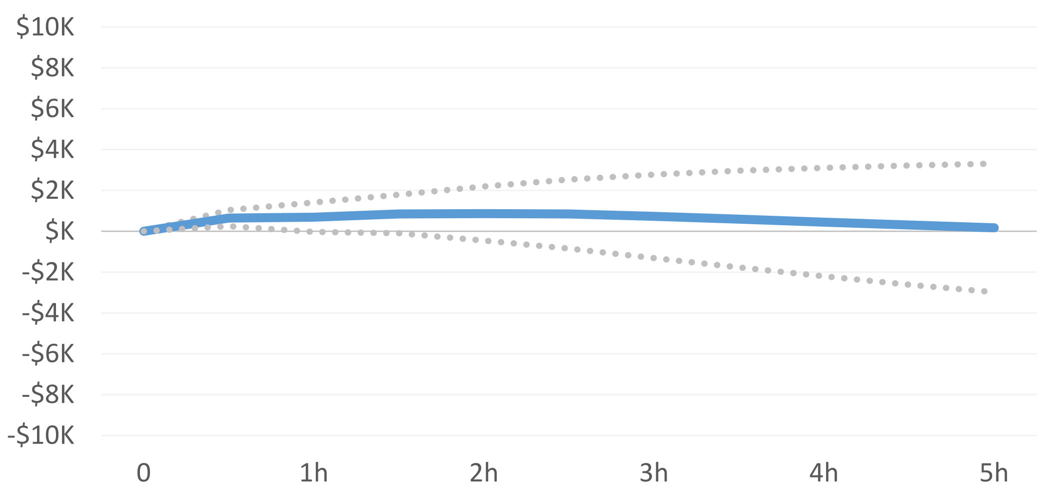


Table I: Summary statistics on stocks, accounts and events

This table shows stock-quarterly counts or stock-quarterly averages for statistics on the stocks, accounts, and events in the sample, during four selected subsample periods: 2005–2006, 2007–2009, 2010–2012, and 2013–2015. For stocks, the statistics are: *Number,* the number of distinct stocks in the sample during the period; *quarterly volume,* the sum of the trading volume (single-counted); *quarterly trades,* the sum of the trades (single-counted); *market cap,* the prices times shares outstanding; *relative bid-ask spread*, the bid-ask spread divided by the price; *inside depth*, the sum of the shares outstanding at the best bid and best ask prices; *volatility*, the 20-day rolling return volatility; *average trade size*, the trade-weighted daily average trade size; and *median trade size*, the trade-weighted daily median trade size.

For accounts, the statistics are by trader type: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter. The statistics are the *number,* the count of unique accounts; and *quarterly total volume,* the sum of the trading volume by account (single-counted).

For events, the statistics are by the type of earnings surprise: *material* if the absolute magnitude of the surprise was greater in absolute value than the median. The statistics are the *number,* the count; and the *absolute size of surprises,* the average distance between the realized earnings per share and the median analyst expectation in absolute value.

	2005–2006	2007–2009	2010–2012	2013–2015
Panel A: Stocks				
Number	247	269	248	230
Quarterly volume	70.9M	110M	86.5M	105M
Quarterly trades	88.8K	244K	270K	377K
Market cap	\$4.8B	\$5.9B	\$6.6B	\$7.5B
Rel. bid-ask spread	33.0 bps	35.1 bps	18.0 bps	24.6 bps
Inside depth	6056.7	6565.2	6427.9	7643.7
20-day rolling volatility	33.54	52.68	31.48	32.61
Trade size	872.0	658.7	373.6	290.1
Median trade size	218.4	203.7	126.2	125.0
Panel B: Accounts				
Earnings informed, number	9.0	23.8	23.1	28.8
quarterly total volume	8.9M	18.6M	16.1M	24.2M
Five-second traders, number	6.4	16.5	23.4	11.6
quarterly total volume	7.0M	17.3M	16.6M	9.3M
Uninformed, number	140.5	189.6	183.5	213.7
quarterly total volume	54.3M	73.2M	53.5M	71.2M
Panel C: Earnings surprises				
Material, number	325	396.3	401	367
|size| of surprises	0.210	0.207	0.246	0.241
Immaterial, number	267.5	268.7	326.3	283.7
|size| of surprises	0.056	0.051	0.039	0.042

Table II: Sample summary statistics on earnings events

This table gives summary statistics on the earnings events in the sample. The events are split into four groups by direction and by materiality. They are split into *positive* and *negative* surprises by whether the reported earnings were greater or lesser than analyst expectations; they are split into *material* and *immaterial* surprises by whether the absolute magnitude of the surprise was above or below the median (for positive and for negative surprises separately). The table reports statistics on: *reported earnings*, the earnings per share reported by companies; *expected earnings*, the median analyst expectation of the earnings per share that would be reported; *surprise earnings*, the difference between reported and analyst expected earnings; *one-day return*, the difference between the close price on the day of the report and the previous day's close price, in percentage points; *two-day return*, the difference between the close price *one day after* the report and the previous day's close price, in percentage points; *seven-day return*, the difference between the close price *six days after* the report and the previous day's close price, in percentage points.

	mean	median	stddev	min	p25	p75	max	count
Panel A: Positive material surprises								
Reported earnings	0.65	0.42	2.01	-1.62	0.19	0.73	47.17	2184
Expected earnings	0.43	0.31	1.09	-18.09	0.11	0.60	24.40	2184
Surprise earnings	0.22	0.07	1.25	0.01	0.04	0.15	29.23	2184
One-day return	1.79	1.14	5.59	-41.41	-0.76	3.79	89.83	2184
Two-day return	2.24	1.64	7.13	-54.45	-0.93	4.94	110.4	2184
Seven-day return	2.57	1.96	9.32	-98.19	-1.54	6.28	96.67	2184
Panel B: Negative material surprises								
Reported earnings	0.10	0.13	1.28	-38.47	-0.03	0.36	6.29	2138
Expected earnings	0.34	0.24	0.92	-8.13	0.05	0.48	32.57	2138
Surprise earnings	-0.24	-0.08	1.37	-42.72	-0.16	-0.04	-0.01	2138
One-day return	-1.75	-1.13	5.65	-78.36	-3.85	0.65	69.23	2138
Two-day return	-2.20	-1.63	6.63	-79.03	-5.01	1.12	61.54	2138
Seven-day return	-3.08	-2.52	9.15	-77.76	-7.16	1.28	76.00	2138
Panel C: Positive immaterial surprises								
Reported earnings	0.49	0.33	1.09	-2.30	0.14	0.60	30.15	1792
Expected earnings	0.43	0.30	0.86	-5.58	0.12	0.57	26.62	1792
Surprise earnings	0.06	0.02	0.42	0.01	0.01	0.04	7.16	1792
One-day return	0.46	0.19	4.15	-33.33	-1.42	2.18	36.88	1792
Two-day return	0.53	0.48	5.43	-49.44	-1.89	2.98	33.84	1792
Seven-day return	0.53	0.41	7.99	-55.90	-3.18	4.04	147.5	1792
Panel A: Negative immaterial surprises								
Reported earnings	0.28	0.19	0.46	-5.78	0.04	0.43	5.14	1718
Expected earnings	0.32	0.22	0.50	-4.93	0.07	0.45	7.28	1718
Surprise earnings	-0.04	-0.02	0.14	-4.35	-0.04	-0.01	-0.01	1718
One-day return	-0.52	-0.37	4.23	-35.00	-2.32	1.37	46.70	1718
Two-day return	-0.62	-0.38	5.38	-40.94	-2.93	1.87	49.25	1718
Seven-day return	-0.92	-0.81	7.57	-45.81	-4.56	2.34	64.18	1718

Table III: The predictability of earnings surprises

This table shows the result of four regressions on the presence of a material earnings surprise or on the size of an earnings surprise using lagged data on the previous surprise. Earnings surprises are *material* if the absolute magnitude of the surprise was greater in absolute value than the median for positive and negative surprises. The size of a surprise is the distance between the earnings per share and the median analyst expectation of the earnings per share.

Columns 1 and 2 show a logit regression predicting materiality status using the materiality status of the previous earnings report for a stock and the size of the earnings surprise of the previous earnings report. Column 1 does not use fixed effects; column 2 uses stock fixed effects.

Columns 3 and 4 show a panel regression predicting the size of an earnings surprise. Column 3 predicts the signed magnitude of the surprise using lagged size and lagged materiality status. Column 4 predicts the absolute magnitude of the surprise using the lagged absolute magnitude of the surprise and the lagged materiality status.

	Materiality		Size	|Size|
	Logit	FE logit	Panel reg.	Panel reg.
Lagged materiality	-0.075	-0.176**		-0.005
	(-1.44)	(-3.21)		(-1.30)
Lagged |size|	0.552**	0.986***		0.129*
	(2.91)	(3.78)		(2.55)
Lagged size			-0.000	
			(-0.01)	
Lagged materiality			0.009	
* sign(size)			(2.07)	
Constant	0.266***		-0.004***	0.080***
	(7.17)		(-10.51)	(24.01)
Observations	7027	7000	7025	7025
Pseudo R2	0.001	0.002		
R2			0.124	0.430
Stock FE	NO	YES	YES	YES
Quarter FE	NO	NO	YES	YES
Within			0.001	0.014

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IV: Sample summary statistics by user groups

This table gives means and standard deviations on metrics computed for all stock-day-accounts in the sample, in which the accounts are divided into the three user groups: *earnings-informed investors*, traders who repeatedly materially trade in the direction of material earnings surprises and do so on the quarter or a quarter adjacent to the surprise; *short-term traders*, traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter; and *uninformed traders*, traders who meet neither of these criteria during the quarter. For the purposes of this table, we exclude traders who are both earnings-informed investors and short-term traders.

The metrics are: *Daily volume*, the daily trading volume of an account; *trades per day*, the number of trades of an account; *median trade size*, the median size of a trade by an account; *percent of trading days traded*, the percent of trading days in a quarter of nonzero trading by an account; *percent aggressive*, the percent volume executed via market orders by an account; *daily directionality*, the end-of-day position of an account divided by its daily volume (an account that bought 9,000 shares and sold 1,000 shares would have a daily volume of 10,000 shares and a directionality of 80); *quarterly directionality*, the end-of-quarter position of an account divided by its quarterly volume.

	Earnings informed	Five-second informed	Un-informed
Daily volume	13045.1	16794.1	6366.8
	(42131.7)	(54068.2)	(27400.2)
Trades per day	38.74	56.76	16.87
	(119.3)	(155.5)	(67.15)
Median trade size	2161.0	1235.2	4075.0
	(22034.8)	(15784.2)	(52522.4)
Percent of trading days traded	0.452	0.545	0.357
	(0.355)	(0.377)	(0.340)
Percent aggressive	42.76	69.01	44.59
	(25.57)	(27.26)	(28.64)
Directionality (daily)	77.63	56.29	72.89
	(22.49)	(33.44)	(30.15)
Directionality (quarterly)	40.49	26.11	39.13
	(32.93)	(32.48)	(36.48)

Table V: Explaining the informational efficiency of prices

This table gives results from a panel regression of two metrics of price efficiency (relative return and jump ratio) on signed trading volumes by earnings-informed, five-second, and uninformed trading volume (in millions of shares) and, in the columns labeled IV, on the unsigned (absolute) five-second trader volumes using the IV setup of Weller (2018). The relative return is the ratio of the return from one day before an earnings report to seven days after and the return from 56 days before an earnings report and seven days after. The jump ratio is the cumulative abnormal return from one day before an earnings report to three days after and the cumulative abnormal return from 23 days before an earnings report to one day before.

In the panel, the cross-sectional variable is the stock, and the time variable is the quarter. The explanatory variables are the two-week-lagged, two-week averaged natural logs of: *market capitalization*, the stock price times the number of shares outstanding; *relative bid-ask spread*, the bid-ask spread in basis points; *price*, the stock price; and *volatility*, the rolling standard deivation of the stock return. Fixed effects are stock and quarter; standard errors in all specifications are clustered by stock and quarter.

In columns labeled "All," the regression is performed using all earnings reports. In the columns labeled "Mat.," the regression is performed on only the earnings reports with material surprises. In the columns labeled "IV," the regression is performed using the IV specification from Weller (2018) in which the lagged log price is an instrument on the absolute value of short-term traders' trading volume.

	Relative return			Jump ratio		
	All	Mat.	IV	All	Mat.	IV
Earn.-inf. volume	-0.0456	-0.0200		0.0270	0.0371	
	(-0.85)	(-0.45)		(0.53)	(0.68)	
Five-sec. volume	0.00698	-0.0310		0.00603	0.0465	
	(0.16)	(-0.47)		(0.17)	(1.01)	
|Five-sec. volume|			-0.405			-0.372
			(-0.71)			(-0.60)
Uninformed volume	-0.0595	-0.0679		0.0449	0.0176	
	(-1.03)	(-1.17)		(1.05)	(0.44)	
Market cap	-1.315	-8.310	0.560	-4.750	0.989	-0.919
	(-0.38)	(-1.87)	(0.30)	(-1.58)	(0.18)	(-0.37)
Rel. bid-ask spread	0.0116	-0.0504	-0.0680	0.0273	0.364	-0.126
	(0.13)	(-0.40)	(-0.69)	(0.22)	(1.78)	(-1.00)
Volatility	-0.580	-0.570	0.136	-2.186**	-3.999***	-1.159
	(-0.87)	(-0.72)	(0.14)	(-3.00)	(-4.07)	(-0.97)
Price	2.112	4.426		5.197	2.142	
	(0.56)	(1.04)		(1.53)	(0.43)	
Constant	0.021***	0.028***	0.020	0.233***	0.243***	0.232***
	(3.68)	(4.86)	(0.67)	(38.44)	(20.30)	(5.82)
Observations	6020	3228	6025	6295	3214	6298
R2	0.040	0.079	0.000	0.039	0.066	0.000
FE	YES	YES	YES	YES	YES	YES
Within	-0.001	-0.000	0.000	0.001	0.002	0.000

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VI: Explaining stock choice by trader type

This table gives results from a panel regression of trading volumes (in millions of shares) by earnings-informed traders and by short-term informed traders on explanatory variables, particularly on industry codes. The panel variable is a stock, and the time variable is a quarter. The explanatory variables are: *uninformed volume*, the trading volume of uninformed traders in millions of shares; log *market capitalization*, the stock price times the number of shares outstanding; log *illiquidity*, the bid-ask spread in basis points; and log *volatility*, the rolling standard deivation of the stock return. We also regress on nine industry dummies as listed. Fixed effects are stock and quarter; standard errors in all specifications are clustered by stock and quarter.

	Earnings-informed investor volume		Five-second trader volume	
	1	2	1	2
Uninf. volume	0.234***	0.140**	0.199***	0.154***
	(12.73)	(3.26)	(14.12)	(7.30)
Market cap	78.89***	-17.50	89.54***	141.2***
	(3.68)	(-0.30)	(5.50)	(4.37)
Illiquidity	0.560	2.346	-0.261	1.458
	(0.49)	(1.45)	(-0.32)	(1.39)
Volatility	7.579	2.967	29.20***	25.91***
	(0.56)	(0.25)	(4.34)	(4.35)
Finance	1.580		-2.470	
	(0.70)		(-1.56)	
Information	2.404		0.183	
	(1.18)		(0.11)	
Manufacture	0.895		-0.590	
	(0.70)		(-0.54)	
Mining	4.987**		0.868	
	(2.85)		(0.85)	
Real Estate	-1.348		-0.642	
	(-1.28)		(-0.50)	
Retail	-1.264		-2.312	
	(-1.11)		(-1.76)	
Transport	-1.068		-4.184**	
	(-0.70)		(-3.49)	
Utilities	0.329		-2.593	
	(0.27)		(-1.86)	
Wholesale	-0.545		0.354	
	(-0.50)		(0.32)	
Constant	17.70***	19.91***	14.49***	14.06***
	(21.31)	(278.00)	(17.27)	(503.57)
Observations	6239	6401	6239	6401
R2	0.564	0.634	0.713	0.765
Quarter FEs	YES	YES	YES	YES
Stock FEs	NO	YES	NO	YES
Within	0.542	0.077	0.695	0.233

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VII: Explaining the order choice of earnings-informed investors

This table shows the results of a regression explaining the weekly percent passive volume of earnings-informed investors during the quarters before material earnings surprises using the weekly average of the explanatory variables: *volume share of five-second traders*, the percent of a stock's trading volume in which a five-second trader was a counterparty; *volume share of uninformed traders*, the percent of a stock's trading volume in which an uninformed trader was a counterparty; *signed return*, the daily return multiplied by the sign of the earnings surprise for the quarter (-1 for negative surprises); *trades*, the daily number of trades for the stock; *trading volume*, the daily number of shares traded for the stock; and *illiquidity*, the relative spread implied by the first principal component of stock relative spread, inside depth, five-minute per-trade price impact, and 30-minute Kyle's lambda.

	Earnings-inf. pct. passive	Earnings-inf. pct. passive	Earnings-inf. pct. passive	Earnings-inf. pct. passive
Volume share of five-sec.	42.91***	38.99***	24.10***	24.43***
	(23.09)	(18.60)	(11.05)	(9.35)
Volume share of uninf.		-1.176		0.461
		(-0.92)		(0.30)
Signed return		-0.078***		-0.088***
		(-3.53)		(-3.76)
Trades		30.23***		2.284
		(4.91)		(0.36)
Trading volume		-0.0844		0.0295
		(-0.61)		(0.31)
Volatility		0.199		0.345***
		(0.98)		(4.27)
Illiquidity		-0.0065		-0.0028
		(-1.25)		(-0.42)
Constant	50.99***	51.52***	53.01***	52.50***
	(176.81)	(52.60)	(250.21)	(42.49)
Observations	46415	46407	46415	46407
R2	0.085	0.095	0.194	0.196
FEs	NO	NO	YES	YES
Within			0.020	0.023

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VIII: Averaged 30-minute vector autoregressions on informed trading and returns

This figure shows the average coefficients from a series of 30-minute VAR models with three lags fit once per stock and fit on a vector of three variables: earnings-informed trading flow, short-term trading flow, and the return. Flow coefficients are multiplied by \$100K for legibility; return coefficients are multiplied by 100 for legibility. The coefficients are averaged over all stocks in the sample, and the t statistic from the standard error of the sample average is reported in parentheses. *earnings-informed investors* are traders who repeatedly materially trade in the direction of material earnings surprises and are also doing so on the quarter of the surprise.

		Earnings informed	Five-second informed	30-min. return
Earnings	Lag 1	42.19***	-0.68***	0.0005
informed		(92.36)	(-7.23)	(1.04)
trading	Lag 2	14.41***	-0.15*	0.0031*
		(71.87)	(-2.15)	(2.08)
	Lag 3	11.03***	-0.14	0.0007
		(68.23)	(-1.91)	(1.16)
Five-second	Lag 1	0.49*	32.07***	0.0017***
informed		(2.20)	(71.35)	(2.71)
trading	Lag 2	-0.22	13.35***	-0.0007
		(-0.87)	(61.16)	(-1.31)
	Lag 3	0.08	9.68***	0.0007
		(0.34)	(55.44)	(1.44)
30-min.	Lag 1	832.53	-2729.63***	-930.1***
return		(1.41)	(-3.74)	(-8.68)
	Lag 2	585.38	-1898.16***	-391.7***
		(1.52)	(-3.49)	(-6.12)
	Lag 3	1035.46***	-1808.28***	-173.6***
		(3.01)	(-3.83)	(-5.26)
Constant		69450.21	4785.25	-4.76***
		(1.90)	(0.20)	(-7.88)
Observations		4599935	4599935	4599935
R2		0.279	0.180	0.021

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IX: Explaining the level of earnings-informed volumes

This figure shows the coefficients from a panel regression on the dollar value of earnings-informed trading volume by stock on explanatory variables. *Earnings-informed investors* are traders who repeatedly materially trade in the direction of material earnings surprises and are also doing so on the quarter of the surprise. The regression is performed at three different intervals of measurement: half-hour, daily and weekly. The regression is peformed twice for each interval: once for all variables expressed in levels, and once in which the variables have been normalized quarterly and expressed in terms of their z-score (so that one unit is equal to one standard deviation). The regression uses stock and quarter fixed effects, and standard errors are double-clustered by stock and quarter.

The explanatory variables are: *standard deviation of uninformed flow*, the standard deviation of the net change in position of uninformed traders during the interval (*uninformed traders* are traders who meet no criteria for informedness during the quarter); *illiquidity*, the relative spread implied by the first principal component of stock relative spread, inside depth, five-minute per-trade price impact, and 30-minute Kyle's lambda; *return volatility*, the standard deviation of five-minute returns during the observation interval; *uninformed volume*, the volume of uninformed trading during the interval; and *the standard deivation of uninformed volume*, the standard deviation of changes in the uninformed volume during the interval.

	Half-hour observations		Daily observations		Weekly observations	
	in levels	in z-scores	in levels	in z-scores	in levels	in z-scores
$\sigma(\Delta$uninf. flow$)$	2.588***	0.421***	5.461***	0.461***	4.618***	0.389***
	(12.79)	(60.21)	(16.78)	(88.20)	(6.63)	(47.66)
Illiquidity	-1119.1***	-0.0303***	-2501.8	-0.0385***	-8142.4	-0.0353***
	(-3.93)	(-10.29)	(-1.38)	(-9.37)	(-1.16)	(-6.71)
Return volatility	483500.2***	0.0641***	-47539.4	0.0826***	-400576.8**	0.0934***
	(6.05)	(22.19)	(-2.01)	(25.32)	(-2.87)	(18.00)
Uninf. volume	0.188***	0.339***	0.212***	0.357***	0.173***	0.398***
	(11.93)	(21.02)	(12.86)	(20.53)	(9.21)	(22.02)
$\sigma(\Delta$uninf. volume$)$	-0.269***	-0.232***	-0.573***	-0.196***	-0.189*	-0.121***
	(-11.50)	(-21.25)	(-10.82)	(-24.97)	(-2.24)	(-15.68)
Constant	52612.0	0.00294***	2330712.6***	0.000151***	12219125.7***	0.00338***
	(1.75)	(6.33)	(6.35)	(7.86)	(5.08)	(16.42)
Observations	5382158	5382147	473031	473030	95663	95661
R2	0.334	0.273	0.611	0.372	0.710	0.384
Within	0.240	0.272	0.335	0.372	0.306	0.383

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table X: Explaining the level of five-second-informed volumes

This figure shows the coefficients from a panel regression on the dollar value of short-term trading volume by stock on explanatory variables. *Five-second traders* are traders whose average daily coefficient of regression of the five-second lagged order imbalance on the five-second return is statistically significantly positive during the quarter. The regression is performed at three different intervals of measurement: half-hour, daily and weekly. The regression is peformed twice for each interval: once for all variables expressed in levels, and once in which the variables have been normalized quarterly and expressed in terms of their z-score (so that one unit is equal to one standard deviation). The regression uses stock and quarter fixed effects, and standard errors are double-clustered by stock and quarter.

The explanatory variables are: *standard deviation of uninformed flow*, the standard deviation of the net change in position of uninformed traders during the interval (*uninformed traders* are traders who meet no criteria for informedness during the quarter); *illiquidity*, the relative spread implied by the first principal component of stock relative spread, inside depth, five-minute per-trade price impact, and 30-minute Kyle's lambda; *return volatility*, the standard deviation of five-minute returns during the observation interval; *uninformed volume*, the volume of uninformed trading during the interval; and *the standard deivation of uninformed volume*, the standard deviation of changes in the uninformed volume during the interval.

	Half-hour observations		Daily observations		Weekly observations	
	in levels	in z-scores	in levels	in z-scores	in levels	in z-scores
σ(Δuninf. flow)	0.621***	0.287***	1.167**	0.248***	2.388*	0.208***
	(5.07)	(45.66)	(2.84)	(31.37)	(2.69)	(31.98)
Illiquidity	-1308.6**	-0.0324***	-291.0	-0.0493***	6122.4	-0.0340***
	(-2.90)	(-8.23)	(-0.10)	(-9.59)	(0.56)	(-6.62)
Return volatility	926079.1***	0.142***	73179.5	0.166***	410579.6	0.186***
	(5.95)	(20.10)	(1.82)	(20.32)	(1.53)	(17.33)
Uninf. volume	0.197***	0.424***	0.217***	0.470***	0.220***	0.449***
	(7.93)	(28.36)	(6.20)	(31.97)	(4.86)	(29.36)
σ(Δuninf. volume)	-0.309***	-0.325***	-0.689***	-0.300***	-0.577**	-0.185***
	(-7.23)	(-25.43)	(-5.38)	(-30.18)	(-3.20)	(-20.24)
Constant	-42019.3	0.00628***	1401271.7	-0.0000356	1987294.9	0.00395***
	(-0.61)	(13.43)	(1.80)	(-1.62)	(0.37)	(22.80)
Observations	5382158	5382147	473031	473030	95663	95661
R2	0.364	0.203	0.609	0.247	0.693	0.281
Within	0.181	0.202	0.253	0.247	0.298	0.280

t statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$