High Price Impact Trades in Dierent Information Environments: Implication for Volatility and Price Eciency

Georges Dionne
HEC Montréal - École des Hautes Études commerciales

High Price Impact Trades in Different Information Environments: Implication for Volatility and Price Efficiency

Georges Dionne[a,b] Xiaozhou Zhou[a,c*]

[a]Canada Research Chair in Risk Management

[b]HEC Montréal

[c]Faculty of Management (ESG), University of Quebec at Montreal

Abstract

We include Limit Order Book (LOB) matchedness as an extended identification of High Price Impact Trades (HPITs). HPITs are trades associated with large price changes relative to their volume proportion, and are related to informed trades. We show that this modification provides finer informed trading identification, and that a stronger presence of HPITs leads to a decline in volatility due to more contrarian informed trades, but this decline varies with information environments. We conclude that HPITs mainly reflect a price correction to initial price (price convergence to equilibrium) of informed trades in a high (low) public disclosure environment.

Keywords: Matched trades, Limit Order Book, Trade size clustering, Price efficiency, Price discovery.

JEL classification: C22 C41 C53 G11

*Corresponding author at: Faculty of Management (ESG), University of Quebec at Montreal (UQAM), 315 rue Sainte-Catherine Est Montreal, Quebec, Canada, H2X 3X2. Tel.: +1 514 987 3000 ext 0999. E-mail addresses: georges.dionne@hec.ca (G. Dionne), zhou.xiaozhou@uqam.ca (X. Zhou).

1 Introduction

Financial markets inherently feature information asymmetry whereby some traders heterogeneously possess superior private information regarding companies' fundamental value. These fundamental-informed traders facilitate the price discovery process through two important reinforcing channels, either by making the security price reflect new private information or by correcting the price distortion caused by uninformed liquidity shocks. Obvious examples include strategically buying or selling the corresponding stocks according to the potential impact of their value-relevant private information, and bringing price back to its fundamental value by trading against uninformed traders. Given the growing importance of HFT (high-frequency trading), another type of informed trader has emerged: belief-based informed traders, who are able to learn ex-post about who might have fundamental-related information (O'Hara (2015)). In practice, belief-based informed traders could also be financial intermediaries and be informed ex-ante.[1] Belief-based informed traders get rewarded by following fundamental-informed traders and trading against uninformed traders.[2] The presence of belief-based informed traders makes the price discovery process more competitive and increases the non-execution risk of limit orders. Consequently, the profits of fundamental-based informed traders are reduced, and informed

[1]Financial intermediaries might be traditional or sophisticated high-frequency market makers who have access to both public information and information related to their customers. The migration to the electronic order-driven market and the use of high-performance computers have driven the rise of high-frequency trading and market making in financial markets. Today, sophisticated market making algorithms can quickly learn about trading motivations by following the footprints that other traders leave in the market. For example, Korajczyk and Murphy (2018) and Van Kervel and Menkveld (2019) show that it takes some time for high frequency market makers to detect fundamental-based trades. Typically, they initially lean against informed traders, but ultimately, they can identify the most informed traders and trade with the wind. In addition, both traditional and high frequency market makers can exploit information related to their customers. Having their customers' whole trading history, market makers possess at least superior private information about the performance of their customers and are in fact well positioned to judge if someone is informed or not. In addition, they often possess information about other brokers. Even though individual brokers cannot identify the exact identity of the traders behind other brokers, they still have a nonfundamental informational advantage. Therefore, belief-based (non-fundamental) information lets market makers better assess if the ongoing trades originate more informed or uninformed traders. Note that, in practice, market makers were not allowed to trade on the information derived from their clients, but enforcement was extremely difficult.

[2]Throughout the paper, we use informed trades to refer to the trades derived from fundamental-based and belief-based traders. Specifically, our definition of fundamental-based traders is in the spirit of the study by Goldstein and Yang (2015), which supposes that traders are heterogeneously informed about one dimension of fundamental values. Further, we classify activists' information as fundamental-related information even though the fundamental value changes when activists purchase a significant number of shares. By definition, activists are outside shareholders who identify a firm for potential value creation through their effort expenditure (Collin-Dufresne and Fos (2015)). Before activists publicly announce their intention to influence corporate governance, they attempt to purchase a significant number of shares in the open market. In contrast, belief-based traders differ from fundamental-based traders by the information they use. Typically, they exhibit reactive trading behavior that acts on information about trades by other traders. Belief-based traders profit when they correctly identify the trading motivations of other investors.

traders (fundamental-based and belief-based) are more likely to submit market orders. In real financial markets, there are also uninformed liquidity traders (institutional or retail),[3] which trade for exogenous discretionary reasons, and decide on which financial market in which to participate according to liquidity costs and underlying information environments (Han et al. (2016)). The proportion of uninformed traders in a financial market determines the total profit for fundamental-based and belief-based informed traders and market makers, and a higher proportion of uninformed traders should attract more belief-based informed traders. Together, informed traders (fundamental-based and belief-based) and uninformed traders constitute the financial markets. The interaction between them creates the dynamics of short-term security prices. How to accurately identify informed trades and recognize which channel (price convergence channel or price correction channel) is used by informed traders to reinforce price discovery, along with informed traders' reactions under different information environments and their implications for market quality, remain important open questions in the empirical finance literature.

To address these questions, we integrate the literature on identification of informed trading, endogenous information acquisition in the financial market, and the impact of informed trading on market quality in different information environments.[4] Our study is carried out in three steps. First, we take indirectly the information embedded in the LOB as a new dimension to identify high price impact trades, and relate them to informed trades. Second, we identify the channel through which these high impact trades reinforce price discovery by empirically evaluating their effect on short-term volatility and return autocorrelation for the stocks traded in different information environments. Third, we provide evidence that, in various information environments, the implications of high price impact trades for price efficiency differ.

Our empirical study uses stocks from three indexes of the Deutsche Börse: DAX, MDAX, and SDAX, which are traded at the Frankfurt Stock Exchange (FSE). The distinction between DAX, MDAX, and SDAX provides a natural setting in terms of information environments to study the role or impact of information-related trades on intraday price formation, volatility and price

[3]In the market microstructure literature, these uninformed trades are defined as discretionary liquidity trades, and have been studied by Admati and Pfleiderer (1988) and Foster (1990). We follow the literature and use the terms "noise trading," "liquidity trading," and "uninformed trading" interchangeably.

[4]Market quality is usually measured by volatility and price efficiency, which assesses how well the price reflects the information that is relevant to the asset's fundamental value. We use price efficiency, market efficiency and informational efficiency interchangeably.

efficiency. Specifically, the DAX index is composed of the 30 largest German companies in terms of market capitalization and Limit Order Book volume (LOB liquidity), which is widely considered as a blue chip market index for the German stock market. In addition, DAX stocks have a large number of analysts, lower forecast dispersion among analysts and smaller forecast error.[5] Based on these facts, it is fair to conclude that 1) the stocks in the DAX index are traded in a high public information environment with high liquidity. 2) DAX stocks are attractive for uninformed liquidity traders because of a lower expected loss of liquidity (Han et al.(2016)). In contrast, the MDAX index is composed of medium-sized companies, which feature a smaller LOB volume, fewer analysts, higher forecast dispersion and larger forecast error. Finally, the SDAX index includes small-sized companies with the smallest LOB volume, the fewest number of analysts, the highest forecast dispersion and the largest forecast error. Compared with large-cap DAX stocks, both MDAX and SDAX stocks are traded in a low public information environment and are less attractive for uninformed liquidity traders because of a higher liquidity cost.

To identify high price impact trades in different information environments, we start by showing the existence of *roundedness* clustering of trade volumes and *matchedness* to the quantity standing in the open LOB. Roundedness clustering either in price or trade size has been a well-recognized phenomenon in many markets (Harris (1991), Grossman et al. (1997), Alexander and Peterson (2007), and Hodrick and Moulton (2005)). Here, we conduct a distinct test of roundedness clustering from a market where the round-lot of a transaction is one unit, in which traders are free to choose their size for both market and limit orders. Our empirical results show that DAX, MDAX, and SDAX stocks feature trade-size clustering on 10, 50, 100 or their multiples. Rounded trades are significantly present in all three indexes and amount to almost 40% of the total trading volume.

Another important, but hardly explored, dimension to distinguish trades is matchedness. With the development of the trading system from the quote-driven market to the order-driven market, trading in financial market has become more transparent and faster than ever. Trading is now fast for all market participants. Even traditional non-high frequency traders have become fast because of their brokers' automatic routing systems, whose latencies for sending a market order directly

[5] The detailed calculations of these variables are presented in Section 3, and the corresponding results are reported in Table 2.

to the exchange are now below one millisecond (O'Hara (2015)). These high-speed traders can continuously observe the dynamics of the open LOB and time their order submissions. Despite the importance of the LOB in order submission strategy and price formation, the role of LOB-based trades has been largely ignored in the literature. In this study, we define a trade as a matched trade when its trade size matches the exact quantities available in the LOB.[6] More specifically, a matched buy (sell) trade is a buy-initiated (sell-initiated) trade that matches the exact cumulative quantity standing on the ask (bid) side of the open LOB. Matched trades do mechanically increase midquote by at least one tick, but they do not increase (decrease) buy (sell) transaction price, which is widely used as a reference price. As shown in Figure 1, the level-1 matched buy market order with 400 shares and the level-1 unmatched buy market order with 100 shares have the same transaction price and price impact, but the former profits from the most beneficial price to the greatest extent. The rationales for using matchedness to identify informed trading are: (i) matched trades largely originate from sophisticated HFT algorithms that could be informed by their high capacity for information processing (Van Kervel and Menkveld (2019)); and (ii) matched trades profit, to the greatest extent, from the most beneficial price available on the market, which could reflect the responsiveness of liquidity demand of informed traders with respect to liquidity supply. In contrast, liquidity traders seek to buy or sell a desired quantity of shares within a given time interval (from one day to several days) and often take VWAP (volume-weighted average price) as their benchmark price. To this end, the optimal strategy involves spreading their trades throughout the whole trading session and matching the volumes associated with these trades to the historical intraday trading volume pattern. Thus, liquidity traders are less sensitive to the exact quantity available in a given level and are more likely to submit unmatched market orders.

[Insert Figure 1 here]

It should be noted that we make no claim that informed traders always use matched trades in their trading strategies. Rather, we view matchedness, coupled with size and roundedness, as a

[6] Because our dataset does not include fully hidden orders and the hidden part of an iceberg order, our matchedness is based on observable limit orders.

convenient criterion to distinguish informed and uninformed trades. To the best of our knowledge, this is the first study that uses LOB matchedness of trade to distinguish trades, and that links matchedness to high frequency informed trades. Our research provides two insights into matched trades: first, matched trades represent more than 50% of total trades for DAX and MDAX stocks, and more than 40% for SDAX stocks. The omission of matchedness in informed trading identification could result in misleading conclusions. For example, our results show that small-matched trades and small-unmatched trades do not contribute to the price discovery process in the same way. The same pattern holds for (un)rounded-matched trades and (un)rounded-unmatched trades regardless of the index, suggesting that informed traders (human or non-human) are wary of market liquidity and carefully monitor the information embedded in the LOB. Second, the information quality of matched trades, measured by the ratio of the cumulative price contribution of matched trades to their corresponding proportion of volume,[7] varies with trade size and information environment. Overall, these findings provide strong evidence that matchedness, as a complement to roundedness, is an important dimension in informed trading identification.

We then classify all trades of the stocks in a given index by size,[8] roundedness, and matchedness, and statistically identify the trade types that have a high price impact.[9] We combine all trade types with a high price impact to construct our high price impact trades (HPITs) measure, which, in general, lead to disproportionately large price discovery[10] relative to their proportion of volume. The HPITs are trades that have relatively larger price impact compared to non HPITs after controlling for the effects of volume. Our first main result is that roundedness and matchedness are important dimensions in HPIT identification, and DAX, MDAX and SDAX stocks do not have the same trade types in their HPITs. In addition, the information quality of HPITs, defined as the ratio of the price discovery of HPITs relative to their proportion of

[7]An alternative interpretation of the information quality of matched trades is the average increase in price contribution given a one-percent increase in HPITs.

[8]In this study, trades of each stock are classified as small, medium or large when the corresponding trade sizes are smaller than the 30th percentile, between the 30th and 70th percentiles, and larger than the 70th percentile of its own trade size distribution, respectively.

[9]For a given type of trade, we calculate the cumulative (aggregated) price changes derived from all trades within this type.

[10]Price discovery attributed to HPITs is the price change associated with HPITs expressed as a percentage of daily price change.

volume,[11] varies across different stock indexes.

Once HPITs are identified, we further analyze their price impact over time at different intervals. Our time series results confirm that the price impact of HPITs over time dominates that of non-HPITs; this relationship holds for all stocks in our indexes. Also, we investigate intraday price contribution dynamics and the corresponding volume (expressed as a percentage of daily trade volumes). For DAX stocks,[12] we show that the contribution to price discovery of HPITs dominates that of non-HPITs during the whole continuous trading session (73.64% for HPITs vs. 26.36 for non-HPITs), accompanied by a striking difference in volume proportion (16.77% for HPITs vs. 83.23% for non-HPITs). More importantly, the beginning-of-the-day (9 a.m.-10 a.m.) contribution of HPITs to the daily price changes is comparable to, but always higher than, that of non-HPITs (20% vs. 15%). However, after 10 a.m., 82.64% of total price discovery is made by HPITs. In addition, the intraday volume from non-HPITs features strong seasonality, while the intraday volume from HPITs is quite stable. This observation confirms our assumption that non-HPITs mainly reflect the trades from uninformed liquidity-chasing traders, and HPITs are largely made by information-based traders. Similar observations are also found for MDAX stocks. However, for SDAX stocks, HPITs contribute less to daily price discovery than non-HPITs do. This could be due to the fact that the extremely wide bid-ask spread of SDAX stocks impedes price correction and incorporation of private information from informed trades. Having found that our identification method can effectively filter out informed trading, we then explore the practicability of our identification method by conducting a simple high frequency identification exercise for all stocks in our sample. We verify that the time required to identify informed trades based on our identification method is, on average, less than two milliseconds. Given that the average trade duration for most liquid DAX stocks is around 9 seconds, our identification approach also exhibits enhanced operating efficiency.

To further explore the market implications of HPITs, we investigate the relationship between HPITs and short-term volatility. Specifically, if HPITs mainly originate from informed traders, we should observe a negative relationship between HPITs and volatility, as predicted by the

[11] Similar to the information quality of matched trades, the information quality of HPITs measures the average increase in price contribution given a one-percent increase in HPITs.

[12] For MDAX and SDAX stocks, we observe similar patterns.

models of Hellwig (1980) and Wang (1993). Our results show that a stronger presence of HPITs does lead to a decline in volatility for most stocks in our sample, which strongly confirms the existence of information content in HPITs and their role as price stabilizers. In addition, we find that this negative effect on volatility increases with stocks' information transparency.[13]

We next look into the rationale behind this difference in volatility decline for the stocks in the three indexes. Specifically, we decompose HPITs into contrarian and herding (driving) HPITs,[14] and conjecture that contrarian HPITs are more likely to be price correction informed trades whereas herding HPITs are more likely to be price convergence informed trades.[15] To date, the empirical literature has mainly focused on the behavior and implications of price convergence informed trades and mostly ignored the role and involvement of price correction informed trades and their relation with information environments. With the increasing prevalence of uninformed trading, especially in the stocks with high public information disclosure and small spread, information about traders' motivations could also have significant implications on volatility. The model of Campbell et al. (1993) shows that a change in stock prices could occur due to information or liquidity-driven pressure. When the price change is due to fundamental information, price reversals are unlikely. However, price reversals are likely when price changes are caused by liquidity-driven pressure. This is precisely what we find with contrarian and herding HPITs. First, contrarian HPITs have a negative effect on return autocorrelation for most DAX, MDAX, and SDAX stocks. Second, herding HPITs have little effect on return autocorrelation for DAX and SDAX stocks and a positive impact on return autocorrelation for most MDAX stocks. In addition, our results show that the negative effect of contrarian HPITs increases with stocks' information transparency (i.e., it is smallest for SDAX stocks and highest for DAX stocks), which seems to confirm the dominance of uninformed liquidity trading and price correction informed trading (fundamental-based and belief-based) in DAX stocks, and an incentive for learning more about the trading motivations of other traders, in line with the prediction of Banerjee et al.

[13] We follow the literature and use the terms "information transparency" and "public disclosure" interchangeably.

[14] Contrarian HPITs refer to HPITs that trade against the current price trend. In a similar way, herding or driving HPITs are HPITs that follow or drive the actual price trend, respectively. In this study, we use the terms "driving" and "herding" interchangeably. The recent study by Van Kervel and Menkveld (2019) defines herding trades as with-wind trades.

[15] Contrarian HPITs should, on average, reflect price correction informed trades, and include fundamental-based and belief-based trades. Accordingly, consider that fundamental-based traders are informed of good news and there is a simultaneous price decrease caused by selling pressure from uninformed liquidity traders. In this situation, both fundamental- and belief-based traders will trade against the current price decrease.

(2018).[16]

Finally, we evaluate the impact of HPITs on short-term price efficiency for stocks traded in different information environments. We follow the recent empirical literature on price efficiency (see, for example, Boehmer and Kelley (2009), Chaboud et al. (2014), Conrad et al. (2015), and Rosch et al. (2016)), and use the variance ratio (Lo and MacKinlay (1989)) and the absolute value of autocorrelation as efficiency measures, which capture different aspects of price efficiency explored in the theoretical literature (Diamond (1985) and Gao and Liang (2013)). Specifically, short-term variance and autocorrelation capture the dynamic dimension of price efficiency by examining how closely the price follows a random walk and how predictable the returns are. After controlling for various market conditions, such as liquidity and volatility, we show that the presence of HPITs makes prices more efficient for DAX stocks, has an insignificant effect on price efficiency for MDAX stocks, and makes prices slightly less efficient for SDAX stocks. Our explanation is that in a market setting where both information transparency[17] and liquidity are already high (DAX stocks), HPITs frequently trade against uninformed trades and correct price distortion. As a result, price efficiency improves rapidly. In contrast, for stocks with less public disclosure and higher private informed trading (SDAX stocks), fundamental-based HPITs could lead to price inefficiency (e.g., price overreaction) without being corrected by informed traders because of a high trading cost. In addition, given that SDAX stocks feature a much wider bid-ask spread, the use of midquote price, as a proxy for expected true price, in price efficiency computation is also questionable.

The rest of the paper is organized as follows. Section 2 presents a literature review and our empirical hypotheses. Section 3 describes our dataset and the Xetra trading system. Section 4 investigates the existence of trade size clustering and the importance of LOB-matched trades. Section 5 examines the importance of roundedness and matchedness in informed trading identification. Section 6 investigates which market conditions are associated with HPITs' incidence. Section 7 assesses the impact of HPITs on volatility and identifies the underlying channels. Sec-

[16] Their theoretical model concludes that in a transparent market, greater public disclosure can discourage private learning about fundamentals while encouraging information acquisition about the trading motivation of others.

[17] Our paper distinguishes information transparency, which makes information acquisition about the fundamental value cheaper, from market transparency, which refers to how much information about the trading process is available to traders.

tion 8 investigates the impact of HPITs on short-term price efficiency, and Section 9 concludes the paper.

2 Literature Review and Main Hypotheses

2.1 Informed Trading and Matchedness

In the theoretical models of Kyle (1985) and Admati and Pfleiderer (1988), informed, risk-neutral speculators endogenously take their price impact into account and trade strategically by spreading their trades over time and selecting the moments when market liquidity is high. Empirically, Barclay and Warner (1993) explore informed traders' choice of trade size, and are the first to propose and validate the well-known stealth trading hypothesis that informed traders concentrate their trades on medium sizes to conceal their information. They find that the cumulative stock-price change is due to medium-size trades. A generalized version of this hypothesis is that if informed traders are the main cause of convergence of the market price to the expected fundamental value, and if these traders concentrate their trades in certain specific types to hide their trading intentions, then most of a stock's cumulative price change should fall within these trade types. Consistent with Barclay et al. (1993), Chakravarty (2001) evaluates the stealth-trading hypothesis by further categorizing trade sizes by initiator (i.e., retail or institutional investors) and posits that institutions are informed traders. Several studies examine the link between stealth trading and trade clustering.[18] Alexander and Peterson (2007) analyze trade size clustering with data from the NYSE and the NASDAQ, and suggest that rounded medium trade sizes have a greater price impact than do unrounded trades. Hodrick and Moulton (2005) study trade size clustering in a rational expectations framework and argue that when many heterogeneous uninformed investors are present, an asset will be traded at an increasing number of

[18] Ball, Torous, and Tschoegl (1985) and Harris (1991) argue that while a more precise price that is mutually acceptable to both the buyer and the seller can be reached by continuing negotiations, the incremental benefit to each side decreases and the exposure of each side to reporting and price risk increases. As a result, clustering will occur as traders will seek to simplify the negotiation process. Another explanation is from a behavioral perspective. Wyckoff (1963) notes that traders think in round numbers and try to trade in round numbers. Niederhoffer and Osborne (1966) argue that the tendency of traders to prefer integers seems to be a fundamental and stable principle of stock market psychology. Ikenberry and Weston (2003) argue that price clustering may be a collective preference by investors to voluntarily trade at particular price levels in order to minimize cognitive processing costs.

distinct sizes as investors' desire to trade exact quantities increases. Similarly, Moulton (2005) uses the data from foreign exchange markets to test the hypothesis that there is less trade size clustering shortly before the end of calendar quarters because portfolio managers seek to align their portfolios more fully with their given objectives. Moreover, the study provides evidence that the price impact of order flow is greater when customers care more about trading precise quantities. Garvey and Wu (2014) examine quantity choice patterns across trading hours and show that traders submit more non-rounded order sizes and more order sizes overall leading up to a day's market close. Studies that use roundedness to classify trades include those of Cai et al. (2006), Menkhoff and Schmeling (2010) and Ascioglu et al. (2011). The increasing number of distinct trade sizes might be related to another important dimension: matchedness, which measures responsiveness of liquidity demand of informed or uninformed traders with respect to liquidity supply. In addition, given that the operations in financial market have become faster than ever before, matchedness provides insight into how high frequency traders are sensitive to liquidity and price impact. On one hand, continuous trading allows traders to react to new informational shock more quickly. On the other hand, infinite trading opportunities under continuous trading regimes lead to less aggressive liquidity provisions because of a higher risk of adverse selection risk. As a result, informed traders suffer severely from price impact in any period and must adopt a more sophisticated strategy design. We conjecture that in a high public disclosure environment, accompanied by a small spread, uninformed traders are not likely to place LOB sensitive trades to meet their given objectives such as VWAP. They are thus less concerned with matched trades. In contrast, informed traders, who are sensitive to both liquidity and price, are likely to submit matched (LOB sensitive) trades when correcting mispricing because matched trades profit, to the greatest extent, from the most beneficial price available on the market. In accordance with the existent literature and our conjecture, we are the first to consider matchedness as a trade attribute, and test the following hypothesis:

> *H1: The information quality of matched trades, measured by the ratio of cumulative price contribution of matched trades to their corresponding proportion of volume, should be higher for stocks with greater public disclosure.*

In this hypothesis high public disclosure stocks corresponds to the stocks with a large number of analysts, lower forecast dispersion among analysts and smaller forecast error.

2.2 Effect of Informed Trades on Price Volatility

Friedman (1953) argues that irrational investors destabilize prices by buying when prices are high and selling when prices are low, whereas rational speculators, by trading against irrational investors (buy when prices are low and sell when high), correct the deviation of prices from fundamentals and stabilize asset prices. Similarly, the noisy rational expectation models of Hellwig (1980) and Wang (1993) argue that volatility increases with uninformed or liquidity trading. Empirically, Avramov et al. (2006) document that the activities of both imitative and nonimitative investors have a significant effect on day-to-day volatility, although in different directions. At the intraday level, Blasco and Corredor (2017) examine the PIN measure and detect that informed trading is a price-stabilizing factor in heavily traded and highly capitalized stocks. Indirectly, using a monthly firm-level PIN measure and excess return, Lai et al. (2014) find a positive correlation between PIN and volatility in international markets. Recently, Collin-Dufresne and Fos (2016) extend Kyle's (1985) model to a case where noise trading volatility follows a general stochastic process, demonstrating that informed traders choose to trade more aggressively when uninformed trade volume is higher and price impact is lower. Therefore, we posit that:

> *H2*: *If HPITs are related to informed trades, they are likely to occur when 1) market liquidity is high, 2) trade intensity is high, and 3) market volatility is high, and reduce volatility. Their negative effect on volatility should be greater for stocks with greater public disclosure.*

2.3 Implications of Contrarian and Herding HPITs

Previous theoretical models follow Hayek's (1945) idea that prices aggregate fundamental-based information dispersed among market participants. In such markets, informed traders behave strategically and lead the market, as Kyle (1985) argues, and noise traders arrive in the market in an exogenous way. The introduction of noise trading in these models mainly aims to provide liquidity to informed traders, solving the "no trade" problem (Milgrom and Stokey (1982)), and imitating the real financial market. However, in practice, the proportion of uninformed traders depends on market information conditions and liquidity costs (Han et al. (2016)). Uninformed trading could also originate from financial institutions,[19] and cause a significant impact on short-term price formation (Cespa and Vives (2015)). Given the importance of uninformed trading in the financial markets, Han et al. (2016) propose an extended rational expectations equilibrium model in which the size of noise trading is endogenously determined by the population size of active liquidity traders in the market. They show that disclosing payoff-relevant public information weakens the information asymmetry problem, and lowers the expected loss of liquidity traders, thereby attracting more such traders. Under these circumstances, there should be more price correction related informed trading for stocks with greater public disclosure. Given the presence of both informed and uninformed (liquidity) traders, Campbell et al. (1993) argue that, in a quote-driven market, price reversals occur as risk-averse market makers absorb order flow from uninformed or liquidity traders. According to their model, a variation in stock price could be caused by informed or uninformed selling pressure. When price change is driven by information, price reversals are unlikely. However, if price change is driven by uninformed liquidity traders, price reversals are likely and liquidity suppliers should be compensated for taking this risk. With exogenous sensational news, Peress and Schmidt (2020) find that on distraction days, trading activity, liquidity and volatility decrease, and prices reverse less. Indeed, in an open limit order book market where everyone could be a liquidity supplier when they know or learn that the incoming orders originate from uninformed traders, especially for high learning capacity computers. Recently, Korajczyk and Murphy (2018) and Van Kervel and Menkveld (2019) show that it takes some time for high frequency traders to detect the fundamental-based trades, but ultimately they are able to identify the most informed traders and follow their trades. This allows us to test our third hypothesis:

[19]For example, Coval and Stafford (2007) find that in response to important redemption requests by clients, fund managers will curtail their positions and engage in "fire sales" for non-informational reasons. Additionally, Henderson et al. (2014) provide evidence that uninformed investment banks may also take large positions in commodity futures to hedge issuance of commodity-linked notes. Recently, with algorithmic trading data, Skjeltorp et al. (2016) find that algorithmic trading by large institutional investors is likely to be uninformed.

H3.1: *If contrarian HPITs mainly reflect price correction informed trading, they should have a negative effect on return autocorrelation, and this negative effect should be more pronounced for stocks with greater public disclosure because of the presence of HFT for these stocks.*

H3.2: *If herding HPITs are mainly related to price convergence informed trading, they should have an insignificant effect on return autocorrelation.*

2.4 Informed Trades and Short-term Price Efficiency

Theoretical models (Diamond (1985), Gao and Liang (2013), Colombo, Femminis, and Pavan (2014), Banerjee et al. (2018), and Dugast and Foucault (2018), among many others), deduce price efficiency as a static precision of the conditional expected price based on fundamental information. Further, a subset of these papers focuses on a "crowding out" effect: greater public disclosure about fundamentals can crowd out private information acquisition, which in turn can reduce price informational efficiency. However, short-term price dynamics and price predictability are largely ignored. We extend the recent empirical literature (Boehmer and Kelley (2009), Chaboud et al. (2014), and Rosch et al. (2016)) on price efficiency to test the effect of HPITs on short-term price efficiency:

H4: *HPITs increase price efficiency for stocks with greater public disclosure and high levels of liquidity.*

3 Xetra Trading System, Ultra High-Frequency and Information Environment Data

The data used in this study are from the Xetra trading system, which is operated by Deutsche Börse at the Frankfurt Stock Exchange (FSE); it has a similar structure to the Integrated Single Book of NASDAQ and Super Dot of NYSE. The Xetra trading system realizes more than 95% of the total transactions at German exchanges.

For highly liquid stocks, there is one dedicated market maker per stock and several sponsors during continuous trading. Similar to other stock trading systems, the Xetra trading system imposes a Price-Visibility-Time Priority condition, where the electronic trading system places the incoming order after checking the price and timestamps of all available limit orders in the LOB. Our database includes 20 levels of LOB information,[20] which means that, by monitoring the LOB, any registered member can evaluate the liquidity supply dynamics and potential price impact of a market order. However, there is no information on the identities of market participants.

The reconstruction of the LOB is predominantly based on two main types of data streams: delta and snapshot. The delta tracks all the possible updates in the LOB such as entry, revision, cancellation, and expiration, whereas the snapshot gives an overview of the state of the LOB and is sent after a constant time interval for a given stock. Xetra original data with delta and snapshot messages are first processed using the XetraParser algorithm, developed by Bilodeau (2013). XetraParser reconstructs the real-time order book sequence including all the information for both auctions and continuous trading by implementing the Xetra trading protocol and Enhanced Broadcast. Then the raw LOB information is put in order and in a readable format for each update time. Useful and accurate information about the state of the LOB and the precise timestamp of order modifications and transactions during continuous trading are also retrieved.

Our study focuses on the component stocks in three market indexes, DAX, MDAX, and SDAX, respectively. The DAX consists of the 30 major German companies listed on the Frankfurt Stock Exchange and is a blue-chip stock market index. MDAX includes 50 component stocks and is a stock index for the listed companies that rank below the companies in the DAX index in terms of market capitalization and order book volume (technology companies excluded). Finally, the SDAX is composed of 50 listed stocks that rank directly below the stocks in MDAX in terms of market capitalization and order book volume. There is a quarterly review to re-rank stocks among these three groups. Using data from the Compustat Global Security Daily files, Table 1 reports the descriptive statistics of daily market variables for DAX, MDAX, and SDAX stocks for

[20] Fully hidden orders and the hidden part of an iceberg order are not observable in our dataset. However, as we observe the state of the LOB before and after the transaction, we can evaluate if a market order hits hidden orders or not. Our backtest results show that fewer than 3% of the market orders run into hidden orders, which represents about 6% of trade volumes.

six months, from February 1, 2013 to July 31, 2013. A decreasing monotonic trend is observed, from DAX to SDAX stocks, for all variables.

[Insert Table 1 here]

The high frequency LOB and trade data used in this study are registered with a timestamp in microseconds. Such precision allows us to identify the state of the LOB just before the trades and to determine whether the trade sizes are perfectly matched to the quantity standing in the open LOB. Table 2 presents the descriptive statistics of trades and information environment variables for DAX, MDAX, and SDAX stocks. To measure the information environment, we use the number of monthly news mentions and three analyst measures as the proxies of the information environment. The monthly number of news is the number of times that the company is mentioned in the news and social media registered by RavenPack. The analyst data are extracted from Institutional Brokers' Estimate System (I/B/E/S) for the period of 2011 to 2015.[21] We take the annual earnings per share (EPS) announcement as our target event. Following Barron et al. (1998), we first compute the number of analysts making forecasts about annual EPS up to the firm's actual announcement date. Second, we focus on the earning forecast dispersion, measured by the standard deviation of the forecasted EPS, standardized by the share price at the beginning of the year (Barron and Stuerke (1998) and Johnson (2004)). The third measure is the forecast error, defined as the absolute difference between the mean forecast EPS and actual EPS, standardized by the price at the beginning of the year (Rajan and Servaes (1997) and Gu and Wang (2005)).[22] We compute these three measures for all stocks and present the mean, median and standard deviation of three groups in Table 2. For a group with more transparent information environment, we expect to observe a larger number of analysts, lower forcast dispersion between analysts and smaller forecast error. It follows from Table 2 that DAX stocks are traded in a high public disclosure environment with low trading costs and high market transparency, while MDAX stocks are traded in a lower disclosure environment with higher trading costs and lower market transparency. Lastly, SDAX stocks feature the lowest disclosure environment and the

[21]To obtain precise and reliable comparison results of different information environments, we select a more extended sample period of analyst data than that of our high frequency transactions and LOB data.

[22]For forecast dispersion and forecast error, we use the last forecast variables before the actual announcement.

highest trading costs.

[Insert Table 2 Here]

4 Cross-sectional Analysis of Trade Size Clustering and LOB-Matched Volume

We first investigate the existence of rounded size clustering in DAX, MDAX and SDAX stocks, and the importance of LOB-matched trades.

4.1 Trade Size Distribution and Presence of Clustering

We start by calculating trade size probability distributions for each stock in the three indexes, and then compute the average probability for each trade size across stocks in the same group. As mentioned above, the round-lot in the Xetra system is one unit; therefore, any integer number may appear in our sample. Figure 2 displays the histogram of trade sizes for DAX stocks (MDAX and SDAX stocks exhibit similar patterns). Three main insights arise from the figure: First, we observe a decreasing trend, which suggests that the larger the trade size, the more infrequently it occurs. Second, the histogram features clustering on rounded numbers. For instance, there are more trades at 20 shares than 19 or 21 shares. Third, the clustering takes place in three levels. In increasing intensity, these levels are multiples of 10, 50 and 100. Specifically, when a trade size is a multiple of both 10 and 50, it will occur more frequently than a trade size that is a multiple of 10 only. For example, the trades at 150 shares occur more often than trades at 140 shares. The same logic applies to trade sizes that are multiples of 10, 50 and 100.[23]

[23] Our analysis does not rule out the possibility of the multiples of 10, 50 and 100 from high-frequency trading. We assume that traders choose roundedness when submitting their orders.

[Insert Figure 2 here]

To test if our observations are statistically significant, we follow Alexander and Peterson (2007) and estimate the following regression separately for the three stock groups:

$$LnFreq_i = \alpha + \beta_5 D5_i + \beta_{10} D10_i + \beta_{50} D50_i + \beta_{100} D100_i + \beta_{LnSize} LnSize_i + \epsilon_i, \quad (1)$$

where $LnFreq_i$ is the natural log of the percentage of trades of size i. $D5_i$, $D10_i$, $D50_i$, $D100_i$ are dummy variables if trade size i is a multiple of 5, 10, 50, 100, respectively, and $LnSize_i$ is the natural log of trade size i measured in numbers of shares.

Panel A, B, and C of Table 3 reports the regression results of trade size clustering for DAX, MDAX, and SDAX stocks, respectively. The coefficients of dummy variables of 5, 10, 50 and 100 are all significant at the 5% level except the dummy variable of size 5 for DAX stocks. The coefficient of $LnSize$, which is significantly negative, confirms the negative slope of the histogram in Figure 2. In addition, a high value of adjusted R^2 suggests that the trade size pattern can be largely explained by trade size clustering and negative slope.

[Insert Table 3 here]

Table 4 presents the cross-sectional mean and standard deviation of clustered size across stocks and days. For each day, on average, the clustered trade-size percentage is significant and stable with a mean of 33.39% and standard deviation of 3.97% for DAX stocks. Similar to the daily cross-sectional mean and standard deviation, the clustered rounded trade percentage remains stable across stocks, with a mean of 33.40%, and standard deviation of 2.84%. The same patterns are found for MDAX and SDAX stocks. The intuition is that traders' choices in terms of roundedness are consistent and remain stable among different groups.

[Insert Table 4 here]

4.2 LOB-matched Trades

In this subsection, we turn to another important dimension of trades, the matchedness of trade size to the quantities available in the LOB. A transaction is initiated by either the buy side or sell side. However, the counterparts of transactions are the limit orders standing in the open LOB.[24] With the development of information technology, the speed of submitting an order has become faster than ever before. For instance, Xetra implemented co-location service that allows traders to connect to the central server with much less latency (13 microseconds). Thus, with this speed advantage, traders can match the exact quantities standing in the open LOB when submitting a market or marketable orders. Note that a rounded trade size might be matched or unmatched depending on the quantity available in the open LOB. The dimension of matchedness is important because it provides insight into high frequency traders and their sensibility to liquidity. When traders have a speed advantage, they can time their trades and match the exact quantities available in the open LOB to profit, to the greatest extent, from the most beneficial price. According to Kyle (1985), informed traders camouflage themselves by splitting large volumes into small ones. Compared with the traders who passively split their desired trade volume, well-equipped high-frequency traders can now actively choose their submission time and trade the most favorable quantity available for them (quantity available at ask or bid side of the LOB). For our LOB-matched-trade identification, one might argue that the marketable orders can also give the illusion of an LOB-matched market order. As we show in Figure 3, a marketable bid (ask) order will both match the exact quantity in the ask (bid) side of the open LOB and increase (decrease) the best bid (ask) price to the price of the matched level. However, a simple buy (sell) market order will only consume the quantity standing in the ask (bid) side without creating a new best bid (ask) price. Therefore, to rule out the marketable orders, we also check the state of the LOB after the transaction to guarantee the accuracy of our matchedness identification.

[Insert Figure 3 here]

Table 5 reports the cross-sectional mean and standard deviation of matched trade sizes calculated

[24] This is not the case for the hidden orders which represent less than 5% of total trades in our sample, suggesting that our results are still robust.

at a daily level. For DAX stocks, the percentage of LOB-matched size is stable with a mean of 52.34%, and a standard deviation of 5.76%. At the stock level, the average percentage of LOB-matched size is 52.34% with a standard deviation of 4.59%. For MDAX and SDAX stocks, we find similar cross-sectional and time-series patterns: the percentage of LOB-matched size exhibits a stable pattern for groups with smaller trade sizes. It is interesting to see that more than half of the trade sizes seek to match the exact quantity available in the LOB to avoid a price impact. In addition, Table 6 reports LOB-level distribution of matched and unmatched trades and their corresponding volume proportions for DAX, MDAX, and SDAX stocks. Typically, most matched and unmatched trades take place at the first level of the open LOB. However, for stocks with greater public disclosure, there are more matched trades than unmatched trades (52.84% Vs. 47.16%), and the opposite is true regarding stocks with less public disclosure (42.07% Vs. 57.93%). It should be noted that a medium- or large-sized matched trade can also take place at the first level of the open LOB.

[Insert Table 5 here]

[Insert Table 6 here]

4.3 Clustered and LOB-matched Trades

Given the patterns and features of the rounded and LOB-matched trades, we consider these two dimensions together and investigate their dynamics across stocks. To do so, we consider the percentage of volumes of four trade types across stocks: UR-UM,UR-M, R-UM, R-M.[25] As shown in Figure 4, the percentage of each trade size category is quite stable across stocks. Based on these observations, our interest now lies in the contribution of these trade types to price discovery, which is examined in next session.

[25]The corresponding trade types are as follows: unrounded and unmatched trades (UR-UM), unrounded and matched trades (UR-M), rounded and unmatched trades (R-UM), and rounded and matched trades (R-M).

[Insert Figure 4 here]

5 Trade Types, HPITs, and Price Discovery

5.1 General Framework

In order to address the issue of the contribution of different trade types to the price discovery process,[26] we follow the rationale that if informed trades are the main cause of stock price changes and concentrate their trades in specific trade types, then most of a stock's cumulative price change should take place on these trade types. After the first empirical work of Barclay et al. (1993), an extensive body of literature attempts to identify informed trading with various datasets from different markets. For exemple, the studies categorize trade type by whether the trades are initiated by institutions (Chakravarty (2001)) or whether the trades feature odd-lot sizes (O'Hara et al. (2014)). With our comprehensive dataset on the transaction and the open LOB, we can further evaluate how important the roundedness, and/or matchedness of trade is in informed trading identification. Also, we take trade size as our first dimension when distinguishing the trades. Trade-size is expressed in relative terms and defined as small, medium and large. The critical values used to categorize the different groups are the 30th and 70th percentiles of trade sizes.

We begin by examining the contribution of the rounded and matched trades separately, and then we consider these two dimensions together to get a clear insight into the price discovery contribution of each trade type. As in the study by O'Hara et al. (2014), we suppose there are N trades for stock s for day t, and each trade can be categorized into one of J groups. In addition, we define the contribution of a given trade as the log difference between the trade's price and the price of the previous transaction.[27] The cumulative price contribution of the trades belonging

[26] We consider the daily price change as the total daily price discovery and take the ratio of the cumulated price change associated with a given category over the full price discovery as the contribution of that trade category.

[27] Our method of computing the price contribution follows Barclay and Warner (1993) and Chakravarty (2001). By this definition, we also ascribe the contribution of the LOB between two consecutive trades to the second trade.

to category j for stock s on day t is defined as

$$PC_j^{s,t} = \frac{\sum_{n=1}^{N} \delta_{n,j} r_n^{s,t}}{\sum_{n=1}^{N} r_n^{s,t}}, \tag{2}$$

where $\delta_{n,j}$ is an indicator variable that takes the value of one if the nth trade falls into size category j, and zero otherwise. Following Barclay and Warner (1993), we weigh each stock's price contribution to mitigate the problem of heteroskedasticity, which may be severe for firms with small cumulative changes. Suppose there are N trades for stock s on day t. The weight for stock s on day t is defined as

$$w^{s,t} = \frac{\mid \sum_{n=1}^{N} r_n^{s,t} \mid}{\sum_{s=1}^{S} \mid \sum_{n=1}^{N} r_n^{s,t} \mid}. \tag{3}$$

The weighted price contribution of trades in size category j on day t is defined as

$$WPC_j^t = \sum_{s=1}^{S} w^{s,t} PC_j^{s,t}. \tag{4}$$

Suppose there are T days in total. The weighted price contribution of trades in size category j is defined as

$$WPC_j = \frac{\sum_{t=1}^{T} WPC_j^t}{T}. \tag{5}$$

To further quantify the contribution of different trade types to daily price discovery and identify HPITs, we estimate the following regression:

$$PC_j^{s,t} = \sum_{j=1}^{k} \alpha_j \times dummy_j + \beta \times PcntVolume_j^{s,t} + \epsilon_j^{s,t} \tag{6}$$

where $dummy_j$ is the dummy variable for category j and $PcntVolume_j^{s,t}$ relates to the volume percentage for stock s on day t in category j. The dependent variable is the cumulative price change of all trades in a given trade type divided by the cumulative price change of all trades during the period. The magnitudes of the denominator affect the level of this variable's sensitivity, especially for stocks with small cumulative price change. The weighting procedure can significantly reduce heteroskedasticity in the dependent variable.

5.2 Trade Types and Price Contribution

Using all transactions from all stocks, we first report the weighted price contribution associated with trade size as in Barclay and Warner (1993). Then we extend our analysis to the weighted price contribution of roundedness and matchedness[28] separately. Finally, we provide a more detailed analysis by jointly considering trade size, roundedness and matchedness. To statistically identify HPITs, we estimate equation (6) with dummy variables of all 12 trade types (3 sizes $\times$ 2 roundedness $\times$ 2 matchedness).

We take size as the first dimension for our trade classification and analysis of weighted price contribution. Specifically, for each stock, trades are first classified as small-, medium- or large-size according to the 30th and 70th percentiles of its own trade size distribution. Then, the corresponding proportion of trades and volumes, and price contribution are computed. Table 7 summarizes the results aggregated by size dimension for the DAX, MDAX and SDAX stocks. It follows that in our dataset it is the small-size trades that are associated with disproportionately large price changes relative to their proportion of volume. However, using a data sample between 1981 and 1984, Barclay and Warner (1993) find that medium-size trades are the trades associated with disproportionately large price changes relative to their proportion of volume. The difference between their findings and ours suggests a migration of informed trades from medium-size to small-size trades. One explanation is that trading cost decreases in financial markets. Informed traders always have to trade off between the gains related to their private information and the costs associated with the trading implementation. In previous quote-driven

[28]At this step of analysis, we keep size (e.g., small, medium, or large) as our first dimension.

markets, traders paid, for each transaction, a high order processing cost charged by financial intermediaries. Thus, the practice of cutting large orders into small ones is very costly. However, the transformation from quote-driven to order-driven market and the introduction of electronic trading reduce dramatically this order processing cost and give informed traders an incentive to place more small orders. One may argue that the decrease in trading cost also gives incentive to liquidity traders to cut their orders. However, note that liquidity traders trade for exogenous discretionary reasons and are less sensible to trading cost decrease. In fact, they are ready to pay for liquidity. Therefore, we can safely link the liquidity cost decrease to the migration of informed trades from medium-size to small-size trades.

To extend our analysis of trade types to roundedness and matchedness, and further examine their corresponding marginal effects, we conduct a two-dimensional analysis by combining separately roundedness or matchedness with size (size-roundedness analysis vs. size-matchedness analysis). Table 8 shows the weighted price contributions of various trade types and the corresponding information quality for the DAX, MDAX and SDAX stocks. It is important to note that Table 8 presents the same findings as Table 7 does, but at a distributive level. To see this, consider that the small-size WPC for DAX in Table 7 (12.53%) is the sum of the small-unrounded and small-rounded WPC for DAX in Panel A1 of Table 8 ($22.26\% - 9.73\% = 12.53\%$). In this typical example, we already notice that the small-rounded trade and small-unrounded trade do not have the same contribution to price discovery process. Therefore, considering all small-size trades in the same way without making any further distinction could be misleading. Panel A1 of Table 8 reports the price contribution of trade types based on size and roundedness for the DAX stocks, and suggests that the small-unrounded trade disproportionately contributes to the cumulative price change.[29] Trades in the medium and large size, regardless of their roundedness, are less informative to the WPC. For instance, large-unrounded and large-rounded trades are responsible for 35.68% and 6.19% of the cumulative price change, respectively. However, these price changes also correspond to 20.03% and 9.71% of the total numbers of transactions, and 46.96% and 21.58% of the total trading volumes. Panels B1 and C1 report the price contribution of different trade types classified by size and roundedness for the MDAX and SDAX stocks. Similar to DAX

[29] By taking the ratio of the WPC of a given category over the corresponding weight of such category in the total volumes and number of trades, we examine which category disproportionately contributes to the cumulative price change.

stocks, small-unrounded trades contribute more to daily price discovery.

Panels A2, B2, and C2 of Table 8 present the price contribution of different trade types according to size and matchedness. The results confirm our first hypothesis that in a high public disclosure environment, as liquidity is already high, uninformed traders are likely to place more unmatched trades to meet their given objectives, and informed traders, who are sensitive to both liquidity and price, are likely to submit LOB-matched trades when correcting mispricing. However, in a low public disclosure environment, accompanied by a liquidity shortage, uninformed traders care more about liquidity and are likely to submit matched trades, and informed traders are likely to submit unmatched trades (e.g., marketable trades). Specifically, for DAX stocks (Panel A2), the information quality of matched trades (1.17) is higher than that of unmatched ones (0.82). The opposite is true for MDAX (0.66 for matched trades vs. 1.37 for unmatched trades) and SDAX (0.68 for matched trades vs. 1.23 for unmatched trades) stocks. However, for all three indexes, we observe a migration of informed trades from matched trades to unmatched ones when trade sizes increase. Take DAX stocks (Panel A2) as an example, with only about 2.68% of the total trade volumes, small-matched trades produce about 17.14% of the cumulative price change, which amounts to an information quality of 6.39. However, small unmatched trades, representing 2.73% of the total trade volumes, result in a WPC of -4.61%.[30] For medium-size trades, matched and unmatched trades have almost the same information quality. While for large-size trades, the information quality of matched trades is dominated by that of unmatched ones. Panels B2 and C2 report the price contribution of different trade types according to size and matchedness for MDAX and SDAX stocks. We observe slightly different but consistent results: 1) the information quality decreases when trade size increases. 2) Even for the small-size group, information quality of matched trades does not dominate any more that of unmatched ones. Overall, our results suggest that matchedness, as a complement to roundedness, is an important dimension in informed trading identication.

In sum, the two-dimensional results of size-roundedness and size-matchedness provide evidence that both are important dimensions in informed trader identification. Further, informed trades are mainly associated with small-size trades, which is consistent with the stealth trading hypoth-

[30] The negative WPC indicates that on average the associated trades run in the opposite direction to price.

esis.

[Insert Table 8 here]

Now, we analyze WPC with all three dimensions: size, roundedness, and matchedness. Table 9 illustrates the WPC of the total of 12 trade types for DAX, MDAX and SDAX stocks. When we compare the ratio of the WPC to its corresponding weight in total trade volumes, the most informative trades for DAX stocks are small-unrounded-unmatched, small-unrounded-matched, small-rounded-matched, and medium-rounded-matched. Surprisingly, rounded-unmatched trades, regardless of their sizes, contribute negatively to cumulative price changes. Our conjecture is that for the DAX stocks, rounded-unmatched trades are predominantly initiated by liquidity-chasing uninformed traders. The results show that the high level of granularity in our trade type analysis is important in informed trading identification. For example, without making distinction between rounded-matched and rounded-unmatched trades, Alexander and Peterson (2007) conclude that US stock markets feature size clustering and medium-rounded trades are more informative ones. For MDAX stocks, the results are similar, except we now find that the small-sized trades, regardless of trade roundedness and matchedness, contribute positively to daily price discovery, and, more importantly, medium-rounded-matched and medium-unrounded-unmatched trades also contribute positively and disproportionately to price discovery. Again, for large-size trades, regardless of their roundedness and matchedness, there is no significant price contribution. For SDAX stocks, we have similar results as MDAX stocks, except that for medium-size trade, only medium-unrounded-unmatched trades make a signficiant contribution to price discovery. In summary, we show that several types of trade do contribute significantly to the daily price discovery process. This implies that informed traders choose different trade types to reveal their information according to the level of information disclosure.

[Insert Table 9 here]

5.3 HPIT Identification

In previous section, we show that the inclusion of matchedness and roundedness in HPITs identification is important, and trade types of HPITs may also change across different indexes. We now turn to statistically identify HPITs for stocks in the three market indexes. To do so, we conduct the weighted-least-squares regressions of the percentage of the cumulative price change occurring in each trade type on dummy variables for the trade types and the percentage of volumes associated with that trade type (equation (6)). If there is no significant contribution of a given trade type to the daily price discovery process, the coefficient of the corresponding dummy variable should not be significantly different from zero.[31]

In Table 10, we present the regression results of the 12 trade types, classified by trade size, roundedness and matchedness, for the stocks in the three market indexes. A positive and significant coefficient for a dummy variable means that the price change related to such type moves in the same direction as a daily price change, while a negative and significant coefficient implies that the price change related to such type moves against the daily price change. For DAX stocks, we show that the coefficients of five dummy variables are significanlly different from zero at the 1% level.[32] The results imply that informed trades in DAX stocks concentrate in various small- and medium-size trade types. In addition, rounded-unmatched trades, regardless of their sizes, do not have contribution to daily price changes, which confirms what we observed in Table 9. As for large-size trades, regardless of roundedness and matchedness, they do not contribute to the price discovery process. One explanation is that the decrease in the order processing cost encourages informed traders to split their trades into tiny pieces, and this is not the case for uninformed traders.

[Insert Table 10 here]

MDAX stocks have six positive and significant informed trade types which include the same five

[31] Note that our analysis focuses on the cumulative price changes of all trades in an associated trade type. Therefore, an estimated coefficient not significantly different from zero for a given type does not mean that there is no informed trades at all in such trade type. Instead, the associated trade type is dominated by uninformed trades.

[32] They are dummy variables for trade types of small-unrounded-unmatched, small-unrounded-matched, small-rounded-matched, medium-unrounded-unmatched, and medium-rounded-matched.

informed trade types as DAX stocks and small-rounded-ummatched trades. As for SDAX stocks, there are only four positive and significant informed trade types. Specifically, for small-size trade types, the small-rounded-matched is the only one that is not informative to daily price, and for medium-size trade types, the medium-unrounded-unmatched is the only one that contributes significantly to the daily price changes. Again, large-size trades, regardless of roundedness and matchedness, do not belong to HPITs for any of the three groups. After identifying high price impact trades, we further attempt to investigate the informational quality of HPITs. The information quality is defined as the ratio of the WPC to its corresponding weight in the total trade volumes.[33] Typically, we aggregate the trade types with positive and significant coefficients from the estimated results of equation (6) as aggregated HPITs. Then, we compute the corresponding aggregated WPC, the corresponding aggregated proportions in total trade numbers and total trade volumes, and take the number-based and volume-based information quality for the aggregated HPITs. Table 11 presents the number-based and volume-based information quality measures for DAX, MDAX and SDAX. One interesting insight is that both ratios decrease with market capitalization and public information disclosure, suggesting that HPITs for DAX stocks contribute more to the price discovery process. Specifically, on average, 1% of HPITs measured by volume contributes to 4.39% total price change for DAX stocks, 3.51% for MDAX and 3.00% for SDAX stocks. As a robustness check, we exclude HPITs that are followed by opposite midquote movements.[34] As shown in Panels a.2 and b.2 of Table 11, the informational quality of HPITs increases, and this increase is more pronounced for the least transparent stocks. These results confirm that competition among informed traders is more intense for DAX stocks, and that informed traders are likely to use market orders.

Note that each HPIT type typically does not exclude uninformed trades: both information-related and uninformed trades could be present in any trade type. However, an HPIT category is the group in which the price contribution of information-related trades dominates that of uninformed traders. Therefore, these results on informational quality suggest that the dominance of information-related trades in such trade types is stronger for DAX stocks and weaker for

[33] We use number-based and volume-based information quality to refer to these two information quality measures.

[34] One may argue that instead of submitting market orders, informed traders can also submit limit orders and wait for liquidity traders' market orders. In that case, after a buy-initiated (sell-initiated) trade, the midquote might decrease (increase).

SDAX stocks. In other words, market capitalisation, liquidity, and informational transparency are main factors to consider when informed traders choose trade types (size, roundedness, and matchedness) for their trades. For large-cap stocks exhibiting high liquidity and transparency, information-related traders are likely to submit small orders, whereas liquidity and uninformed traders are likely to use large-size orders because of a lower expected loss of liquidity (Han et al.(2016)). However, regarding stocks with a liquidity shortage, information-related and uninformed traders have to consider the price impact of their trades, which limits their choice of trades in terms of size, roundedness, and matchedness. As a result, information-related and uninformed traders of less liquid stocks are likely to submit similar market orders, which are more likely to be found in the same trade types.[35]

[Insert Table 11 here]

5.4 Price Impact of HPITs Over Time

After providing strong evidence that HPITs contribute significantly to daily price changes, we now turn our attention to the temporal price impact of HPITs. Specifically, we use effective price impact to compare the price impact of HPITs and non-HPITs over time:

$$Effective\,price\,impact_{i,t} = \frac{D_{i,t}(Midquote_{i,t+m} - Midquote_{i,t})}{Midquote_{i,t}}, \tag{7}$$

where $D_{i,t}$ is the direction of the trade for stock i at tth trade that has $D_{i,t} = +1$ if the trade is buy-initiated and -1 if it is sell-initiated, $Midquote_{i,t}$ is midquote of the best bid and ask before the execution time, and $Midquote_{i,t+m}$ is midquote m minutes after the execution time. Table

[35]Note that based on our regression results, trade types of HPITs vary with stock indexes. In what follows, DAX HPITs contain small-unrounded-unmatched, small-unrounded-matched, small-rounded-matched, medium-unrounded-unmatched, and medium-rounded-matched trades. MDAX HPITs are composed of small-unrounded-unmatched, small-unrounded-matched, small-rounded-unmatched, small-rounded-matched, medium-unrounded-unmatched, and medium-rounded-matched trades. Finally, SDAX HPITs include small-unrounded-unmatched, small-unrounded-matched, small-rounded-unmatched, and medium-unrounded-unmatched trades.

12 compares the price impact of HPITs and non-HPITs over time for DAX stocks after various time intervals. The empirical results suggest that the temporal price impact of HPITs dominates that of non-HPITs for stocks in three indexes, and that the difference is more pronounced for less transparent and illiquid stocks. The intuition is that less transparent and illiquid markets feature a lower proportion of noise traders, and that it is more difficult for informed traders to camouflage their trading intention in such markets, which results in a significant difference between informed and uninformed trading. The opposite holds for more transparent and liquid markets. Also, our results show that SDAX HPITs' price impact is higher than that of DAX HPITs, suggesting again a lower proportion of noise traders for SDAX stocks.

[Insert Table 12 here]

5.5 Intraday Dynamics of HPITs

Once we have identified HPITs, we further evaluate their hourly contribution to price discovery during the trading day, as shown in Panel A of Figure 5. We first compute the hourly price contribution by taking the ratios of hourly price change over daily price change, and decompose the resulting hourly price contribution into those associated with HPITs and non-HPITs. For DAX stocks, the contribution to price discovery of HPITs dominates that of non-HPITs during the whole continuous trading session. More specifically, between 9 a.m. and 10 a.m. (during the beginning of a trading session), the contribution of HPITs is comparable to, but still higher than, that of non-HPITs (20% vs. 15%). Strikingly, non-HPITs produce about 15% of the daily price change with around 13% of total trading volumes. In contrast, HPITs contribute 20% of overall price change, with only 2% of total trading volumes. Also, for the time bins after 10 a.m., 82.64% of total price contributions come from HPITs.[36] We examine another important factor in the identification of informed trading: the corresponding volumes for HPITs and non-HPITs. Panel B of Figure 5 shows the evolution of hourly trading volumes (expressed as a percentage of

[36]Table 13 reports that for the DAX index the total daily price contribution and the HPITs daily price contribution after 10 a.m are 53.14% and 64.30%, respectively.

daily trade volumes) for HPITs and non-HPITs. The trading volumes of non-HPITs are much larger than those of HPITs. More importantly, the hourly volumes of non-HPITs change a lot during the day and exhibit a strong seasonality pattern. That is, the highest trading volumes arrive at the beginning and the end of trading day. However, the hourly trading volumes of HPITs are around 2% per hour, which is relatively small, and quite stable during the day. The different intraday pattern of trading volumes for HPITs and non-HPITs suggests that uninformed liquidity traders are more likely to time their trades than are information-related traders.

[Insert Figure 5 here]

[Insert Table 13 here]

As Table 13 illustrates, a similar trend is found for MDAX stocks, but the dominance of HPITs over non-HPITs is less pronounced than that of DAX stocks. Surprisingly, for SDAX stocks, even though the information quality of HPITs is always higher than that of non-HPITs (2.997 vs. 0.693), the daily price contribution of HPITs is less than that of non-HPITs (39.95% vs. 60.05%). One possible explanation is that information-related trades are impeded by a high trading cost, a serious obstacle faced by intraday traders. Generally, the net profit of intraday informed trades is the difference between the gains derived from their belief- or fundamental-based information and the trading costs related to the order execution. In a market with a lower trading cost, information-related traders can get rewarded easily and have more incentive to trade against uninformed traders. In contrast, in a less liquid market that features a higher trading cost, informed traders have less incentive to trade against uninformed ones. To qualitatively investigate the relationship between trading cost and the contribution of HPITs across different markets, we present, in Figure 6, the intraday evolution of average relative bid-ask spread, which is defined as the ratio of bid-ask spread to midquote price. Two interesting insights arise from this figure. First, on average, the best bid-ask spread of SDAX stocks is much larger than those of DAX and MDAX stocks. More precisely, the spread of SDAX stocks is almost six times and three times as large as that of DAX and MDAX stocks, respectively. This means that information-related traders in SDAX stocks have to bear an extremely high cost before getting rewarded.

Second, the average spreads for stocks in different indexes decrease during the trading day, with an exception in the middle of the trading session. These findings seem to confirm that: 1) most of the information is diffused at the beginning of the trading session; and 2) at the opening, the market exhibits a higher degree of information asymmetry, and liquidity providers face a high risk of adverse selection. To protect themselves, liquidity providers increase the bid-ask spread.

[Insert Figure 6 here]

6 Modeling for HPIT Incidence

We further investigate which market conditions are associated with incidence of HPITs. To capture the well-documented persistence of trade type (buy-initiated trades are more likely to be followed by buy-initiated trades) and various market characteristics, we apply an auto-logistic model (Cox et al., (1981)) in which the dynamics of stationary variables is modeled by an autoregressive structure and exogenous variables. More specifically, we adopt a structure called the GLARMA (Generalized Linear Autoregressive Moving Average) binary model, which is a general structure for moving average-type behavior (Shephard (1995)). The use of market condition variables is straightforward in that informed trading is determined or driven by market conditions. Theoretical models (Admati and Pfleiderer (1988), Collin-Dufresne and Fos (2016)) predict that informed traders trade more aggressively when market liquidity, trade intensity, and market volatility are high. To capture these market conditions, we use average spread, price range, trading volume, and trading intensity as our trading variables. Average spread and price range are proxies for liquidity and volatility, respectively. Given intraday price discovery dynamics, we also include the first-hour dummy in our modeling.

$HPIT_t$ is a bivariate variable that takes the values of 0 or 1 to indicate whether the kth trade belongs to HPITs or not. Given that the log-likelihood function of the auto-logistic model is concave, numerical optimization can be done easily and reliably. However, high-frequency data often exhibit a slow decay for longer lags in an autoregressive structure. Thus, there is a trade-off between bias and variance, i.e., inference with too few parameters may be biased, while that with too many parameters may cause precision and identification problems. To solve this, the auto-logistic model for HPITs is defined as:

$$f\left(HPIT_t = 1\,|F_{t-1}\right) = p(\theta_t), \text{where } \; p(\theta_t) = \frac{\exp(\theta_t)}{1+\exp(\theta_t)} \tag{8}$$

$$\text{and } \theta_t = c + g_t + \Delta^{'} M_{t-1}, \quad g_t = \sum_{j=1}^{p} \gamma_j g_{t-j} + \sum_{j=1}^{l} \lambda_t HPIT_{t-j}\,.$$

$$\Delta^{'} M_{t-1} = \delta_1 \times Range_{t-1} + \delta_2 \times AvgSpread_{t-1} + \delta_3 \times TradeVolume_{t-1} + \delta_4 \times TradeDensity_{t-1} + \delta_5 \times FirstHour_{t-1}$$

$$\text{Consequently, } f(HPIT_t = 0\,|F_{t-1}) = \frac{1}{1+\exp(\theta_t)},$$

where M_{t-1} is the vector of market-related variables known at $t-1$. In this logistic modeling, the parameter θ_t is time-varying and depends on both its own lag variables, such as lags of g_t and $HPIT_t$, and some market-related variables. Table 14 reports the estimated results of the auto-logistic model of HPITs for DAX stocks. The coefficients of the "GLAR" part are positive and significant with a mean of 0.974, which suggests a cluster effect: HPITs are more likely to be followed by another HPIT. In addition, the estimation results confirm several theoretical predictions regarding informed trading incidence. HPITs are more likely to take place when trading conditions are characterized by high volatility ($\delta_1 > 0$ for 16 out of 30 DAX stocks), high liquidity ($\delta_2 < 0$ for 25 out of 30 DAX stocks), greater trading volume ($\delta_3 > 0$ for 29 out of 30 DAX stocks), and higher intensity ($\delta_4 > 0$ for all stocks). The model is validated by ROC (Receiver Operating Characteristic) and count accuracy.[37] We obtain around 60% for both tests for the proposed binary model. The summary results for three indexes are presented in Table 15. For MDAX stocks, the results are similar to those of DAX stocks and confirm the association of HPITs and informed trading behaviors. However, for SDAX stocks, the market condition

[37]ROC evaluates binary model accuracy at various threshold settings (Swets (1986), (1988)) by considering Type I and Type II errors. Count accuracy measures the in-sample accuracy predicted by the model.

variables are less associated with HPIT incidence, suggesting that fewer noise traders and less market transparency make informed trading less sensitive to market trading conditions.

[Insert Table 14 here]

[Insert Table 15 here]

7 Impact of HPITs on Intraday Volatility

7.1 Impact on Intraday Volatility

Up to now, we show how important HPITs are in daily price contribution and how to identify them. We next turn our attention to their implications for short-term volatilty. The noisy rational expectation model of Hellwig (1980) argues that volatility increases with uninformed trading. In his model, information is aggregated into price by the actions of risk-averse, heterogeneous agents who, individually, have no influence on prices. Rational informed investors stabilize prices by taking positions whenever prices deviate from their fundamentals, i.e. take long (short) position when the price is lower (higher) than fundamentals. As the proportion of informed investors increases, their impact on price increases, leading to a decrease in the deviation of price from its fundamental value. However, if the number of uninformed or liquidity traders increases, there will be an increase in volatility caused by uninformed trading. Wang (1993) also provides a model of asymmetric information and shows that the conditional volatility of prices increases with uninformed trading. In sum, both Hellwig's (1980) and Wang's (1993) models show that volatility decreases with informed trading and increases with liquidity/uninformed trading. Therefore, if HPITs are associated with informed trading, our results should be in line with the theoretical models. That is, we should find that an increase in HPITs leads to a decrease in volatility. In addition, because it is hardly possible for traders' risk aversion and market settings to change on a daily basis, it is possible to rule out such "macro factors" and reasonable to consider in terms of a close connection between information trading and volatility.

In order to examine the impact of HPITs on intraday volatility, we analyze the effect of the proportion of HPITs on the 15-min conditional volatility. Given that high-frequency data behaves very differently from low-frequency data, before estimating the model, we first remove seasonality by following a regression approach as did Dufour and Engle (2000). Moreover, the Ljung-Box statistics with 15 lags on the deseasonalized returns and the corresponding volatilities reject independence at all significance levels for most of the stocks in the sample. Thus, taking the model efficiency and parsimony into consideration, we estimate the model with an EGARCH(1,1) for all DAX, MDAX, and SDAX stocks:

$$r_i = \sigma_i \cdot \varepsilon_i \tag{9}$$

$$log(\sigma_i^2) = \omega + \sum_{j=1}^{p} \alpha_j g(Z_{i-j}) + \sum_{j=1}^{q} \beta_j log(\sigma_{i-j}^2) + \gamma HPIT\%_{i-1} \tag{10}$$

with $g(Z_i) = Z_i + \lambda(|Z_i| - E(|Z_i|))$, and where r_i is ith 15-min deseasonalized return, $HPIT\%_{i-1}$ relates to the proportion of HPITs for the period $i-1$, and ε_i is a normally distributed random variable. The parameters β and λ capture the autocorrelation in volatility. γ measures the impact of HPITs on volatility. After the estimation, the model is validated again by Ljung-Box statistics (with 15 lags) of the standardized residuals and squared standardized residuals.

Table 16 shows the estimation results of the proposed model for DAX stocks. The results suggest that 1) there is a high persistence in volatility given that the parameter β has a mean of 0.862. 2) 29 out of 30 DAX stocks have a negative γ, statistically significant at the 1% level. 3) The proposed model effectively captures the dynamics of volatility, which is validated by Ljung-Box statistics.[38] Similar results are obtained for the MDAX and SDAX stocks. For the sake of brevity, we only present a summary of the estimated parameters in Table 17, instead of full estimation results. In sum, HPITs have negative effect on volatility. However, this negative effect varies across different stock indexes. Specifically, this negative effect decreases, in the absolute term,

[38] The 5% critical value for the Ljung-Box with 15 lags is 24.99. A statistic less than 24.99 means that we cannot reject the null hypothesis that the time series is autocorrelated.

from 2.24 for DAX stocks to 1.11 for MDAX stocks, and 0.42 for SDAX stocks.

[Insert Table 16 here]

[Insert Table 17 here]

One plausible explanation for this difference in the impact on volatility is the difference in information condition of DAX, MDAX, and SDAX stocks. More specifically, DAX stocks are large-cap stocks and have greater public disclosure, which leads to favorable trading conditions such as more transparency, higher liquidity and more market turnover. The channel under which more disclosure leads to more market turnover and liquidity trading is highlighted in the theoretical model of Han et al. (2016). The intuition is that greater public disclosure lowers the expected loss of liquidity traders, thereby attracting more such traders. Implicitly, the proportion of traders that approximately know the (expected) fundamental value of the stocks is higher than that for the less transparent MDAX and SDAX stocks. Furthermore, a greater public disclosure market discourages private learning about fundamentals but encourages information acquisition about trading motivations of other traders (Banerjee et al. (2018)). Therefore, traders have a greater incentive to identify the trading motivation of other traders. For example, when the stock price deviates from its (expected) fundamental value caused by buying or selling pressure from uninformed or liquidity traders, belief-based traders will act as informed traders by taking the contrarian trades. Also, because of the favorable trade condition in the DAX stocks such as smaller bid-ask spread and a large number of uninformed traders, informed traders can easily get rewarded by providing liquidity to uninformed traders.

Generally, as mentioned above, belief-based traders do not rule out traditional or sophisticated high-frequency market makers who have access to both public information and information related to their customers. In the modern financial market, sophisticated market making algorithm can learn fast about trading motivations of other traders by preying on the footprints they leave in the market (Korajczyk and Murphy (2018) and Van Kervel and Menkveld (2019)). In addition, having all the trading history of their customers and participating actively in the interdealer markets, the market makers possess private information about other market participants and are in a good position to judge if someone is informed or not.[39]

However, for less liquid medium-cap stocks there is more information asymmetry and a larger bid-ask spread. Thus, the proportion of uninformed traders is lower compared with large-cap DAX stocks, and the short-term price deviations caused by liquidity trading are less likely. However, even if this price distortion still exists, it is difficult for informed traders and belief-based traders to get rewarded when faced with a large transaction cost (a large bid-ask spread). The same arguments also hold for the least liquid SDAX stocks. That is, SDAX stocks are characterized by the highest bid-ask spread, the lowest proportion of uninformed traders, the least likely short-term price deviations. Overall, in line with previous theoretical models, the negative relationship between HPITs and intraday volatility suggests that HPITs are rewarded for acting as price stabilizers. However, this negative effect decreases when information asymmetry and bid-ask spread increase, which are consistent with the idea that informed traders have to limit their involvement as price stabilizer in presence of unfavorable trade conditions.

7.2 Information Conveyed by HPITs: Autocorrelation Test

Thus far, we have empirically shown that a higher proportion of HPITs leads to a decline in volatility. We now provide more detailed empirical evidence on the channel through which informed trading could reduce volatility and how the magnitude of this decline in volatility varies with information environments. To do so, we define contrarian (herding) HPITs as trades that are against (after) the current price trend. Specifically, buy (positive) HPITs in the presence of decreasing price and sell (negative) HPITs during a price increase are designated as contrarian HPITs. Similarly, buy (positive) HPITs during a price increase and sell (negative) HPITs in the presence of a price decrease are defined as herding HPITs. By definition, contrarian (herding) HPITs involve price correction (price convergence) informed trading. If our conjecture that HPITs are sent by informed traders is correct, we should have negative coefficients for contrarian

[39] Trading ahead of one's own customers is illegal, but the mentioned strategy entails trading after or against one's customers.

HPITs, and insignificant ones for herding HPITs in equation (11). That is, contrarian HPITs lead to a price reversal and herding HPITs have no effect on return autocorrelation. In addition, $\delta_{1,i} + \delta_{2,i}$ capture the net effect of informed trading behavior in a given information environment. More specifically, when $\delta_{1,i} + \delta_{2,i} < 0$, this suggests that informed traders concentrate more on trading against uninformed traders than trading with their own private information to make price converge to the fundamentals. The opposite is true when $\delta_{1,i} + \delta_{2,i} > 0$.

We estimate the following regression by controlling for the trading volume[40] and other market variables:

$$\begin{aligned} R_{i,t} = & (\delta_{1,i} HPIT\%_{\,i,\text{t-1}} * \mathtt{1}_{(contrarian)} + \delta_{2,i} HPIT\%_{\,i,\text{t-1}} * \mathtt{1}_{(herding)} + \delta_{3,i} Volume_{t-1}) R_{i,t-1} \\ & + \psi_{i,1} Range_{i,t-1} + \psi_{i,2} Spread_{i,t-1} + \psi_{i,3} Price_{i,t-1} \\ & + \sum_{k=2}^{5} \beta_{i,k} R_{i,t-k} + \sum_{k=1}^{N} \alpha_{i,k} D_{kt} + e_{i,t} \end{aligned} \tag{11}$$

where $Range_{i,t-1}$ is the range between maximum and minimum price for stock i during the period $t-1$. $Spread_{i,t-1}$ and $Volume_{t-1}$ are the average spread and the total trade volume during the period $t-1$, respectively. $\delta_{1,i}$ and $\delta_{2,i}$ measure the impact of contrarian and herding HPITs on return's autocorrelation. Moreover, $\mathtt{1}_{(contrarian)}$ and $\mathtt{1}_{(herding)}$ are dummy variables that take the value of 1 if the price change of the HPITs goes against or follows the price trend at $t-1$. Finally, D_{kt} is the dummy variable that takes the value of 1 if period t belongs to hour k of the trading session. Table 18 presents the results for DAX stocks. The estimated δ_1 for contrarian HPITs are negative and significant for all stocks at the 1% level of confidence. Thus, contrarian HPITs of the DAX stocks do lead to price reversals. However, δ_2 has an average around zero and 27 out of 30 stocks have insignificant coefficients, which suggests an insignificant effect of herding HPITs on return for the DAX stocks. The total effect of HPITs, measured by $\delta_1 + \delta_2$, is also negative and significant for 27 out of 30 stocks, which is consistent with the result that HPITs lead to a decline in volatility. Therefore, based on these empirical results, we conclude that in a high-level public information environment, most of the time informed traders trade against uninformed traders.

[40] In particular, Volume is included in regression 11 to control for result of Campbell et al. (1993) that volume contributes to return reversals.

[Insert Table 18 here]

For the sake of brevity, we present a summary of the estimated parameters for stocks in the three indexes in Table 19, instead of full estimation results. For MDAX stocks, both coefficients of contrarian and herding HPITs have an impact on short-term returns. Specifically, all stocks have negative and significant δ_1, and 29 out of 50 stocks have positive and significant δ_2. This result suggests that herding HPITs lead to a price continuation for more than half of MDAX stocks.[41] In terms of total effect, 45 out of 50 stocks have negative and significant $\delta_1 + \delta_2$, which confirms the dominance of price correction informed trading. From the perspective of information acquisition, the results also provide support for complementarity in learning: in a less public information disclosure environment when prices do not fully reflect available public information, traders have the incentive to learn both dimensions of information and make profits by incorporating fundamental-based information into the current price and correcting price distortion created by uninformed traders. Finally, for SDAX stocks, 31 out of 48 stocks have negative and significant δ_1 for contrarian HPITs, which presents the lowest ratio of negative and significant δ_1 among three indexes. Also, 40 out of 48 stocks have insignificant δ_2 for herding HPITs, which suggests that herding HPITs have little effect on return autocorrelation.

[Insert Table 19 here]

Information-related trades can occur in two situations, one is when the market price deviates from its fundamental price, and actual spread allows informed traders to get rewarded by correcting the price distortion. The second one is when the informed traders have private information and try to take the maximum profit before this private information becomes public. For large-cap stocks such as DAX stocks, the first situation is more often observed. For SDAX stocks, the second situation is more likely. The MDAX stocks are in-between. Our results show that 1) the average effect of HPITs for DAX stocks is -0.903, suggesting that a one-percent increase in HPITs would reduce the autocorrelation by approximately 0.01 on average; 2) The effect is less

[41] The theoretical model of Kyle (1985) does not imply the autocorrelation of return in the presence of informed traders. However, empirically, due to market friction or information quality, we could observe a positive autocorrelation.

pronounced for MDAX stocks with decreases of 0.0036, and the smallest (0.0013) for the SDAX stocks. In sum, our results of autocorrelation test provide more evidence that contrarian HPITs act as a price stabilizer and lead to return reversal for DAX, MDAX and SDAX stocks. However, the size of this reversal decreases when information asymmetry and bid-ask spread increase.

8 Impact of HPIT on Price Efficiency

So far, we have empirically shown that HPITs lead to a decline in intraday volatility by making more contrarian trades, and explained why this decline in volatility is not the same across different groups identified by the difference in their information setting. We next turn our attention to the effect of HPITs on price efficiency. One of the fundamental roles of the financial market is to facilitate the price discovery process, which also means making stock prices to reflect fundamentals quickly. Thus the questions of how prices reflect the fundamental values and how price efficiency is affected by various market settings have drawn strong interest among academics, practitioners, and regulators. However, theoretical and empirical finance do not always have the same measurements and conclusions of market efficiency, depending on their focus. Specifically, theoretical models emphasize on the static precision of the conditional expected price based on fundamental information (Diamond (1985), Gao and Liang (2013), Colombo et al. (2014), Banerjee et al. (2018), and Dugast and Foucault (2018)), while empirical studies attempt to assess the dynamics aspect of efficiency, that is, statistically, how closely stock prices follow a random walk (Lo and MacKinlay (1988), Boehmer and Kelly (2009), Chaboud et al. (2014), Conrad et al. (2015), and Rosch et al. (2016)). Given this nuanced divergence in measurement, our study follows the empirical finance literature and uses variance ratio- and autocorrelation-based measurements for price efficiency.

8.1 Variance Ratio Evidence

The first measurement we use for price efficiency is derived from the variance ratio proposed by Lo and MacKinlay (1989). According to their notation, x_t represents a log price process,[42] and there are n non-overlapping long-horizon intervals in the measurement interval and q non-overlapping short-horizon intervals in each long-horizon interval. Moreover, each interval is equally spaced so that there exist $T = nq$ returns in the measurement interval. In such a setting, the estimate of the mean drift in prices is equal to:

$$\hat{\mu} = \frac{1}{nq} \sum_{k=1}^{nq} (x_k - x_{k-1}) = \frac{1}{nq}(x_{nq} - x_0), \tag{12}$$

and the estimates of the variance are as follow

$$\overline{\sigma}_a^2(q) = \frac{1}{nq-1} \sum_{k=1}^{nq} (x_k - x_{k-1} - \hat{\mu})^2, \tag{13}$$

$$\overline{\sigma}_c^2(q) = \frac{1}{m} \sum_{k=q}^{nq} (x_k - x_{k-q} - q\hat{\mu})^2, \tag{14}$$

where $m = q(nq-q+1) \times (1-\frac{q}{nq})$, and $\overline{\sigma}_a^2$ and $\overline{\sigma}_c^2(q)$ are short and large interval return variances, respectively.

If prices follow a random walk process, variances should be linear in the measurement interval. This implies that the ratio of scaled large interval return variance over short interval return variance, $\overline{\sigma}_c^2(q)/\overline{\sigma}_a^2$, should be equal to one. Specifically, the test based on the random walk hypothesis is

[42] We use midquote price instead of trade price to avoid the negative autocorrelation caused by the bid-ask bounce.

$$M_r(q) \equiv \frac{\overline{\sigma}_c^2(q)}{\overline{\sigma}_a^2} - 1 = 0. \tag{15}$$

For our dataset, we take 30 seconds and 5 minutes as our short and large intervals, respectively. Implicitly, q is equal to 10. Further, we compute the ratio of variance over 2-hour and 4-hour measurement intervals. Precisely, taking the example of a 2-hour measurement interval, there are 240 non-overlapping short and 24 non-overlapping large intervals. To avoid the degeneration of the variance ratio, we require at least 30 nonzero short interval returns in each 2-hour measurement interval. We choose 30 seconds as our short intervals because the interval should be short enough to capture the high-frequency dynamics in price changes and provide sufficient observations to compute the variance. This interval also needs to be long enough to avoid high-frequency noise.

To examine the effect of HPITs on price efficiency, we take the absolute value of $M_r(q) - 1$ as our efficiency measure and run the following fix-effect and random-effect panel regressions:

$$\begin{aligned} \mid M_r(q) - 1 \mid_{i,t} = {} & \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) \\ & + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t} \,, \end{aligned} \tag{16}$$

$$\begin{aligned} \mid M_r(q) - 1 \mid_{i,t} = {} & \mu + \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) \\ & + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t} \,. \end{aligned} \tag{17}$$

As mentioned above, we include *Range* to control for volatility and *Spread* for liquidity. If HPITs are information-related trades, according to the random walk hypothesis, the future price should be less predictable because more information is incorporated in the price. In other words, the presence of HPITs helps to incorporate information into the price and will make the future prices less predictable or more likely to follow a random walk process. Our dependent variable is the absolute value of $M_r(q) - 1$ and the minimum value of zero corresponds to a pure random walk process. Therefore, if our conjecture is correct, we expect a negative effect of HPITs on the dependent variable.

Table 20 reports the results of regressions (16) and (17). For DAX stocks, an increase in HPITs significantly results in price efficiency at the 1% level. However, this effect decreases and becomes insignificant for MDAX stocks. Although the presence of more HPITs still makes the future price more efficient, it is not statistically significant at the 5% level. Finally, regarding the least liquid SDAX stocks, more HPITs adversely affect price efficiency and make it more predictable at the 5% level. Furthermore, as mentioned above, informed traders consider both expected profits and trading costs before submitting their orders. It follows from Table 20 that an increase in the spread also makes the future price more efficient for both DAX and MDAX stocks at the 1% level. However, for the least liquid SDAX stocks, which already have extremely wide bid-ask spread, the change in spread has little impact on efficiency.

[Insert Table 20 here]

We draw two main inferences from the results. First, both HPITs and bid-ask spread play a role in price efficiency. More specifically, an increase in HPITs makes the price more efficient by adding more information into the price and make it less predictable. However, another channel through which price efficiency can be improved is to increase the bid-ask spread. To understand the relationship between spread and price efficiency, consider that the expected fundamental value of the stock is p_0, which is different from the current midquote price mq_0, and there exists a spread s_0 between the best ask and the best bid price. When s_0 is so large that the expected fundamental price, p_0, falls in the interval $(mq_0 - \frac{s_0}{2}, mq_0 + \frac{s_0}{2})$, this discourages information-related trades because the gain from the information cannot cover the transaction cost. As a result, prices remain efficient without trading activity.

The second inference we draw from these results is that the way information is incorporated in the price varies with liquidity levels and information environments. Combined with the results found in subsection 7.2 for DAX stocks, HPITs mainly reflect price correction informed trading. Given greater public disclosure and a higher proportion of liquidity traders, competitiveness between ex-ante and ex-post informed traders is high. When prices deviate from the expected fundamental value, these informed-traders have to react quickly to get rewarded. Because this process makes more information be incorporated in the stock price, price efficiency is improved rapidly. However, for illiquid stocks, HPITs might be followed by uninformed trades, and a high bid-ask spread impedes the price correction from contrarian HPITs. In consequence, price efficiency deteriorates, and prices are less likely to follow a random walk. Another explanation for this adverse effect is that the expected fundamental prices may fall in the interval $(mq_0 - \frac{s_0}{2}, mq_0 + \frac{s_0}{2})$, but the efficiency is only computed with the mid-quote price, which might not be a reliable proxy for the expected fundamental value when the bid-ask spread is wide.

8.2 Autocorrelation Evidence

The variance ratio measures only one facet of price informational efficiency. More generally, one concern about high-frequency information-related traders is that they cut their large volumes into small ones and span them during a longer horizon, which may cause autocorrelation. We thus access the impact of HPITs on a more general measure of price efficiency: the autocorrelation of high-frequency return. Specifically, we investigate the effect of HPITs on the absolute value of the first-order autocorrelation of five-second returns every two hours. If HPITs are information-related trades, the returns should be less autocorrelated because more information is incorporated in the price, which suggests a negative effect of HPITs on absolute autocorrelation coefficient. The following panel regression is estimated for all DAX, MDAX, and SDAX stocks:

$$\begin{aligned} \mid \rho \mid_{i,t} = \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) \\ + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{18}$$

$$\begin{aligned} \mid \rho \mid_{i,t} = \mu + \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) \\ + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{19}$$

Table 21 presents the results of a panel regression of the absolute autocorrelation coefficient on HPITs and other market variables. Similar to the results for the variance ratio, for DAX stocks, a higher proportion of HPITs decreases the intraday return autocorrelation. This effect is significant for both fixed effect and random effect panel regressions. A less significant effect is found for MDAX stocks and an insignificant effect for SDAX stocks. Specifically, HPITs of SDAX stocks have a positive impact on autocorrelation, which confirms the slight price deviation after fundamental-based HPITs.

The results on autocorrelation-based efficiency provide more evidence on how HPITs act as a price stabilizer for the DAX, MDAX, and SDAX. Specifically, this effect depends on the characteristics of trading and information environments. For large-cap, liquid stocks such as DAX stocks, price correction informed trades reduce the return autocorrelation, whereas for medium- and small-cap stocks, given that a wide bid-ask spread impedes price correction, the role played by HPITs as price stabilizer is less pronounced. As a result, lesser price efficiency is observed in MDAX and SDAX.

[Insert Table 21 here]

9 Conclusion

We identify trades that have disproportionately large cumulative price changes relative to their proportion of volume (HPITs), and relate them to informed trades. Our results suggest that HPITs might be price correction based or price incorporation based, depending on stocks' information environments. We also show that the market implications of HPITs vary significantly with information environments. Typically, we consider stocks in the DAX, MDAX and SDAX indexes, traded at the Frankfurt Stock Exchange, as a natural setting in terms of information environments to study the type of information conveyed by HPITs.

To precisely identify HPITs, we first confirm the existence of trade size clustering and LOB matchedness. Typically, trade sizes feature clustering on 10, 50, 100 or their multiples, which accounts for more than 40% of total trades, and LOB-matched trades represent more than 50% of total trades. In addition, our empirical results of matched trades confirm our first hypothesis that the information quality of matched trades should be higher for stocks with greater public disclosure. The two-dimensional results of size-roundedness and size-matchedness provide evidence that both are important dimensions in informed trading identification. In addition, we show that HPITs not only have a higher cross-sectional price impact, but also a higher price impact over time

With the above identified informed trading, we validate the second hypothesis that a stronger presence of HPITs leads to a decline in volatility for all three groups of stocks. However, this negative effect increases with the level of stocks' public disclosure (i.e., it is highest for DAX stocks and lowest for SDAX stocks). To further explore the trading behavior conveyed by HPITs, we decompose HPITs into contrarian and herding HPITs, and show that contrarian HPITs are responsible for this decline in volatility and lead to price reversals. Again, this negative effect increases with the level of stocks' public disclosure. Our empirical findings support our third hypothesis that 1) contrarian HPITs are more present in the greater public disclosure market and trade against uninformed traders; and 2) herding HPITs are mainly related to fundamental-based information and have an insignificant effect on return autocorrelation.

Finally, we use variance-ratio and autocorrelation-based price efficiency measures to test our fourth hypothesis that HPITs increase price efficiency for stocks with greater public disclosure and high levels of liquidity. Our results show that for DAX stocks, the increase in HPITs improves price efficiency significantly. Similar results are found for MDAX stocks, but are not significant. For the least liquid SDAX stocks, the presence of HPITs slightly reduces market efficiency. We provide two possible explanations for this phenomenon: First, for stocks with more private-information traders and a wide bid-ask spread, informed trades might be followed by uninformed trades that lead a price to deviate from its fundamental level. Further, a high bid-ask spread impedes the price correction made by informed traders. Therefore, price efficiency deteriorates. Second, the midquote price we use to compute the efficiency measure might not be a reliable proxy for the expected true price, especially for the least liquid stocks with a large bid-ask spread. More research is needed to find a better proxy.

Figure 1: Matched Orders Vs. Unmatched Orders

Panel A

Initial Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10	400
300	9.95	10.03	300
50	9.90	10.07	2000
6000	9.88	10.42	1000
300	9.84	10.45	500

Matched Buy Market Order Volume = 400 (Level 1)

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10.03	300
300	9.95	10.07	2000
50	9.90	10.42	1000
6000	9.88	10.45	500
300	9.84		

Matched Buy Market Order Volume = 700 (Level 1 and 2)

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10.07	2000
300	9.95	10.42	1000
50	9.90	10.45	500
6000	9.88		
300	9.84		

Panel B

Initial Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10	400
300	9.95	10.03	300
50	9.90	10.07	2000
6000	9.88	10.42	1000
300	9.84	10.45	500

Unatched Buy Market Order Volume = 100 (Level 1)

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10	300
300	9.95	10.03	300
50	9.90	10.07	2000
6000	9.88	10.42	1000
300	9.84	10.45	500

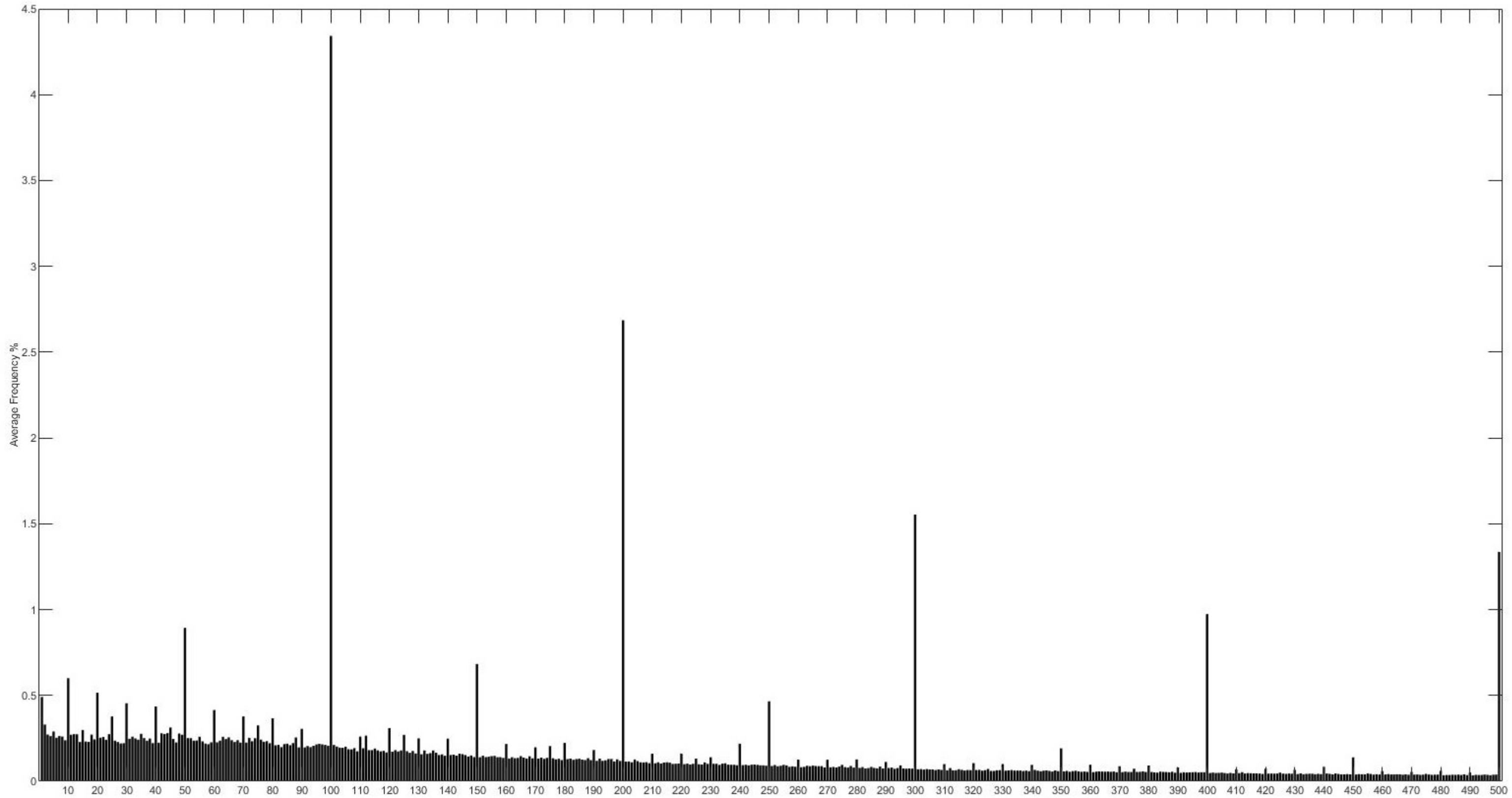
Unmatched Buy Market Order Volume = 500 (Level 1 and 2)

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10.03	200
300	9.95	10.07	2000
50	9.90	10.42	1000
6000	9.88	10.45	500
300	9.84		

The figure presents and compares the different outcomes of matched and unmatched orders. Panel A reports the impact of the level-1 matched buy market order and that of levels 1 and 2. Similarly, Panel B illustrates the impact of the level-1 unmatched buy market order and that of levels 1 and 2. Note that the level-1 matched buy market order with 400 shares and the level-1 unmatched buy market order with 100 shares have the same transaction price and price impact.

Figure 2: Histogram of Trade Sizes for DAX Stocks



The figure presents the cross-sectional average probability distribution of trade size of DAX30 stocks. The odd lot in the German Stock Exchange is one, and the size is truncated at 500 shares.

Figure 3: Marketable Orders Vs. Market Orders

Initial Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10	400
300	9.95	10.03	300
50	9.90	10.07	2000
6000	9.88	10.42	1000
300	9.84	10.45	500

Market Order

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
100	9.98	10.07	2000
300	9.95	10.42	1000
50	9.90	10.45	500
6000	9.88		
300	9.84		

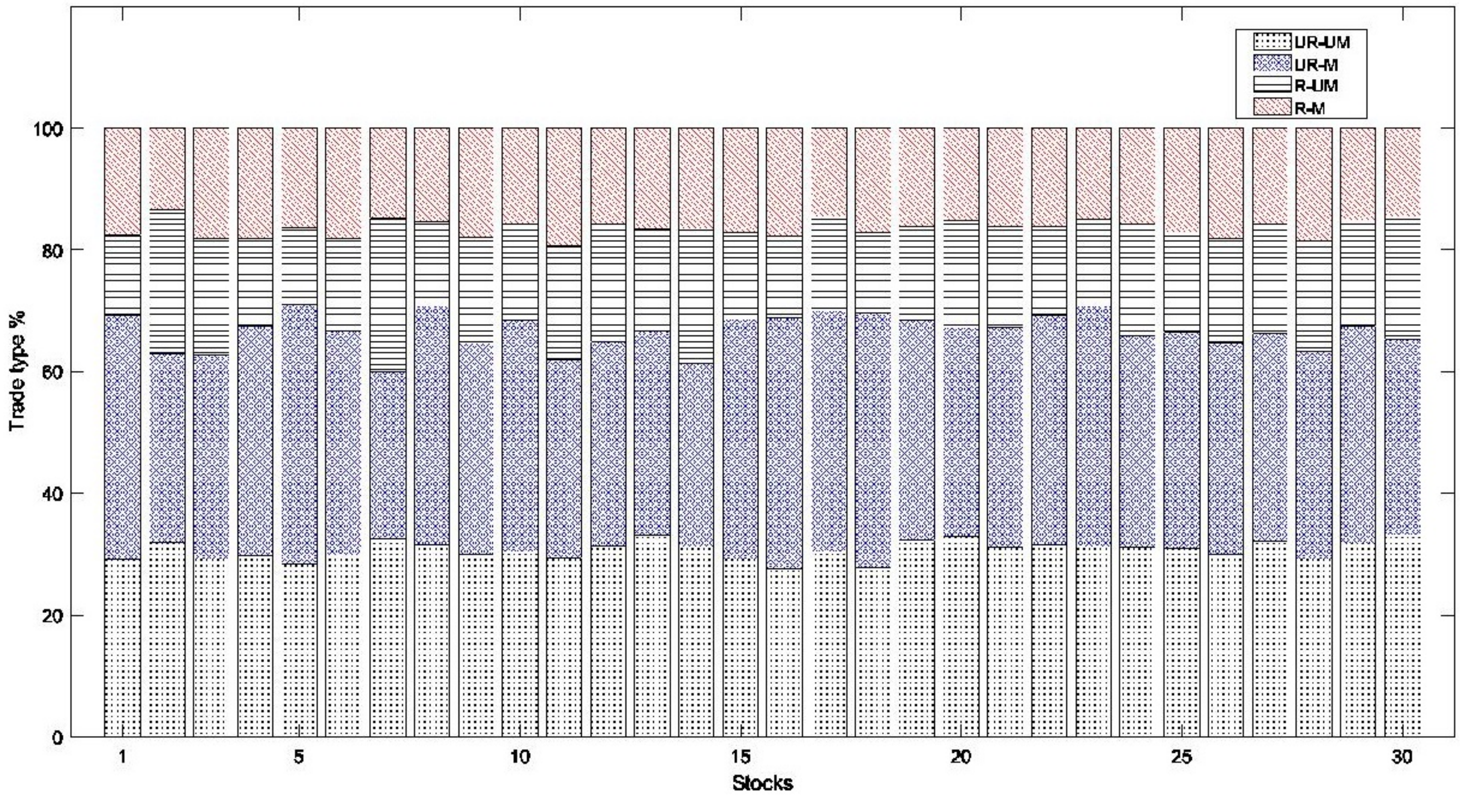
Marketable Order

New Order Book

Bid Volume	Bid Price	Ask Price	Ask Volume
300 (YOU)	10.05	10.07	2000
100	9.98	10.42	1000
300	9.95	10.45	500
50	9.90		
6000	9.88		

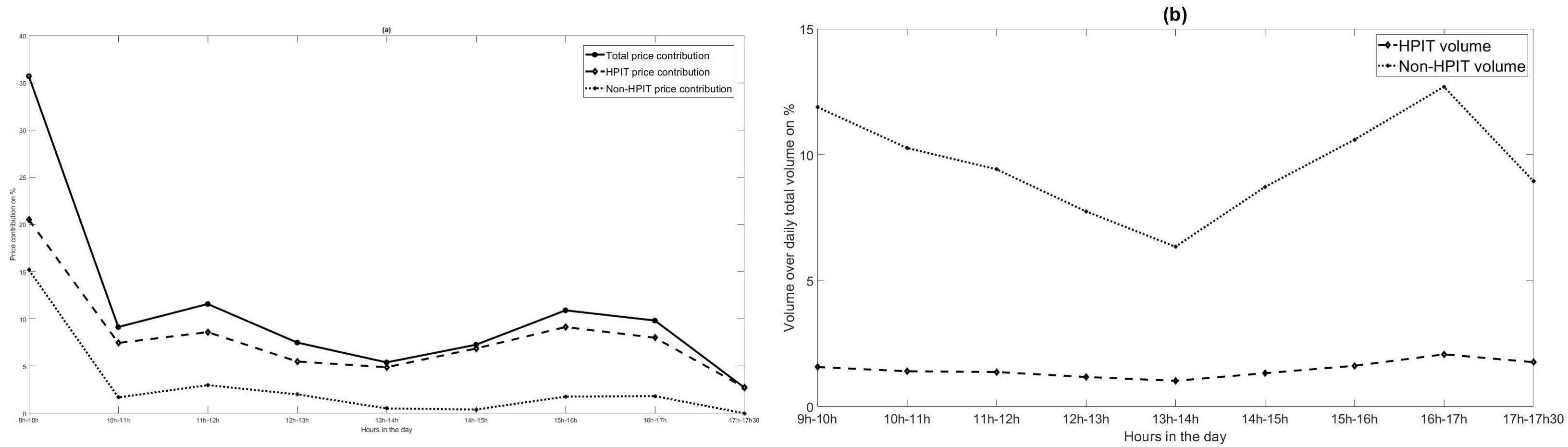
The figure presents and compares the different outcomes of market and marketable orders. The market order is submitted with volume of 700 shares. The marketable order is submitted with volume of 1 000 shares and price of 10.05 Euros.

Figure 4: Composition of Roundedness and Matchedness



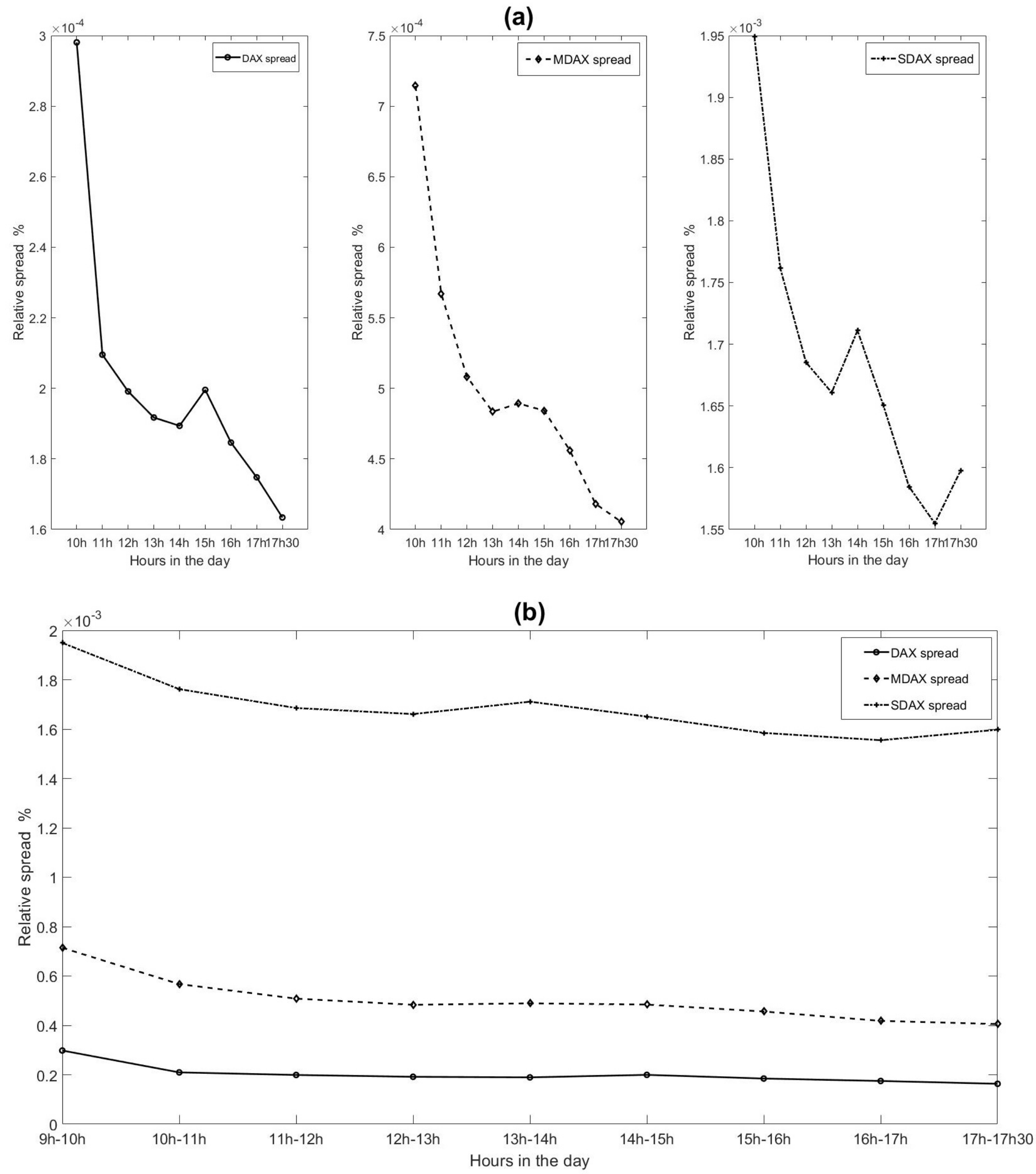
The figure illustrates the proportions of different trade types categorized by roundedness and matchedness for 30 DAX stocks. UR-UM, UR-M, R-UM, R-M indicate, respectively, unrounded-unmatched, unrounded-matched, rounded-unmatched, rounded-matched trades.

Figure 5: Daily Dynamics of HPITs and Volumes for DAX Stocks



Panel (a) illustrates the intraday evolution of price changes related to HPITs (dashed line), Non-HPITs (dotted line) and the sum of two (solid line). Panel (b) compares the intraday evolution of corresponding trading volumes (dashed line for HPITs and dotted line for Non-HPITs) for DAX30 stocks, from February 1, 2013 to July 31, 2013.

Figure 6: Average Intraday Dynamics of Spread for DAX, MDAX and SDAX Stocks



Panel (a) illustrates separately the intraday evolution of average relative bid-ask spread for DAX (solid lines), MDAX (dashed lines), and SDAX (dotted lines) stocks. Panel (b) compares the intraday evolution of average relative bid-ask spread for DAX (solid lines), MDAX (dashed lines), and SDAX (dotted lines) stocks. The sample period covers 6 months from February 1, 2013 to July 31, 2013. Relative bid-ask spread is defined as the ratio of bid-ask spread to midquote price.

Table 1: Summary Statistics for DAX, MDAX, and SDAX Stocks

	DAX	MDAX	SDAX
A. Sample			
Number of days	125	125	125
Number of stocks	30	50	50
B. Daily market			
Avg. Market Capitalization (in billion Euros)			
Mean	3362.67	464.00	60.29
Median	2318.76	279.04	45.32
Standard deviation	2647.32	604.67	50.62
Avg.Daily Price (in Euros)			
Mean	62.81	48.22	27.60
Median	57.48	34.09	16.44
Standard deviation	43.07	45.54	41.74
Avg.Daily Trading Volume (in million shares)			
Mean	4.08	0.35	0.21
Median	2.09	0.19	0.03
Standard deviation	6.71	0.52	0.58
Avg. Daily Turnover (in percentage)			
Mean	0.50%	0.31%	0.25%
Median	0.41%	0.26%	0.16%
Standard deviation	0.29%	0.17%	0.38%
Avg. Daily Return (in percentage)			
Mean	0.04%	0.03%	-0.07%
Median	0.03%	0.02%	-0.03%
Standard deviation	0.24%	0.13%	0.48%

This table reports the statistics for the average market capitalization (in billion Euros), the average daily price (in Euros), the average daily trading volume (in million shares), the average daily turnover (in percentage) defined as the trading volume over the outstanding shares and the average daily (log) return for the stocks in DAX, MDAX and SDAX indexes, from February 1, 2013 to July 31, 2013. All data are from the Compustat Global Security Daily files and based on the primary issues.

Table 2: Trade and Information Environment Statistics for DAX, MDAX, and SDAX Stocks

	DAX	MDAX	SDAX
A. Sample			
Number of days	125	125	125
Number of stocks	30	50	50
B. Trade environment			
Best bid-ask spread (basis points)			
Mean	5.45E-04	1.57E-03	5.20E-03
Median	6.02E-04	1.43E-03	4.68E-03
Standard deviation	1.49E-04	6.11E-04	2.45E-03
LOB depth ask (cum.5-level)			
Mean	11508	1388	2667
Median	3405	925	1544
Standard deviation	27368	1204	4770
LOB depth bid (cum.5-level)			
Mean	11873	1338	2854
Median	3390	913	1207
Standard deviation	29515	1107	5818
Shares/trade			
Mean	668	209	481
Median	259	144	251
Standard deviation	1138	182	744
Volumes (€)/trade			
Mean	17283.21	5982.99	4074.04
Median	15142.64	5467.92	3972.78
Standard deviation	6219.74	2054.80	1042.54
Duration (second)/per trade			
Mean	9.40	46.37	268.03
Median	8.87	35.49	294.24
Standard deviation	4.57	28.88	108.74
Daily number of trades			
Mean	4527	1025	185
Median	3954	1012	125
Standard deviation	2149	479	143
C. Information environment			
Monthly Number of news per stock in average			
Mean	606	104	21
Median	236	41	16
Standard deviation	752	288	14
Number of analysts			
Mean	29.77	2.88	2.02
Median	30	3	2
Standard deviation	4.55	0.45	0.51
Forecast dispersion			
Mean	0.008	0.009	0.032
Median	0.005	0.005	0.006
Standard deviation	0.015	0.014	0.146
Forecast error			
Mean	0.007	0.013	0.055
Median	0.002	0.004	0.005
Standard deviation	0.015	0.023	0.226

This table reports the statistics for trading and information environment variables. The best bid-ask spread is the relative bid-ask spread on percentage. LOB depth ask (bid) is the cumulative quantity available for the first three levels at the ask (bid) side of the LOB. Duration/trade is the time between two consecutive trades. The monthly number of news is the number of times that the company is mentioned in the mass media and the news data are from the RavenPack dataset. Finally, trades hit by hidden orders is the proportion of market orders that are matched with iceberg or hidden orders embedded in the open LOB.

Table 3: Estimated Results for Clustering for DAX, MDAX, and SDAX Stocks

	Panel A: DAX stocks						
Variables	Cons.	D5	D10	D50	D100	LnSize	Adjusted R^2
Coeff.	1.612	0.008	**0.127**	**0.207**	**1.316**	**-1.701**	0.8701
Stand.dev	0.189	0.010	0.016	0.036	0.058	0.019	
	Panel B: MDAX stocks						
Variables	Cons.	D5	D10	D50	D100	LnSize	Adjusted R^2
Coeff.	3.120	**0.038**	**0.080**	**0.251**	**1.430**	**-2.008**	0.8903
Stand.dev	0.382	0.022	0.031	0.072	0.113	0.046	
	Panel C: SDAX stocks						
Variables	Cons.	D5	D10	D50	D100	LnSize	Adjusted R^2
Coeff.	-0.708	**0.050**	**0.090**	**0.626**	**1.308**	**-1.339**	0.8405
Stand.dev	0.181	0.016	0.021	0.046	0.071	0.021	

Panel A, B and C report the results of regression, $LnFreq_i = \alpha + \beta_5 D5_i + \beta_{10} D10_i + \beta_{50} D50_i + \beta_{100} D100_i + \beta_{LnSize} LnSize_i + \epsilon_i$, for the stocks in DAX, MDAX and SDAX indexes, respectively. $LnFreq_i$ is the natural log of the percentage of trades of size i, $D5_i$, $D10_i$, $D50_i$, $D100_i$ are dummy variables if trade size i is a multiple of 5, 10, 50, 100, respectively, and $LnSize_i$ is the natural log of trade size i measured in numbers of shares. Bold entries indicate that the coefficients are significant at the 5% level.

Table 4: Dynamics of Trade-size Clustering for DAX, MDAX, and SDAX Stocks

Year-Month		201302-201307, DAX	201302-201307, MDAX	201302-201307, SDAX
A. Sample				
Number of days		125	125	125
Number of stocks		30	50	50
B. Statistics				
Daily cross-sectional percent %	mean	33.39 (1.02)	33.20(1.21)	41.04 (2.64)
	std	3.97 (0.51)	5.88 (1.03)	10.03 (1.44)
Individual stock time-series percent %	mean	33.40 (2.92)	33.20 (3.49)	41.04 (6.63)
	std	2.84 (0.61)	4.81 (1.26)	9.75 (2.52)

This table reports measures of the cross-sectional distribution of size clustering around multiples of 10, 50 and 100 shares. The cross-sectional mean and standard deviation of this distribution is calculated daily. Each cell contains the mean and the standard deviation of the daily means and the standard deviations over the same period. The time-series mean and standard deviation are calculated for each stock. Each cell contains the mean and standard deviation of the time-series means and standard deviations across stocks.

Table 5: Dynamics of LOB-matched Trade-size for DAX, MDAX, and SDAX Stocks

Year-Month		201302-201307, DAX	201302-201307, MDAX	201302-201307, SDAX
A. Sample				
Number of days		125	125	125
Number of stocks		30	50	50
B. Statistics				
Daily cross-sectional percent %	mean	52.34 (2.08)	52.23 (2.79)	42.00 (5.07)
	std	5.76 (0.97)	7.68 (1.21)	9.94 (2.47)
Individual stock time-series percent %	mean	52.34 (4.06)	52.23 (4.05)	42.00 (2.38)
	std	4.59 (0.78)	7.05 (1.22)	10.07 (1.32)

This table reports measures of the cross-sectional distribution of LOB-matched sizes. The cross-sectional mean and standard deviation of this distribution is calculated daily. Each cell contains the mean and the standard deviation of the daily means and the standard deviations over the same period. The time-series mean and standard deviation are calculated for each stock. Each cell contains the mean and standard deviation of the time-series means and standard deviations across stocks.

Table 6: Matched and Non-matched Trades Distributions

			BID					ASK			
	Level of open LOB		>3	3	2	1	0	1	2	3	>3	Total
DAX	Nb.Trade	Matched	0.001%	0.001%	0.009%	26.450%	N/A	26.369%	0.009%	0.001%	0.001%	52.84%
		Non_Matched	0.142%	0.393%	1.839%	20.951%	0.151%	21.372%	1.810%	0.379%	0.130%	47.16%
	Volume	Matched	0.010%	0.003%	0.018%	25.757%	N/A	25.629%	0.019%	0.003%	0.011%	51.45%
		Non_Matched	1.062%	1.551%	3.791%	17.988%	0.099%	17.934%	3.786%	1.516%	0.995%	48.55%
MDAX	Nb.Trade	Matched	0.003%	0.002%	0.012%	25.927%	N/A	25.817%	0.013%	0.003%	0.004%	51.78%
		Non_Matched	0.305%	0.458%	2.049%	21.177%	0.151%	21.409%	2.021%	0.438%	0.281%	48.22%
	Volume	Matched	0.024%	0.007%	0.024%	25.337%	N/A	25.607%	0.026%	0.010%	0.036%	51.07%
		Non_Matched	2.492%	1.367%	3.299%	17.386%	0.115%	17.678%	3.315%	1.396%	2.547%	48.93%
SDAX	Nb.Trade	Matched	0.015%	0.008%	0.051%	21.290%	N/A	20.637%	0.028%	0.011%	0.026%	42.07%
		Non_Matched	0.879%	1.056%	4.135%	22.830%	0.151%	23.896%	3.789%	0.857%	0.712%	57.93%
	Volume	Matched	0.076%	0.032%	0.073%	19.884%	N/A	19.433%	0.063%	0.035%	0.091%	41.69%
		Non_Matched	8.420%	2.951%	6.588%	16.200%	0.079%	17.122%	6.093%	2.421%	6.145%	58.31%

This table reports the distribution of matched and non-matched trades and their corresponding volume proportions at different levels of the open LOB for DAX, MDAX, and SDAX stocks. For non-matched trades, the level is the highest level that is partially depleted. For instance, for DAX stocks, 1.839% of trades involve sell-initiated trades that consume the total liquidity at the best bid and part of the liquidity at the second bid. Trade at level zero relates to a transaction inside the best bid and ask. There are two possible scenarios: 1) a market buy (sell) order trades against a hidden sell (buy) limit order with a price lower (higher) than the best ask (bid), and 2) both buy and sell orders arrive at the market simultaneously and are matched automatically.

Table 7: Price Contribution of Trade Sizes for DAX, MDAX, and SDAX Stocks

	DAX			MDAX			SDAX		
Size	WPC	Share of Trades	Share of Volume	WPC	Share of Trades	Share of Volume	WPC	Share of Trades	Share of Volume
Small	12.53%	30.30%	5.41%	35.12%	31.11%	5.63%	20.80%	31.02%	5.24%
Medium	45.60%	39.95%	26.05%	39.71%	40.01%	27.55%	35.33%	39.89%	26.23%
Large	41.87%	29.75%	68.54%	25.17%	28.88%	66.82%	43.88%	29.09%	68.52%

This table reports the weighted price contribution for trade type classified by size for DAX, MDAX, and SDAX stocks. WPC is the weighted price contribution. The share of trades (share of volume) relates to the percentage of trades (volume) in each size category.

Table 8: Price Contribution of Roundedness-Size And Matchedness-Size for DAX, MDAX, and SDAX Stocks

	DAX							
	Panel A1: Price Contribution and Trade Roundedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UR	RD	UR	RD	UR	RD	UR	RD
Small	22.26%	-9.73%	20.25%	10.05%	3.55%	1.86%	6.26	-5.24
Medium	37.90%	7.70%	24.62%	15.33%	16.10%	9.95%	2.35	0.77
Large	35.68%	6.19%	20.03%	9.71%	46.96%	21.58%	0.76	0.29
Total	95.84%	4.16%	64.91%	35.09%	66.61%	33.39%	1.44	0.12
	Panel A2: Price Contribution and Trade Matchedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UM	MD	UM	MD	UM	MD	UM	MD
Small	-4.61%	17.14%	16.34%	13.96%	2.73%	2.68%	-1.69	6.39
Medium	19.25%	26.35%	17.09%	22.87%	10.82%	15.23%	1.78	1.73
Large	24.29%	17.57%	13.49%	16.26%	34.09%	34.45%	0.71	0.51
Total	38.94%	61.06%	46.92%	53.08%	47.64%	52.36%	0.82	1.17

	MDAX							
	Panel B1: Price Contribution and Trade Roundedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UR	RD	UR	RD	UR	RD	UR	RD
Small	24.67%	10.45%	23.11%	8.00%	4.06%	1.57%	6.07	6.66
Medium	25.66%	14.05%	27.09%	12.93%	18.48%	9.07%	1.39	1.55
Large	14.09%	11.08%	19.11%	9.76%	44.25%	22.57%	0.32	0.49
Total	64.43%	35.57%	69.31%	30.69%	66.79%	33.21%	0.96	1.07
	Panel B2: Price Contribution and Trade Matchedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UM	MD	UM	MD	UM	MD	UM	MD
Small	20.14%	14.99%	15.77%	15.34%	2.82%	2.81%	7.13	5.34
Medium	24.68%	15.03%	18.77%	21.24%	12.78%	14.77%	1.93	1.02
Large	20.58%	4.59%	13.05%	15.83%	32.17%	34.64%	0.64	0.13
Total	65.40%	34.60%	47.59%	52.41%	47.77%	52.23%	1.37	0.66

	SDAX							
	Panel C1: Price Contribution and Trade Roundedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UR	RD	UR	RD	UR	RD	UR	RD
Small	13.97%	6.82%	22.33%	8.68%	3.57%	1.67%	3.92	4.07
Medium	23.36%	11.97%	25.28%	14.61%	16.60%	9.64%	1.41	1.24
Large	20.57%	23.31%	17.50%	11.60%	40.02%	28.50%	0.51	0.82
Total	57.91%	42.09%	65.11%	34.89%	60.19%	39.81%	0.96	1.06
	Panel C2: Price Contribution and Trade Matchedness							
	WPC		Share of Trades		Share of Volume		Info.Quat	
Trade type	UM	MD	UM	MD	UM	MD	UM	MD
Small	14.67%	6.12%	17.67%	13.34%	2.97%	2.28%	4.95	2.69
Medium	29.52%	5.81%	22.27%	17.62%	14.60%	11.63%	2.02	0.50
Large	27.21%	16.67%	16.34%	12.76%	40.43%	28.10%	0.67	0.59
Total	71.41%	28.59%	56.28%	43.72%	57.99%	42.01%	1.23	0.68

This table reports the weighted price contribution for trade type classified by roundedness and matchedness, respectively, for DAX, MDAX, and SDAX stocks. WPC is the weighted price contribution. The share of trades (share of volume) relates to the percentage of trades (volume) in each size category. UR and RD relate to unrounded and rounded trades, respectively. Similarly, UM and MD stand for unmatched and matched trades. Info.Quat is the information quality measured by the ratio of WPC to the corresponding proportion of volume.

Table 9: Price Contribution, Roundedness And Matchedness of Trade Sizes for DAX, MDAX, and SDAX Stocks

		WPC				Share of Trade				Share of Volume			
		UR		RD		UR		RD		UR		RD	
DAX		UM	MD	UM	MD	UM	MD	UM	MD	UM	MD	UM	MD
	Small	12.61%	9.65%	-17.22%	7.49%	10.58%	9.68%	5.77%	4.28%	1.70%	1.85%	1.03%	0.83%
	Medium	27.72%	10.18%	-8.47%	16.17%	10.30%	14.32%	6.78%	8.55%	6.63%	9.47%	4.19%	5.76%
	Large	27.97%	7.71%	-3.68%	9.87%	8.71%	11.32%	4.78%	4.93%	22.37%	24.58%	11.72%	9.86%
		UR		RD		UR		RD		UR		RD	
		UM	MD	UM	MD	UM	MD	UM	MD	UM	MD	UM	MD
MDAX	Small	16.02%	8.65%	4.12%	6.33%	11.80%	11.31%	3.97%	4.03%	2.03%	2.04%	0.80%	0.77%
	Medium	20.51%	5.15%	4.17%	9.87%	12.46%	14.62%	6.31%	6.62%	8.38%	10.11%	4.40%	4.67%
	Large	15.50%	-1.41%	5.08%	5.99%	8.06%	11.05%	4.99%	4.77%	20.06%	24.18%	12.11%	10.46%
		UR		RD		UR		RD		UR		RD	
		UM	MD	UM	MD	UM	MD	UM	MD	UM	MD	UM	MD
SDAX	Small	9.56%	4.41%	5.11%	1.71%	12.69%	9.64%	4.98%	3.70%	2.01%	1.56%	0.96%	0.72%
	Medium	20.87%	2.49%	8.65%	3.32%	13.47%	11.81%	8.80%	5.81%	8.81%	7.79%	5.79%	3.84%
	Large	13.46%	7.11%	13.75%	9.55%	8.92%	8.58%	7.41%	4.18%	21.36%	18.66%	19.07%	9.43%

This table reports the weighted price contribution for each order category classified by size, roundedness, and matchedness. WPC is the weighted price contribution. The share of trades (share of volume) relates to the percentage of trades (volume) in each size type. UR and RD relate to unrounded and rounded trades, respectively. Similarly, UM and MD stand for unmatched and matched trades.

Table 10: Regression Results of Price Discovery and Roundedness and Matchedness of Trade Sizes for DAX, MDAX, and SDAX

	Trade types		DAX		MDAX		SDAX	
			Param.	P-value	Param.	P-value	Param.	P-value
Small	Unrounded	Unmatched	**0.004**	<0.01	**0.003**	<0.01	**0.003**	<0.01
		Matched	**0.003**	<0.01	**0.001**	<0.01	**0.001**	0.036
	Rounded	Unmatched	-0.006	<0.01	**0.001**	<0.01	**0.001**	0.028
		Matched	**0.002**	<0.01	**0.001**	<0.01	3.47E-04	0.384
Medium	Unrounded	Unmatched	**0.008**	<0.01	**0.003**	<0.01	**0.004**	<0.01
		Matched	0.002	0.336	-0.001	0.063	-0.002	0.014
	Rounded	Unmatched	-0.003	<0.01	0.000	0.401	0.001	0.350
		Matched	**0.005**	<0.01	**0.001**	<0.01	-2.61E-04	0.662
Large	Unrounded	Unmatched	0.006	0.150	1.06E-05	0.976	-0.003	0.098
		Matched	-0.001	0.848	-0.004	<0.01	-0.005	<0.01
	Rounded	Unmatched	-0.003	0.126	-0.001	<0.01	-0.002	0.014
		Matched	0.002	0.381	-4.10E-04	0.049	-1.62E-04	0.845
Constant			0.014	0.443	**0.015**	<0.01	**0.038**	<0.01
Adjusted R^2			0.0091		0.0076		0.0049	
Obs			44,904		74,856		41,940	

This table reports the results of weighted least square regressions of WPC on the percentage of the transaction (volume) and dummies based on roundedness, matchedness, and size, $CumPC_j^{s,t} = \sum_{j=1}^{k} \alpha_j \times dummy_j + \beta \times PcntVolume_j^{s,t} + \epsilon_j^{s,t}$, for DAX, MDAX, and SDAX stocks. Bold entries indicate that the coefficients are significant at the 5% level. Obs is the number of observations in the regression. From the sample, we exclude the days that have the same open and close prices.

Table 11: The Informational Quality of HPITs for DAX, MDAX and SDAX

	Panel a.1: WPC of HPITs			Panel b.1: Informational Quality of HPITs		
	WPC	Trade%	Volume%		WPC/Trade%	WPC/Volume%
DAX	73.6%	43.4%	16.8%	DAX	1.697	4.392
MDAX	65.5%	50.2%	18.7%	MDAX	1.305	3.508
SDAX	40.0%	40.8%	13.3%	SDAX	0.980	2.997
	Panel a.2: WPC of adjusted HPITs			Panel b.2: Informational Quality of adjusted HPITs		
	WPC	Trade%	Volume%		WPC/Trade%	WPC/Volume%
DAX	66.07%	37.02%	14.40%	DAX	1.785	4.589
MDAX	59.94%	41.43%	15.61%	MDAX	1.447	3.839
SDAX	39.02%	31.66%	10.66%	SDAX	1.232	3.660

Panel (a.1) reports the weighted price contribution (WPC) of HPITs for DAX, MDAX and SDAX stocks based on the regression, $CumPC_j^{s,t} = \sum_{j=1}^{k} \alpha_j \times dummy_j + \beta \times PcntVolume_j^{s,t} + \epsilon_j^{s,t}$. Trade% is the ratio of HPITs numbers over total trade numbers, and volume% is the ratio of HPITs volumes over total trade volumes. Panel (b.1) represents the corresponding informational qualities, measured by WPC/Trade% and WPC/Volume%. Panels (a.2 and b.2) report the weighted price contribution (WPC) of adjusted HPITs and corresponding informational qualities, respectively. To obtain adjusted HPITs, we exclude trades that are not in the same direction as the resulting midquote change.

Table 12: Temporal Price Impact of HPITs and Non-HPITs

	DAX Price Impact (bps)			MDAX Price Impact (bps)			SDAX Price Impact (bps)		
Minutes	HPITs	Non-HPITs	p-value	HPITs	Non-HPITs	p-value	HPITs	Non-HPITs	p-value
1	0.0313	0.0098	<0.01	0.0356	0.0096	<0.01	0.4818	0.1807	<0.01
2	0.0311	0.0097	<0.01	0.0374	0.0100	<0.01	0.5035	0.1875	<0.01
3	0.0307	0.0095	<0.01	0.0380	0.0102	<0.01	0.5314	0.2008	<0.01
5	0.0307	0.0093	<0.01	0.0383	0.0102	<0.01	0.5602	0.2066	<0.01
7	0.0312	0.0094	<0.01	0.0385	0.0103	<0.01	0.5863	0.2136	<0.01
10	0.0308	0.0093	<0.01	0.0378	0.0102	<0.01	0.5905	0.2140	<0.01
15	0.0305	0.0092	<0.01	0.0372	0.0099	<0.01	0.5975	0.2149	<0.01
20	0.0313	0.0095	<0.01	0.0374	0.0099	<0.01	0.6142	0.2187	<0.01
30	0.0307	0.0094	<0.01	0.0372	0.0099	<0.01	0.6530	0.2323	<0.01
60	0.0299	0.0093	<0.01	0.0387	0.0101	<0.01	0.6981	0.2323	<0.01

This table presents the temporal price impact of HPITs and Non-HPITs per unit for various intervals for DAX, MDAX, and SDAX stocks. Price impact (in basis points) is the effective price impact divided by the corresponding average volume. The effective price impact is defined as: *Effective price impact*$_{i,t}$ = $\frac{D_{i,t}(Midquote_{i,t+m} - Midquote_{i,t})}{Midquote_{i,t}}$. The p-value is for two-sample t-test with the null hypothesis that the means of the two samples are equal

Table 13: Intraday Price Discovery and Volume Proportion of HPITs and non-HPITs for DAX, MDAX and SDAX Stocks

Index	DAX						MDAX						SDAX					
	Price discovery (%)			Volume proportion (%)			Price discovery (%)			Volume proportion (%)			Price discovery (%)			Volume proportion (%)		
Trading hour	Total	HPITs	Non.HPITs	Total	HPITs	Non.HPITs	Total	HPITs	Non.HPITs	Total	HPITs	Non.HPITs	Total	HPITs	Non.HPITs	Total	HPITs	Non.HPITs
9h-10h	35.70%	20.51%	15.19%	15.71%	2.69%	13.02%	33.19%	17.73%	15.46%	12.86%	2.25%	10.61%	27.71%	9.43%	18.28%	13.46%	1.57%	11.89%
10h-11h	9.14%	7.45%	1.68%	12.23%	2.10%	10.13%	12.80%	8.55%	4.25%	11.03%	2.08%	8.95%	12.40%	4.93%	7.47%	11.67%	1.40%	10.27%
11h-12h	11.57%	8.59%	2.98%	10.93%	1.84%	9.10%	11.28%	7.89%	3.40%	10.51%	2.02%	8.49%	11.62%	5.88%	5.74%	10.80%	1.37%	9.42%
12h-13h	7.49%	5.48%	2.01%	8.47%	1.45%	7.03%	7.28%	5.45%	1.83%	8.68%	1.71%	6.97%	8.42%	3.03%	5.39%	8.93%	1.18%	7.75%
13h-14h	5.39%	4.88%	0.51%	7.56%	1.30%	6.26%	5.93%	4.32%	1.61%	8.10%	1.52%	6.58%	6.35%	2.65%	3.70%	7.37%	1.02%	6.35%
14h-15h	7.26%	6.86%	0.40%	9.18%	1.58%	7.60%	7.06%	5.55%	1.51%	9.38%	1.81%	7.57%	7.44%	3.37%	4.07%	10.05%	1.33%	8.72%
15h-16h	10.90%	9.13%	1.77%	12.01%	2.03%	9.98%	9.64%	7.20%	2.44%	12.04%	2.24%	9.81%	8.83%	2.66%	6.18%	12.22%	1.62%	10.60%
16h-17h	9.82%	8.00%	1.82%	14.45%	2.35%	12.10%	9.06%	6.08%	2.98%	15.88%	2.90%	12.98%	9.46%	3.70%	5.76%	14.78%	2.07%	12.70%
17h-17h30	2.73%	2.73%	-0.01%	9.46%	1.44%	8.03%	3.76%	2.73%	1.02%	11.52%	2.14%	9.38%	7.77%	4.30%	3.47%	10.72%	1.76%	8.96%
Total	100.00%	73.64%	26.36%	100.00%	16.77%	83.23%	100.00%	65.50%	34.50%	100.00%	18.67%	81.33%	100.00%	39.95%	60.05%	100.00%	13.33%	86.67%

This table reports intraday price discovery and volume proportion of HPITs and non-HPITs for DAX, MDAX and SDAX stocks. HPITs are high price impact trades and non-HPITs are the other trades. Price discovery is the associated price change expressed as a percentage of daily price change, and volume proportion is the associated volume expressed as a percentage of daily trade volumes.

Table 14: Estimation Results for HPITs Incidence

Param	γ_1	λ_1	λ_2	c	δ_1	δ_2	δ_3	δ_4	δ_5	ROC	Nb_Acc
ADS	0.975***	0.373***	-0.309***	-1.520***	0.158***	-2.280***	0.047***	4.685***	-0.002	0.566	0.579
ALV	0.985***	0.365***	-0.324***	-1.522***	-0.037*	1.016**	0.029***	2.889***	0.048***	0.553	0.596
BAS	0.974***	0.318***	-0.262***	-1.134***	-0.031	-0.232	0.017***	1.712***	0.013**	0.542	0.568
BAYN	0.978***	0.382***	-0.329***	-1.369***	0.015	-1.038***	0.032***	3.175***	-0.019***	0.556	0.571
BEI	0.977***	0.386***	-0.330***	-1.338***	0.114**	-1.848***	0.033***	3.272***	0.028**	0.559	0.577
BMW	0.918***	0.382***	-0.254***	-1.308***	0.051*	-0.519	0.074***	7.405***	-0.048***	0.556	0.572
CBK	0.994***	0.478***	-0.448***	-2.763***	-1.035***	-8.065***	0.025***	2.485***	0.002	0.715	0.656
CON	0.976***	0.350***	-0.289***	-1.445***	0.016	0.791***	0.025***	2.517***	-0.023**	0.562	0.569
DAI	0.982***	0.354***	-0.306***	-1.562***	-0.029	-1.194**	0.038***	3.843***	-0.024***	0.557	0.590
DB1	0.967***	0.387***	-0.306***	-1.235***	0.078**	-0.318	0.009**	0.889***	-0.006	0.566	0.575
DBK	0.975***	0.358***	-0.296***	-1.425***	-0.167***	-1.742**	0.034***	3.437***	-0.039***	0.557	0.576
DPW	0.957***	0.426***	-0.333***	-1.005***	0.525***	-7.746***	0.018***	1.759***	-0.022*	0.562	0.608
DTE	0.965***	0.411***	-0.328***	-1.450***	-0.542**	-14.597***	0.037***	3.735***	0.073***	0.568	0.587
EOAN	0.981***	0.364***	-0.308***	-1.219***	0.427***	-24.055***	0.002	0.225***	-0.008	0.567	0.617
FME	0.979***	0.357***	-0.298***	-1.378***	0.182***	-1.558**	0.014***	1.446***	-0.002	0.559	0.577
FRE	0.982***	0.325***	-0.280***	-1.295***	0.033	-1.048***	0.029***	2.888***	-0.021*	0.548	0.568
HEI	0.972***	0.363***	-0.302***	-1.128***	0.080*	-1.568***	0.025***	2.528***	0.002	0.550	0.576
HEN3	0.973***	0.384***	-0.318***	-1.282***	0.046	-2.017***	0.026***	2.599***	0.003	0.563	0.575
IFX	0.977***	0.362***	-0.301***	-1.460***	-0.314	16.386***	0.018***	1.782***	-0.046***	0.567	0.587
LHA	0.979***	0.421***	-0.367***	-1.464***	0.203	-8.463***	0.037***	3.688***	0.000	0.565	0.583
LIN	0.983***	0.345***	-0.303***	-1.198***	0.014	-1.078***	0.010**	1.029***	0.120***	0.549	0.571
LXS	0.990***	0.325***	-0.295***	-1.651***	0.001	-0.369	0.031***	3.066***	0.033***	0.554	0.559
MRK	0.954***	0.422***	-0.336***	-0.888***	0.215***	-1.872***	0.025***	2.497***	0.096***	0.558	0.568
MUV2	0.978***	0.363***	-0.309***	-1.324***	0.053**	-1.693***	0.032***	3.210***	0.127***	0.563	0.577
RWE	0.971***	0.384***	-0.309***	-1.378***	-0.006	-3.779***	0.021***	2.089***	0.040***	0.567	0.584
SAP	0.969***	0.378***	-0.306***	-1.250***	-0.008	-1.867***	0.030***	2.974***	-0.016**	0.557	0.594
SDF	0.985***	0.352***	-0.309***	-1.597***	-0.088*	-1.555**	0.033***	3.288***	0.000	0.568	0.585
SIE	0.968***	0.369***	-0.299***	-1.323***	0.093***	-1.145***	0.034***	3.401***	-0.008	0.556	0.566
TKA	0.974***	0.389***	-0.329***	-1.439***	0.073	1.092	0.043***	4.259***	0.025**	0.559	0.587
VOW3	0.980***	0.344***	-0.290***	-1.416***	0.039***	-0.488***	0.018***	1.849***	0.000	0.558	0.572
Mean	0.974	0.374	-0.312	-1.392	0.005	-2.428	0.028	2.820	0.011	0.564	0.582
Min	0.918	0.318	-0.448	-2.763	-1.035	-24.055	0.002	0.225	-0.048	0.542	0.559
Max	0.994	0.478	-0.254	-0.888	0.525	16.386	0.074	7.4049	0.127	0.715	0.656

This table reports the estimated results of the auto-logistic model for HPIT arrivals, $f\left(HPIT_t = 1\,|F_{t-1}\right) = p(\theta_t)$, where $p(\theta_t) = \frac{exp(\theta_t)}{1+exp(\theta_t)}$, $\theta_t = c+g_t+\Delta' M_{t-1}$, $g_t = \sum_{j=1}^{p} \gamma_j g_{t-j} + \sum_{j=1}^{l} \lambda_t HPIT_{t-j}$, and $\Delta' M_{t-1} = \delta_1 \times Range_{t-1} + \delta_2 \times AvgSpread_{t-1} + \delta_3 \times TradeVolume_{t-1} + \delta_4 \times TradeDensity_{t-1} + \delta_5 \times FirstHour_{t-1}$, for DAX stocks. The results remain qualitatively similar for MDAX and SDAX stocks. ROC signifies the Receiver Operating Characteristic test. Nb_Acc is the Count accuracy that takes 50% as the threshold to have value one. ***, ** and * denote either coefficient estimates that are significantly different from zero or test statistics that are significant at 1%, 5%, and 10%, respectively.

Table 15: Summary of HPITs Arrivals Modeling for DAX , MDAX and SDAX Stocks

	DAX			MDAX			SDAX		
Param	Average	Std	Nb_Sig/30	Average	Std	Nb_Sig/50	Average	Std	Nb_Sig/49
γ_1	0.974	0.001	30	0.971	0.002	50	0.894	0.028	47
λ_1	0.374	0.006	30	0.312	0.012	50	0.400	0.054	48
δ_1	0.005	0.068	16	0.320	0.141	31	0.624	0.774	21
δ_2	-2.428	0.947	25	-3.461	0.660	47	-0.926	1.805	16
δ_3	0.028	0.004	29	0.015	0.004	42	0.010	0.015	20
δ_4	2.821	0.007	30	-0.130	0.05	34	-0.345	0.750	21
δ_5	0.011	0.009	19	0.083	0.022	43	0.131	0.127	20
ROC	0.564	0.028		0.581	0.016		0.629	0.048	
Nb_Acc	0.582	0.018		0.569	0.010		0.636	0.038	

The table compares the estimated results of the auto-logistic model for HPIT arrivals, $f\left(HPIT_t = 1 \left| F_{t-1} \right.\right) = p(\theta_t)$, where $p(\theta_t) = \frac{exp(\theta_t)}{1+exp(\theta_t)}$, $\theta_t = c + g_t + \Delta' M_{t-1}$, $g_t = \sum_{j=1}^{p} \gamma_j g_{t-j} + \sum_{j=1}^{l} \lambda_t HPIT_{t-j}$, and $\Delta' M_{t-1} = \delta_1 \times Range_{t-1} + \delta_2 \times AvgSpread_{t-1} + \delta_3 \times TradeVolume_{t-1} + \delta_4 \times TradeDensity_{t-1} + \delta_5 \times FirstHour_{t-1}$, for DAX, MDX and SDAX stocks. ROC signifies the Receiver Operating Characteristic test. Nb_Acc is the Count accuracy that takes 50% as the threshold to have value one. For SDAX stocks, we excluded the GXI stock, which had few observations.

Table 16: The Effect of HPITs of DAX Stocks on 15-min Conditional Volatility

Param	ω	α	ϑ	β	γ	Q(15)	Q2(15)
ADS	-1.913***	-0.003	0.259***	0.844***	-2.038***	19.44	4.74
ALV	-2.002***	-0.045***	0.226***	0.836***	-1.304***	18.44	11.76
BAS	-1.900***	-0.025***	0.168***	0.845***	-2.703***	29.02	12.66
BAYN	-2.356***	-0.017**	0.216***	0.807***	-1.753***	17.84	5.99
BEI	-4.302***	0.018*	0.360***	0.657***	-0.616***	7.74	5.31
BMW	-0.113***	-0.011***	0.051***	0.990***	-3.372***	11.22	17.98
CBK	-0.696***	-0.041***	0.204***	0.935***	-4.686***	14.64	71.82
CON	-2.370***	-0.003	0.291***	0.800***	-1.856***	18.39	11.54
DAI	-2.081***	-0.011	0.255***	0.824***	-2.936***	12.62	6.70
DB1	-0.999***	0.083***	0.246***	0.915***	-3.134***	16.61	2.19
DBK	-1.197***	-0.034***	0.167***	0.896***	-4.077***	14.32	14.07
DPW	-1.998***	-0.013	0.314***	0.840***	-0.674***	14.10	9.60
DTE	-0.907***	0.025***	0.181***	0.927***	-3.290***	17.74	8.50
EOAN	-0.862***	-0.006	0.201***	0.928***	-2.202***	24.65	9.59
FME	-1.231***	0.007	0.277***	0.900***	-1.026***	27.20	30.15
FRE	-1.896***	-0.011	0.249***	0.847***	0.042	21.28	2.74
HEI	-1.995***	0.004	0.262***	0.831***	-1.139***	23.82	3.95
HEN3	-2.813***	0.025**	0.220***	0.773***	-0.526***	17.33	6.47
IFX	-1.241***	-0.016*	0.262***	0.894***	-2.893***	12.09	13.43
LHA	-2.668***	-0.009	0.378***	0.772***	-1.817***	12.11	7.08
LIN	-2.146***	-0.053***	0.167***	0.831***	-1.201***	28.36	12.72
LXS	-2.338***	-0.023***	0.313***	0.800***	-3.273***	21.53	8.09
MRK	-2.213***	0.005	0.253***	0.821***	-1.225***	10.92	5.29
MUV2	-1.404***	-0.014	0.237***	0.886***	-1.587***	15.93	11.83
RWE	-0.101***	0.000	0.080***	0.991***	-3.750***	16.02	21.43
SAP	-0.052***	-0.015***	0.039***	0.996***	-2.682***	14.97	44.75
SDF	-0.154***	-0.033***	0.129***	0.986***	-3.136***	14.90	89.48
SIE	-0.282***	-0.039***	0.099***	0.977***	-4.240***	32.80	5.47
TKA	-3.460***	0.030***	0.459***	0.697***	-1.886***	17.01	4.04
VOW3	-2.295***	-0.117***	0.413***	0.805***	-2.265***	12.06	4.27
Mean	-1.67	-0.01	0.23	0.86	-2.24	17.84	15.45
Min	-4.30	-0.12	0.05	0.66	-4.69	7.74	2.19
Max	-0.05	0.08	0.16	1.00	0.04	32.80	89.48

This table reports the estimated results of the EGARCH model, $log(\sigma_i^2) = \omega + \sum_{j=1}^{p} \alpha_j g(Z_{i-j}) + \sum_{j=1}^{q} \beta_j log(\sigma_{i-j}^2) + \gamma HPIT_{\text{i-1}}\%$, for 15-min deseasonalized returns for DAX stocks. The results remain qualitatively similar for the 30-min interval. Q(15) and Q2(15) relate to Ljung-Box statistics on 15 lagged standardized residuals and squared standardized residuals derived from the model. The 5% critical value is 24.99. ***, ** and * denote either coefficient estimates that are significantly different from zero or test statistics that are significant at the 1%, 5%, and 10%, respectively.

Table 17: Summary of the Effect of HPITs on 15-min Conditional Volatility for DAX , MDAX and SDAX Stocks

	DAX			MDAX			SDAX		
Param	Average	Std	Nb_Sig/30	Average	Std	Nb_Sig/50	Average	Std	Nb_Sig/48
ω	-1.666	1.025	30	-1.86	0.87	50	-2.04	1.67	45
α	-0.011	0.033	18	-0.01	0.03	30	-0.02	0.06	31
θ	0.233	0.099	30	0.29	0.07	50	0.26	0.12	47
β	0.862	0.085	30	0.84	0.07	50	0.81	0.17	48
γ	-2.241**	1.204	29	-1.11**	0.51	50	-0.42**	0.53	40
Q(15)	17.84	5.99	13	28.20	13.54	6	34.22	23.81	10
Q2(15)	15.45	19.89	24	11.45	7.4	42	11.92	9.80	39
Q(15)_raw	27.54	19.21		38.90	19.00		53.78	146.40	
Q2(15)_raw	170.45	221.21		252.82	197.21		153.84	279.90	

The table compares the estimated results of the EGARCH model, $log(\sigma_i^2) = \omega + \sum_{j=1}^{p} \alpha_j g(Z_{i-j}) + \sum_{j=1}^{q} \beta_j log(\sigma_{i-j}^2) + \gamma HPIT_{\text{i-1}}\%$, for DAX, MDX and SDAX stocks.

Q(15) and Q(15)_raw relate to Ljung-Box statistics on 15 lagged standardized residuals derived from the model and raw deseasonalized returns. Q2(15) and Q2(15)_raw are Ljung-Box statistics on 15 lagged squared standardized residuals derived from the model and raw deseasonalized returns. The 5% critical value is 24.99. Nb_Sig is the number of stocks with significant parameters. ** indicates that the t-statistics of the differences of the parameters between the actual group and the other two groups are significant at the 5% level. For SDAX stocks, we excluded two stocks that had fewer than 50 observations.

Table 18: Effect of HPITs on Autocorrelation for DAX Stocks

	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,3}$	$\delta_{i,4}$	$\psi_{i,1}$	$\psi_{i,2}$	$\psi_{i,3}$	$\psi_{i,4}$	$\delta_{i,1}+\delta_{i,2}$	Adjusted R^2
ADS	-0.7965***	-0.2587	0.0249***	-0.0106	0.0004	-0.0150	0.0011***	-0.0008	-1.0552***	0.145
ALV	-1.1133***	-0.0812	-0.0106	0.0290	-0.0016***	-0.0090	0.0014***	0.0008	-1.1945***	0.244
BAS	-1.0273***	-0.0052	0.0118	-0.0397	-0.0021**	0.0316	0.0010***	0.0014	-1.0325***	0.140
BAYN	-0.6011***	0.1881	0.0821***	0.0850**	0.0005	0.0012	0.0009***	-0.0020	-0.4130*	0.121
BEI	-0.9144***	-0.0126	0.0034	-0.0221	0.0019	-0.0206	0.0010***	0.0004	-0.9270***	0.144
BMW	-1.1628***	-0.0575	-0.0121	-0.0471**	-0.0009	0.0380	0.0009***	0.0005	-1.2203***	0.093
CBK	-0.5340***	0.6621***	0.0086	0.0249	0.0022	0.1569	0.0038***	-0.0003	0.1281	0.360
CON	-0.8454***	-0.1152	0.0260	0.0264	0.0004	-0.0029	0.0012***	0.0007	-0.9605***	0.124
DAI	-1.1257***	-0.2624	0.0050	-0.0821*	0.0005	-0.0170	0.0016***	0.0005	-1.3881***	0.184
DB1	-0.7204***	0.0608	-0.0022	-0.0160	0.0048	-0.0586**	0.0012***	0.0027	-0.6596**	0.169
DBK	-1.1506***	-0.2265	0.0283	-0.0758	-0.0036**	0.0364	0.0020***	0.0007	-1.3771***	0.249
DPW	-1.1992***	-0.0518	0.0545***	0.0106	-0.0018	-0.0637	0.0010***	0.0016	-1.2510***	0.177
DTE	-1.3341***	-0.2429	0.0378	-0.0022	-0.0009	-0.0116	0.0011***	0.0012	-1.5770***	0.197
EOAN	-0.8597***	-0.3674**	0.0065	0.0285	-0.0120**	-0.1999	0.0016***	-0.0007	-1.2271***	0.272
FME	-1.1554***	-0.1251	0.0021	-0.0176	-0.0018	0.0207	0.0009***	0.0001	-1.2805***	0.135
FRE	-0.5119***	0.0809	0.0540**	-0.0244	-0.0011	-0.0071	0.0008***	-0.0018	-0.4309**	0.115
HEI	-1.1860***	0.3942**	0.0502*	0.0325	-0.0004	-0.0256	0.0008***	0.0029*	-0.7918***	0.112
HEN3	-0.8704***	0.2582	-0.0190	0.0533**	0.0005	-0.0265*	0.0008***	0.0004	-0.6122***	0.093
IFX	-1.0782***	-0.0730	0.0867***	0.0045	-0.0134	-0.1552	0.0014***	0.0015	-1.1512***	0.177
LHA	-0.5453***	0.1934	0.0454	-0.0102	-0.0046	-0.1954**	0.0015***	-0.0007	-0.3519	0.160
LIN	-0.7941***	-0.5138**	0.0321	0.0213	-0.0009**	-0.0142	0.0009***	0.0011	-1.3079***	0.160
LXS	-0.3159**	-0.0278	0.0628**	0.0043	0.0019	-0.0024	0.0018***	0.0004	-0.3437*	0.209
MRK	-0.6054***	0.5134***	0.0754***	0.0103	-0.0002	-0.0079	0.0010***	-0.0011	-0.0920	0.189
MUV2	-0.8635***	-0.2252	0.0662**	0.0724***	-0.0006	0.0029	0.0012***	0.0010	-1.0887***	0.189
RWE	-0.9280***	-0.5442***	0.0067	0.0036	-0.0034	-0.1124*	0.0014***	0.0024***	-1.4722***	0.184
SAP	-1.2399***	-0.2036	-0.0302	-0.0325	0.0012	-0.0219	0.0008***	-0.0006	-1.4434***	0.108
SDF	-0.9489***	-0.0649	0.0053	-0.0383**	-0.0145***	0.0464	0.0014***	0.0016*	-1.0138***	0.348
SIE	-1.0153***	-0.2589	-0.0175	-0.0366	-0.0024	-0.0255	0.0011***	-0.0017	-1.2742***	0.195
TKA	-0.5232***	-0.2820	0.0016	0.0389	-0.0007	0.0042	0.0021***	0.0009	-0.8052***	0.207
VOW3	-1.1286***	-0.3820*	0.0148	-0.0211	-0.0008	0.0100	0.0018***	0.0008	-1.5106***	0.289
Bonferroni p-value	<0.001	0.050	0.003	0.254	<0.001	1	<0.001	0.186	<0.001	
Nb significant	30	7	9	6	6	4	30	3	27	
Mean	-0.9031	-0.0677	0.0234	-0.0010	-0.0018	-0.0215	0.0013	0.0005	-0.9708	0.1829
Median	-0.9212	-0.0771	0.0133	0.0007	-0.0008	-0.0103	0.0011	0.0006	-1.0720	0.1768
Min	-1.3341	-0.5442	-0.0302	-0.0821	-0.0145	-0.1999	0.0008	-0.0020	-1.5770	0.0932
Max	-0.3159	0.6621	0.0867	0.0850	0.0048	0.1569	0.0038	0.0029	0.1281	0.3596

The table reports the estimated results of the following model:

$$R_{i,t} = (\delta_{1,i} HPIT\%_{i,t\text{-}1} * 1_{(contrarian)} + \delta_{2,i} HPIT\%_{i,t\text{-}1} * 1_{(herding)} + \delta_{3,i} Volume_{t-1}) R_{i,t-1} + \psi_{i,1} Range_{i,t-1} + \psi_{i,2} Spread_{i,t-1} + \psi_{i,3} Price_{i,t-1} + \sum_{k=2}^{5} \beta_{i,k} R_{i,t-k} + \sum_{k=1}^{N} \alpha_{i,k} D_{kt} + e_{i,t},$$

for DAX stocks. Bonferroni p-value is the p-value based on the Bonferroni correction and is calculated as $min(1, min(p_1, ..., p_n) \times n)$, where p_i is the p-value of a given parameter from the estimation of the model for the ith stock. A Bonferroni p-value less than 0.05 suggests that the null hypothesis is rejected jointly across all stocks at the 5% significance level or higher. Nb significant is the number of stocks for which the results are statistically significant at the 10% level. ***, ** and * denote either coefficient estimates that are significantly different from zero or test statistics that are significant at the 1%, 5%, and 10% levels, respectively.

Table 19: Summary of the Effect of HPITs on Autocorrelation for DAX, MDAX and SDAX Stocks

	Avg. δ_1	Nb. Sig.Pos./Total	Nb. Sig.Neg./Total	Avg.δ_2	Nb. Sig.Pos./Total	Nb. Sig.Neg./Total	Avg.$\delta_1 + \delta_2$	Nb. Sig.Pos./Total	Nb. Sig.Neg./Total
DAX	-0.903**	0/30	**30/30**	-0.068*	3/30	4/30	-0.971**	0/30	**27/30**
MDAX	-0.790**	0/50	**50/50**	0.215**	**29/50**	1/50	-0.575**	0/50	**45/50**
SDAX	-0.393**	0/42	**31/48**	0.005*	4/48	4/48	-0.387**	0/48	**15/48**

The table compares the estimated results of the following model:

$$R_{i,t} = (\delta_{1,i} HPIT\%_{i,t\text{-}1} * 1_{(contrarian)} + \delta_{2,i} HPIT\%_{i,t\text{-}1} * 1_{(herding)} + \delta_{3,i} Volume_{t-1}) R_{i,t-1} + \psi_{i,1} Range_{i,t-1} + \psi_{i,2} Spread_{i,t-1} + \psi_{i,3} Price_{i,t-1} + \sum_{k=2}^{5} \beta_{i,k} R_{i,t-k} + \sum_{k=1}^{N} \alpha_{i,k} D_{kt} + e_{i,t},$$

for DAX, MAX and SDAX stocks. Nb. Sig.Pos./Total (Nb. Sig.Neg./Total) is the number of stocks for which the corresponding parameter is positive (negative) and significant over the total number of stocks in the group. Bold entries indicate that the results are representative for the majority of stocks. ** indicates that the t-statistics of the differences of the parameters between the actual group and the other two groups are significant at the 5% level. * specifies that the t-statistic of the difference of the parameters between the actual group and MDAX stocks is significant at the 5% level. For SDAX stocks, we excluded two stocks that had fewer than 50 observations.

Table 20: The Effect of HPITs on Market Efficiency (Variance Ratio-based Measure)

	DAX		MDAX		SDAX	
$HPIT_{t-1}$	-0.46***	-0.36***	-0.05	-0.06	0.10**	0.09**
	(-2.75)	(-3.03)	(-0.84)	(-1.23)	(2.13)	(2.45)
$Price_{t-1}$	-0.123***	0.014	-0.037	0.013	-0.020	0.004
	(-2.60)	(1.10)	(-0.80)	(1.60)	(-1.46)	(1.02)
$Range_{t-1}$	-0.034***	-0.041***	-0.010	-0.015*	-0.016***	-0.001
	(-4.20)	(-3.79)	(-1.16)	(-1.78)	(-3.93)	(-0.54)
$Spread_{t-1}$	-0.091***	-0.026***	-0.022***	-0.009***	-4.0E-04	1.60E-04
	(-8.40)	(-4.59)	(-8.91)	(-6.99)	(-0.73)	(0.45)
$Constant$		0.536***		0.488***		0.371***
		(10.85)		(15.71)		(21.16)
Effet	Fixed	Random	Fixed	Random	Fixed	Random
Observation	7484	7484	12476	12476	6990	6990
Adjusted R^2	0.053	0.018	0.023	0.012	0.003	0.004
No.tickers	30	30	50	50	49	49

The table presents the fix- and random-effect panel regression results on variance-ratio based market price efficiency:

$$\mid M_r(q) - 1 \mid_{i,t} \quad = \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t},$$
$$\mid M_r(q) - 1 \mid_{i,t} \quad = \mu + \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t} \ ,$$

for DAX, MDAX, and SDAX stocks with 4h measurement interval. $HPIT_{i,t-1}$ is the the proportion of HPITs for stock i during the period $i-1$, $Range_{i,t-1}$ relates to the range between maximum and minimum price, and $Spread_{i,t-1}$ and $Price_{i,t-1}$ are the average spread and price. Results remain qualitatively similar for 2h measurement interval. No.tickers is the number of tickers used in estimation. For SDAX stocks, we excluded the ticker HBH3 (HORNBACH HOLD.VZO O.N) that had only 15 trades on daily average. ***, ** and * denote either coefficient estimates that are significantly different from zero or test statistics that are significant at the 1%, 5%, and 10% levels, respectively.

Table 21: The Effect of HPITs on Market Efficiency (Autocorrelation-based Measure)

	DAX		MDAX		SDAX	
$HPIT_{t-1}$	-0.10**	-0.06*	-0.03*	-0.02	0.02	0.02
	(-2.41)	(-1.74)	(-1.74)	(-1.17)	(1.07)	(1.25)
$Price_{t-1}$	-0.019**	0.002	0.006	-0.001	2.66E-05	0.004**
	(-2.02)	(0.80)	(0.51)	(-0.36)	(0.026)	(2.01)
$Range_{t-1}$	-0.002	-0.005***	7.26E-05	0.001	-0.002	-3.0E-04
	(-1.02)	(-2.69)	(0.02)	(0.18)	(-0.82)	(-0.12)
$Spread_{t-1}$	-0.016***	-0.005***	-0.003***	-0.002***	0.001	0.001**
	(-9.30)	(-4.52)	(-6.76)	(-5.10)	(1.41)	(2.46)
$Constant$		0.122***		0.129***		0.092***
		(12.80)		(15.61)		(14.96)
Effet	Fixed	Random	Fixed	Random	Fixed	Random
Observation	7484	7484	12476	12476	6990	6990
Adjusted R^2	0.032	0.011	0.011	0.007	0.001	0.006
No.tickers	30	30	50	50	49	49

The table presents the fix- and random-effect panel regression results on variance-ratio based market price efficiency:

$$\begin{aligned} \mid \rho \mid_{i,t} &= \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t}, \\ \mid \rho \mid_{i,t} &= \mu + \alpha_i + \gamma_t + \beta_1 \times HPIT_{i,t-1} + \beta_2 \times log(Price_{i,t-1}) + \beta_3 \times Range_{i,t-1} + \beta_4 \times Spread_{i,t-1} + \epsilon_{i,t}, \end{aligned}$$

for DAX, MDAX, and SDAX stocks with 4h measurement interval. $HPIT_{i,t-1}$ is the the proportion of HPITs for stock i during the period $i-1$, $Range_{i,t-1}$ relates to the range between maximum and minimum price, and $Spread_{i,t-1}$ and $Price_{i,t-1}$ are the average spread and price. Results remain qualitatively similar for 2h measurement interval. No.tickers is the number of tickers used in estimation. For SDAX stocks, we excluded the ticker HBH3 (HORNBACH HOLD.VZO O.N) that had only 15 trades on daily average. ***, ** and * denote either coefficient estimates that are significantly different from zero or test statistics that are significant at the 1%, 5%, and 10% levels, respectively.

Acknowledgements

We thank Yann Bilodeau for his help in constructing the dataset and comments. We also thank Yuferova Darya, Antonia Kirilova, Vineet Virmani, Jun Wen, and participants at the 2019 Cross Country Perspective in Finance Symposium, 2019 China Conference of CEA(UK/Europe), the 31th Asian FA Annual Meeting, the 27th Conference on the theories and practices of securities and financial markets and 2019 Paris December Finance Meeting for their remarks. This work was supported by the Social Sciences and Humanities Research Council in Canada [435-2012-1503]; the Canada Foundation for Innovation [219145]; and the Fonds de Recherche sur la Société et la Culture de Québec [2019-NP-253336].

References

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

Alexander, G.J., Peterson, M.A., 2007. An analysis of trade-size clustering and its relation to stealth trading. Journal of Financial Economics 84, 435 – 471.

Ascioglu, A., Comerton-Forde, C., McInish, T.H., 2011. Stealth trading: The case of the Tokyo stock exchange. Pacific-Basin Finance Journal 19, 194 – 207.

Avramov, D., Chordia, T., Goyal, A., 2006. The impact of trades on daily volatility. Review of Financial Studies 19, 1241–1277.

Ball, C.A., Torous, W.N., Tschoegl, A.E., 1985. The degree of price resolution: The case of the gold market. Journal of Futures Markets 5, 29–43.

Banerjee, S., Davis, J., Gondhi, N., 2018. When transparency improves, must prices reflect fundamentals better? Review of Financial Studies 31, 2377–2414.

Barclay, M.J., Warner, J.B., 1993. Stealth trading and volatility. Journal of Financial Economics 34, 281 – 305.

Barron, O.E., Kim, O., Lim, S.C., Stevens, D.E., 1998. Using analysts' forecasts to measure properties of analysts' information environment. Accounting Review 73, 421–433.

Barron, O.E., Stuerke, P.S., 1998. Dispersion in analysts' earnings forecasts as a measure of uncertainty. Journal of Accounting, Auditing & Finance 13, 245–270.

Bilodeau, Y., 2013. Xetraparser [computer software]. HEC Montréal .

Blasco, N., Corredor, P., 2017. The information environment, informed trading, and volatility. Journal of Behavioral Finance 18, 202–218.

Boehmer, E., Kelley, E.K., 2009. Institutional investors and the informational efficiency of prices. Review of Financial Studies 22, 3563–3594.

Cai, B.M., Cai, C.X., Keasey, K., 2006. Which trades move prices in emerging markets?: Evidence from china's stock market. Pacific-Basin Finance Journal 14, 453 – 466.

Campbell, J.Y., Grossman, S.J., Wang, J., 1993. Trading volume and serial correlation in stock returns. Quarterly Journal of Economics 108, 905–939.

Cespa, G., Vives, X., 2015. The beauty contest and short-term trading. Journal of Finance 70, 2099–2154.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., Clara, V., 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. Journal of Finance 69, 2045–2084.

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

Collin-Dufresne, P., Fos, V., 2015. Do prices reveal the presence of informed trading? Journal of Finance 70, 1555–1582. doi:10.1111/jofi.12260.

Collin-Dufresne, P., Fos, V., 2016. Insider trading, stochastic liquidity, and equilibrium prices. Econometrica 84, 1441–1475.

Colombo, L., Femminis, G., Pavan, A., 2014. Information acquisition and welfare. Review of Economic Studies 81, 1438–1483.

Conrad, J., Wahal, S., Xiang, J., 2015. High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics 116, 271–291.

Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics 86, 479 – 512.

Cox, D.R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., Lauritzen, S.L., 1981. Statistical Analysis of Time Series: Some Recent Developments [with Discussion and Reply]. Scandinavian Journal of Statistics 8, 93–115.

Diamond, D.W., 1985. Optimal release of information by firms. Journal of Finance 40, 1071–1094.

Dufour, A., Engle, R.F., 2000. Time and the price impact of a trade. Journal of Finance 55, 2467–2498.

Dugast, J., Foucault, T., 2018. Data abundance and asset price informativeness. Journal of Financial Economics 130, 367–391.

Foster, F.D., Viswanathan, S., 1990. A theory of the interday variations in volume, variance, and trading costs in securities markets. Review of Financial Studies 3, 593–624.

Friedman, M., 1953. The case for flexible exchange rates. in Milton Friedman (ed), Essays in Positive Economics, University of Chicago Press, Chicago, IL .

Gao, P., Liang, P., 2013. Informational feedback, adverse selection, and optimal disclosure policy. Journal of Accounting Research 51, 1133–1158.

Garvey, R., Wu, F., 2014. Clustering of intraday order-sizes by uninformed versus informed traders. Journal of Banking and Finance 41, 222 – 235.

Goldstein, I., Yang, L., 2015. Information diversity and complementarities in trading and information acquisition. Journal of Finance 70, 1723–1765.

Grossman, S.J., Miller, M.H., Cone, K.R., Fischel, D.R., Ross, D.J., 1997. Clustering and competition in asset markets. The Journal of Law & Economics 40, 23–60.

Gu, F., Wang, W., 2005. Intangible assets, information complexity, and analysts' earnings forecasts. Journal of Business Finance & Accounting 32, 1673–1702.

Han, B., Tang, Y., Yang, L., 2016. Public information and uninformed trading: Implications for market liquidity and price efficiency. Journal of Economic Theory 163, 604 – 643.

Harris, L., 1991. Stock price clustering and discreteness. Review of Financial Studies 4, 389–415.

Hayek, F.A., 1945. The use of knowledge in society. American Economic Review 35, 519–530.

Hellwig, M.F., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory 22, 477 – 498.

Henderson, B.J., Pearson, N.D., Wang, L., 2014. New evidence on the financialization of commodity markets. Review of Financial Studies 28, 1285–1311.

Hodrick, L., P.C., M., 2005. Clustering in US stock prices after decimalization. Unpublished working paper, Rice University. .

Ikenberry, D., Weston, J., 2003. Liquidityg. Unpublished working paper, Columbia University .

Johnson, T.C., 2004. Forecast dispersion and the cross section of expected returns. Journal of Finance 59, 1957–1978.

Korajczyk, R.A., Murphy, D., 2018. High-frequency market making to large institutional trades. Review of Financial Studies 32, 1034–1067.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lai, S., Ng, L., Zhang, B., 2014. Does PIN affect equity prices around the world? Journal of Financial Economics 114, 178 – 195.

Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Lo, A.W., MacKinlay, A.C., 1989. The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. Journal of Econometrics 40, 203–238.

Menkhoff, L., Schmeling, M., 2010. Whose trades convey information? Evidence from a cross-section of traders. Journal of Financial Markets 13, 101 – 128.

Milgrom, P., Stokey, N., 1982. Information, trade and common knowledge. Journal of Economic Theory 26, 17 – 27.

Moulton, P.C., 2005. You can't always get what you want: Trade-size clustering and quantity choice in liquidity. Journal of Financial Economics 78, 89 – 119.

Niederhoffer, V., Osborne, M.F.M., 1966. Market making and reversal on the stock exchange. Journal of the American Statistical Association 61, 897–916.

O'Hara, M., 2015. High frequency market microstructure. Journal of Financial Economics 116, 257 – 270.

O'Hara, M., Yao, C., Ye, M., 2014. What's not there: Odd lots and market data. Journal of Finance 69, 2199–2236.

Peress, J., Schmidt, D., 2020. Glued to the tv: Distracted noise traders and stock market liquidity. Journal of Finance 75, 1083–1133.

Rajan, R., Servaes, H., 1997. Analyst following of initial public offerings. Journal of Finance 52, 507–529.

Rosch, D.M., Subrahmanyam, A., van Dijk, M.A., 2016. The dynamics of market efficiency. Review of Financial Studies 30, 1151–1187.

Shephard, N., 1995. Generalized linear autoregressions. Working paper, Nuffield College, University of Oxford .

Skjeltorp, J.A., Sojli, E., Tham, W.W., 2016. Flashes of trading intent at NASDAQ. Journal of Financial and Quantitative Analysis 51, 165–196.

Swets, J.A., 1986. Indices of discrimination or diagnostic accuracy: Their rocs and implied models. Psychological Bulletin 99(1), 100–117.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–1293.

Van Kervel, V., Menkveld, A., 2019. High-frequency trading around large institutional orders. Journal of Finance 74, 1091–1137.

Wang, J., 1993. A model of intertemporal asset prices under asymmetric information. Review of Economic Studies 60, 249–282.

Wyckoff, 1963. The psychology of stock market timing. Prentice-Hall, Englewood Cliffs, NJ .