Essex Finance Centre
Working Paper Series

Working Paper No 85: 12-2023

"High Frequency Trading and Stock Herding"

"Servanna Mianjun Fu, Neil M. Kellard, Thanos Verousis and Iordanis Kalaitzoglou"

Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ
Web site: http://www.essex.ac.uk/ebs/

High Frequency Trading and Stock Herding

Servanna Mianjun Fu [a, *], Neil M. Kellard [a], Thanos Verousis [b], Iordanis Kalaitzoglou [c]

[a] Essex Business School, [b] Vlerick Business School, [c] Université de Nantes

November, 2023

Abstract

Using Trade and Quote (TAQ) data to infer variation in High frequency Trading (HFT) for the US equity markets and HFT start and colocation dates for a sample of 10 international exchanges, we find that increases in HFT activity lead to a significant increase in stock herding. The effect of HFT on herding is more pronounced for large-cap stocks, higher liquidity periods and during more volatile days. HFT activities are strongly associated with non-fundamental herding and encourage information cascades that induce price inefficiencies, suggesting changes to market design might be warranted.

Keywords: High Frequency Trading; HFT; Herding; Colocation; Information cascades; Fundamental information

JEL Codes: G23; G4; G14

---

[*] Corresponding author.
E-mail address: servanna.fu@essex.ac.uk (S. M. Fu).

1. Introduction

Algorithmic Trading (AT) and High Frequency Trading (HFT) now account for approximately 50% of trading volume in U.S. stock markets (Breckenfelder, 2020 and NASDAQ, 2023). The explosion of AT and HFT activities has fuelled an intense debate between HFT firms, institutional investors, regulators and academics about the effects of the HFT 'arms race' on financial markets.[2] Whilst AT and HFT potentially improve liquidity and price discovery, recent studies suggest that HFT activity increases comovement in returns and liquidity, discourages information acquisition and reduces long-term price informativeness.[3] In this paper, we examine a new prospective destabilising effect of HFT activities, namely an increase in stock herding.

Herding refers to the tendency of market participants to suppress their own beliefs and imitate the actions of others. In the social learning literature (see Banerjee, 1992 and Bikhchandani, Hirshleifer and Welch, 1992), whether an individual would have chosen differently if she had acted on her own beliefs alone distinguishes informational cascades from herding. Informational cascades imply that individuals ignore their own set of beliefs when making a decision, simply imitating others, whereas herd behaviour occurs when a large number of individuals take the same decision, not necessarily imitating others (Çelen and Kariv, 2004). We conjecture that the high frequency arms race causes significant increases in stock herding.

There are several reasons to believe that HFT induces herding. First, HFT trading strategies are highly correlated (see Chaboud et al., 2014 and Boehmer, Li and Saar, 2018). Second, High Frequency Traders (HFTs) trade as back-runners, that is, they observe past order

---

[2] Aquilina, Budish and Neill (2022) lists the regulatory investigations into HFT, including a list of proposals to curb HFT by introducing speed bumps.

[3] See Malceniece, Malcenieks and Putniņš (2019); Weller (2018); Gider, Schmickler and Westheide (2019); Breckenfelder (2020); Aquilina et al., (2022).

flow of fundamental investors and trade in the same direction (see Yang and Zhu, 2019). Third, HFTs engage in short-term predatory trading, that is, they tend to trade in the direction of institutional investors before eventually changing direction and trading against them (see Van Kervel and Menkveld, 2019). Fourth, HFT activity has enhanced price discovery, implying that HFTs trade on the same fundamental information (see Brogaard, Hendershott and Riordan, 2014; Chaboud, Chiquoine, Hjalmarsson and Vega, 2014). Fifth, as HFTs trade in order to adjust their inventory, they create noise (see Benos and Sagade, 2012), thereby inducing non-fundamental herding. Despite the above though, to date, there is no evidence in the extant literature on the impact of HFT on stock herding.

Our paper fills this gap in the literature. We rely on two datasets. The first consists of all Trade and Quote (TAQ) data for the constituents of the S&P100 from January 2015 to December 2017. Second, we use the index constituents from 10 exchanges in nine countries around the world. We employ both the HFT effective start date and the colocation date i.e., the date each exchange started housing the trading firms' computer servers within the exchange's data centre, as the HFT start date. The use of colocation dates complements our analysis in the U.S. market as it can be seen as a quasi-natural experiment associated with a significant increase in HFT activity (see Aitken Cumming and Zhan, 2015). As the timing of colocation decisions vary across exchanges and due to the staggered introduction of colocation services, we avoid any potential identification bias in our results which are also not confounded by a single unrelated event.

Our main finding is that increased HFT activity is strongly associated with stock herding. In particular, HFT-related herding is more pronounced during more volatile periods and periods of higher liquidity. This result is robust to the proxy employed for HFT. Interestingly, when we decompose trading activity into more active and less active HFT, we show that less active HFT inspires localised herding (anti-herding)[4] causing cross-sectional dispersion to increase beyond rational levels. We also demonstrate that herding is more pronounced among large-cap stocks by considering the liquidity provision by HFT and the channel through which return comovement is impacted. This is in line with the literature that HFTs are typically more active in this context (see Brogaard, Carrion, Moyaert, Riordan, Shkilko and Sokolov, 2018, and Jarnecic and Snape, 2014).

Subsequently, we investigate if HFT is related to fundamental or non-fundamental herding. Non-fundamental herding induces information cascades that are fragile and unstable whereas, in fundamental herding, the actions of market participants are still informative (see Çelen and Kariv, 2004). We show that HFT-related herding is strongly associated with non-fundamental herding. In line with our prior results, we show that the link between HFT activities and non-fundamental herding is stronger for large-cap stocks. Lastly, we use colocation dates and HFT effective start dates (see Aitken, Cumming, and Zhan, 2015) as exogenous shocks to HFT participation. Analogously to the results for the U.S. market, we show that HFT effective start dates, and colocation dates are associated with stock herding.

We contribute to the behavioural finance literature as well as the HFT literature. First, by decomposing market activities into more and less active HFT, we produce evidence that herding is distinctly related to more HFT participation. Second, by documenting significant evidence after HFT start dates, we provide further evidence into how market microstructure affects financial market outcomes (see Gider, Schmickler and Westheide, 2019). Third, by identifying that HFT is related to non-fundamental herding, we add to the growing body of evidence that HFT activities induce price inefficiencies (see Weller, 2018 and Gider et al., 2019). Finally, our findings show further support for the notion that correlated trading strategies

[4] Anti-herding implies that market participants largely ignore market-wide information and trade against the market consensus (see Sibande et al., 2021 and Gębka and Wohar, 2013).

and not diffusion of market-wide information may be responsible for an increase in comovements in returns (see Malceniece, Malcenieks and Putninš, 2019).

Our results have implications for policymakers and investors. Exchanges around the world are keen to attract HFTs as they are a valuable source of revenue, thus offering colocation services in exchange for trading fees (see Jørgensen, Skjeltorp and Ødegaard, 2018). Our analysis shows that there is a potential downside to the increased liquidity that comes with HFT in the form of increased price inefficiencies and fragility. Our finding that HFTs herd implies there is a level of systematic risk in the market that is associated with correlated trading strategies. From an investor perspective, by identifying the existence of anti-herding, we signal that there is insufficient diversification within a group of market participants (see Gębka and Wohar, 2013).

The rest of the paper is organized as follows. In Section 2, we review the relevant literature on stock herding, AT and HFT. In Section 3, we discuss the sample, variables, and methods and in Section 4, we present the analysis. In Section 5, we conclude the paper.

2. Literature review and hypothesis development

2.1 Herding

Conceptually, herding implies that investors suppress their private signals and resort or mimic their peers' trading behaviour following observation of their activities (Hirshleifer and Teoh, 2003). To the extent that investors fully suppress their own private information, herding leads to inefficiencies as equilibrium prices do not reflect the entire set of available information (see Banerjee, 1992). In that respect, whilst the decision to herd may be a rational one, i.e., an investor with little or no information will be better off following others rather than acting on their own private information, nevertheless it leads to inefficiencies. Equally, investors may herd as a response to irrational motivations. For example, Tedeschi, Iori and Gallegati (2012) and Kellard et al. (2017) demonstrate the importance of networks in the emergence of herding and Shleifer and Summers (1990) show that investors who suffer from systematic biases may induce herding. Relatedly, Cui, Gebka and Kallinterakis (2019) show that closed-end fund herding is mostly noise driven, reflecting the notion that the closed-end fund clientele mainly consists of retail investors (and thus is more susceptible to biases).

Theoretically, Christie and Huang (1995) demonstrate that, under rational asset pricing expectations, the dispersion of individual returns increases with the absolute value of market returns as individual stocks differ in their sensitivity to the market return. When however investors suppress their own beliefs in favour of the market consensus (i.e., the market return), the return dispersion around the market return is relatively low. In the presence of herding, return dispersion around the market return will increase at a decreasing rate or even be negative (see Chang, Cheng and Khorana, 2000; Andrikopoulos Kallinterakis, Leite Ferreira and Verousis, 2017; Voukelatos and Verousis, 2019; Bernales, Verousis and Voukelatos, 2019; Benkraiem, Bouattour, Galariotis and Miloudi, 2021).

An important distinction refers to the presence of fundamental versus non-fundamental herding (see Bikhchandani and Sharma, 2000). Imagine a situation where a group of investors observe the same set of (public or quasi-public) information and there is relative certainty about the quality of this information. It is reasonable to assume that this group of investors will interpret the information in the same way and possibly trade in the same direction. This will give rise to "spurious" or fundamental herding in the sense that investors did not suppress their own beliefs with the intention to follow others. Imagine now the same set of investors receive information that is however difficult to assess in terms of quality. In this scenario, each investor assesses this information independently of other investors and assessments are not shared between investors. Instead, each investor is only able to observe the actions of the other investors. In this situation, "intentional" or non-fundamental herding may arise, indicating that investors mimic each others' actions with intent. Non-fundamental herding is noise-driven and affected by sentiment (see also Galarioties et al., 2015 and Cui et al., 2019). Importantly, non-fundamental herding is fragile (Bikchandani and Sharma, 2000), causes information cascades and leads to inefficiencies (see Çelen and Kariv, 2004).

2.2 Algorithmic Trading, High Frequency Trading, and HFT start

The European Securities and Markets Authority (ESMA) defines AT as follows: "trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price or quantity of the order or how to manage the order after its submission, with limited or no human intervention" (see ESMA MiFID II Review Report, 2021, p.12). There is a general agreement amongst academics that AT improves liquidity and informational efficiency (see Boehmer, Fong and Wu, 2021; Chaboud et al., 2014; Hendershott, Jones and Menkveld, 2011; Chakrabarty and Pascual, 2022). However, AT is also associated with higher adverse selection costs for slow traders (Chaboud et al., 2014) and a reduction in information acquisition (Weller, 2018).

Relatedly, HFT is defined as, "*an algorithmic trading technique characterised by: (a) infrastructure intended to minimise network and other types of latencies, including at least one of the following facilities for algorithmic order entry: co-location, proximity hosting or high-speed direct electronic access; (b) system-determination of order initiation, generation, routing or execution without human intervention for individual trades or orders; and (c) high message intraday rates which constitute orders, quotes or cancellations.*" (ESMA MiFID II Review Report, 2021, p.14). The literature on HFT commonly finds that HFT has a positive impact on liquidity and informational efficiency, while facilitating price discovery (see Hasbrouck and Saar, 2013; Brogaard et al., 2014; Boehmer et al., 2018). However, research work on HFT faces a moving target, with firms engaging in a high-frequency arms race worth an estimated $5 billion annually (Aquilina et al., 2022). Critics of HFT argue that this arms race is "socially wasteful" (Budish, Cramton and Shim, 2015) and "socially excessive" (Biais, Foucault and Moinas, 2015).

In order to examine the effects of HFT on stock exchanges, it is necessary to identify the HFT start date across different exchanges. Unfortunately, there is no established HFT start date corresponding to these various exchanges and therefore, previous literature commonly use colocation service start date in exchanges as the proxy of HFT emergence (Aitken, Cumming, and Zhan, 2017; Boehmer, Fong, and Wu, 2020), while Aitken, Cumming, and Zhan (2015, 2017) also propose HFT effective date as another proxy of HFT emergence.[5] Colocation refers to HFTs and associated firms locating their servers in the same building as the stock exchange servers, providing a higher speed to the flow of time-sensitive information (Brogaard, Hagströmer, and Nordén, 2015). Previous studies employ colocation service announcements as an indicator of AT and HFT activity and examine international differences (Boehmer, Fong, and Wu, 2020), market liquidity (Frino, Mollica, and Webb, 2014), severity of end-of-day price dislocation and average trade size influence (Aitken, Cumming, and Zhan, 2015 and 2017), and HFT firm trading performance (Baron, Brogaard, Hagströmer, and Kirilenko, 2019).

2.3 Hypothesized effects of HFT on herding

There are compelling reasons to believe that HFT activities are likely associated with significant herding. For example, HFT firms pursue several similar strategies, ranging from market making (Menkveld, 2013) to opportunistic strategies, including predatory trading (see Hagströmer and Nordén, 2013; Ye, Yao and Gai, 2013; O'Hara, 2010). Importantly, several studies show that strategies are highly correlated across HFTs, stocks and markets (see

[5] HFT effective date usually appears few years before colocation date. This probably due to HFT established themselves at locations proximate to exchanges earlier before colocation service offered (Aitken, Cumming, and Zhan, 2017).

Chaboud et al., 2014; Boehmer et al., 2018; Menkveld, 2013), therefore giving rise to co-movement in returns and liquidity (see Malceniece et al., 2019; Li, Yin and Zhao, 2020). These co-movements are likely to lead to further stock herding, especially over short-term intervals. Indeed, Jarrow and Protter (2012) describe a mechanism where HFT activity can induce herding. In their setting, HFTs unknowingly coordinate on the basis of a common signal. Their actions, albeit independent, create their own short-term momentum in the market and induce a type of herding behaviour. In addition, the introduction of colocation services would increase competition among HFT market makers and increase intensity of HFT activities (Hendershott, Jones, and Menkveld, 2011). Exchanges with colocation services expect to have more intensive HFT activity, therefore, we expect to find stronger herding evidence from these exchanges. On this basis, our primary hypothesis is the following:

Hypothesis 1: HFT activity induces stock herding.

In particular, if HFT is positively associated with herding, we would expect to see more herding when there is more HFT participation. High frequency traders are known for their ability to provide liquidity to the market due to their rapid trading activities (Brogaard et al., 2018). They use sophisticated algorithms to enter and exit positions, which can enhance market liquidity by reducing bid-ask spreads and increasing trading volume (Hasbrouck, 2007). The literature identifies that HFTs are more likely to trade in large-size stocks (see Brogaard et al., 2018), as these stocks provide greater opportunities for executing their high frequency strategies efficiently (Hendershott and Riordan, 2013). Therefore, when liquidity is abundant, the impact of HFT on price formation becomes more pronounced, potentially leading to stronger herding effects. Also, the literature suggests that HFTs participate more in the market when price volatility is high (see Jarnecic and Snape, 2014). During extreme market conditions characterized by heightened volatility, market participants may become more prone to herd as they react to rapid price movements (Bikhchandani et al., 1992). We therefore propose the following hypothesis:

Hypothesis 2: The association between HFT activity and herding is stronger for large stocks during high liquidity periods.

Further, we have reasons to believe that both fundamental and non-fundamental sources of herding are at play in the market. HFTs exploit the same set of public information, thus following the same set of market signals, which may induce fundamental herding. Given HFTs enhance price discovery also indicates that they trade on fundamental information. On the other hand, HFT trading strategies are highly correlated (see Chaboud et al., 2014; Boehmer et al., 2018) suggesting non-fundamental herding. HFTs aim to close the trading day flat, implying their trading decisions may also be idiosyncratic (Benos and Sagade, 2012). Back-running and predatory trading (see Yang and Zhu, 2019; Van Kervel and Menkveld, 2019) may also lead to an increase in non-fundamental herding. Ultimately, however, we are agnostic regarding the nature of stock herding that is induced by HFT activities. We leave this as an empirical research question and therefore:

Hypothesis 3: HFT activity is associated with both fundamental and non-fundamental herding.

3. Data, variables, and methods

3.1 Data

The utilization of two datasets, one encompassing U.S. stocks and the other - international stocks, is of importance to our study, providing a robust platform to discernibly investigate the relationship between HFT and stock herding across varied market settings. It contributes to a more comprehensive understanding of the role HFT plays in shaping market dynamics such as herding. The U.S. market, with its high HFT penetration and mature infrastructure, provides an essential setting to observe the effects of established HFT practices on stock herding behavior. This domestic sample thus offers a basis for analyzing the potential causal link between HFT activities and stock herding, particularly in a market where HFT accounts for a significant proportion of total trading volume.

Of course, one cannot straightforwardly draw causal inferences from the analysis of the U.S. sample. For example, there may be a latent factor influencing both HFT activity and herding. Seeking to examine causality in more depth, and using an event study approach, this paper uses HFT effective start dates and colocation introduction as start dates of AT and HFT for our international sample. Considering a range of international markets, with varying degrees of HFT adoption, provides a useful comparative dimension to our analysis. The potential variations in HFT induced herding across these markets could illuminate the impact of different market structures, regulatory environments, and trading practices.

We obtain TAQ data for the constituents of the S&P100 from January 2015 to December 2017. This dataset contains information on company name, trade date, millisecond time stamp, types (i.e., trade or quote), trade and quote price, quote size and trade volume. We drop the first and last 10 minutes of the trading day to minimize any overnight and market closing effects. All stocks are sorted at the end of each year according to their market capitalisation and daily data of the Fama-French return factors is collected from Kenneth R. French's Data Library. We average log returns per minute during the trading day.

The sample of international stocks includes the main index constituents for 10 exchanges in nine countries, including Australia, Canada, Germany, Japan, India (Bombay and NSE), Sweden, Switzerland, United Kingdom, and the United States. At the end of each year, we update the constituents for each index and for each exchange. The dataset is collected from Refinitiv Eikon. We follow Aitken et al. (2015) to identify the colocation date and HFT effective date for each exchange.[6] We list the colocation start dates and HFT effective date for each exchange in Table 1. Given that the HFT effective date always precedes the colocation date, we select data three years before the HFT effective date and three years after the colocation date for each exchange (see Figure 1）.

[Insert Table 1 here]

[Insert Figure 1 here]

3.2 Variables

Similarly to Malceniece et al. (2019), Hendershott et al. (2011) and Conrad, Wahal, and Xiang (2015) we use quote and trade message traffic to infer HFT activity.[7] In particular, we estimate *QuoteUpdates* as the average number of quote (price and quantity) changes at best prices per minute and *HFT_trades* as the average negative trading volume (in USD 100) divided by the total number of messages per minute (see Conrad et al., 2015; Hendershott et al., 2011; Boehmer et al., 2020). An increase in *QuoteUpdates* and *HFT_trades* imply an increase in HFT activity. In Figure 2, we plot the time-series of *QuoteUpdates* and *HFT_trades* separately across terciles. In line with previous literature, HFTs are more active in large capitalisation stocks (see Brogaard et al., 2018 and Malceniece et al., 2019).

[6] Several recent studies use colocation dates in order to infer AT/HFT start dates in an international setting. See Boehmer, Fong and Wu, 2021; Frino, Mollica and Webb, 2014; Aitken et al., 2015 and 2017; Baron, Brogaard, Hagströmer and Kirilenko, 2019; Gider et al., 2019.
[7] Conrad et al. (2015) and Boehmer, Fong, and Wu (2020) show that HFTs mainly act through quotes rather than trades.

For the sample of international stocks, we generate two dummy variables: *HFT_start*[8] which equals zero before the HFT start date and one afterwards and *Col_start*[9] which equals zero before the colocation date and one afterwards. As the timing of colocation decisions vary across exchanges, we avoid any potential identification bias in our results. Further, due to the staggered introduction of colocation services, our results are not confounded by a single unrelated event.

3.3 Methods

Methodologically, the empirical herding literature focuses on two domains. The first examines the presence of herding by institutional investors (see Bennett, Sias, and Starks, 2003; Gavriilidis, Kallinterakis, and Ferreira, 2013). On the other hand, most empirical studies use aggregate market data to investigate herding towards the market consensus (see Chang et al., 2000; Chiang and Zheng, 2010; Bohl, Branger and Trede, 2017; Chong, Bany-Ariffin, Matemilola and McGowan, 2020; Duygun, Tunaru and Vioto, 2021). In this paper, we follow the second empirical approach.

Importantly, Christie and Huang (1995) and Chang et al. (2000) demonstrated that in the absence of herding, the relationship between cross-sectional return dispersion and absolute market returns is expected to be linear and positive. If herding is present, the relationship between cross-sectional return dispersion and absolute market returns is expected to be non-linear i.e., cross-sectional return dispersion will decrease for large absolute values of market return. As a consequence, Chang et al. (2000) proposed the following specification to detect herding:

[8] Unlike colocation dates, the *HFT_start* is from Aitken et al. (2015) using trade size and cancellations to identify HFT effective start date to refer the impact of HFT on the marketplace. It is expected as high frequency traders establish their servers close to exchanges before the introduction of colocation services.
[9] *Col_start* refers to the actual introduction of colocation services from exchanges, which is a commitment to offer low-latency infrastructure to attract higher intensity of AT/HFT (Boehmer et al., 2020).

$$CSAD_t = \beta_0 + \beta_1 |R_{m,t}| + \beta_2 R_{m,t}^2 + \varepsilon_t \tag{1}$$

with CSAD denoting the Cross-Sectional Absolute Deviation of returns, estimated as follows:

$$CSAD_t = \frac{1}{N} \sum_{i=1}^{N} |R_{i,t} - R_{m,t}| \tag{2}$$

where N is the number of stocks i included in the cross-section at time t, and $R_{i,t}$ and $R_{m,t}$ are the stock return and market return for each day t, respectively.

Under the no-herding scenario (i.e., rational pricing assumptions), β_1 is expected to be positive and β_2 is expected to be statistically insignificant. Positive β_1 represents a linear, positive relationship between absolute market return and the dispersion of individual returns. In other words, if the market return swings are larger (regardless of whether they're positive or negative), the individual stock returns should also disperse more widely. A statistically significant and positive β_1 reflects that greater market movements lead to more individual return dispersion in a fully rational and efficient market. If β_1 is significant and negative, it could potentially suggest anti-herding behavior, as it would mean that the dispersion of individual returns decreases when the market return increases in absolute terms.

β_2 captures any non-linear effects. Specifically, β_2 is negative and significant if the null of no-herding is rejected, suggesting that the dispersion of individual returns decreases when the market return is at extremely high or low levels. This implies that in times of extreme market conditions, individual returns tend to cluster around the market return, indicating herding behavior. However, if the value of β_2 is statistically significantly positive, this indicates the case of excessively high cross-sectional return dispersion and implies 'anti-herding' (Christie and Huang, 1995; Chiang and Zheng, 2010; Gebka and Wohar, 2013, Sibande et al., 2021). Unlike market-wide herding, investors in this latter context appear to largely ignore information conveyed by the market-wide price movements. Instead, they focus on the dominant views from subset of market participants in an excessive and exaggerated way.[10]

In order to investigate the hypothesis that HFT activity induces herding, we extend (1) as follows:

$$CSAD_t = \beta_0 + \beta_1 |R_{m,t}| + \beta_2 R_{m,t}^2 + \beta_3 HFT_{activity_t}$$
$$+ \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t \quad (3)$$

where $HFT\_activity$ is a dummy variable that takes the value of one when $QuoteUpdates_t$ (HFT_trades_t) is within the top 10% of its distribution and zero otherwise.[11] If HFT activity induces herding (anti-herding), we would expect β_4 to be statistically significant and negative (positive). β_3 estimates the base influence of HFT activity on the dispersion of individual returns, independent of the state of the market. If β_3 is statistically significant and positive, it suggests a potential tendency towards market fragmentation or individualistic trading behaviors (anti-herding) as a result of HFT activity. Conversely, a significant and negative β_3 implies that heightened HFT activity tends to be associated with a narrower dispersion of individual returns, hinting at a tendency towards herding.[12]

---

[10] This behaviour of market participants in this context is often attributed to localised herding, retreating from the market during market stress, and overconfidence (Gebka and Wohar, 2013).

[11] We conducted a regression analysis that incorporated the interaction term $HFT\_activity_t |R_{m,t}|$, however, the findings indicate that this interaction term did not exhibit statistical significance.

[12] If both β_3 and β_4 are significant and have the same sign, it suggests a consistent influence of HFT activity on herding (or anti-herding) behavior, regardless of market conditions. On the other hand, if β_3 and β_4 are significant but have opposite signs, it suggests that the results can change depending on the market conditions. If β_3 is positive (indicating anti-herding under general market conditions) and β_4 is negative (indicating herding under extreme market movements), it implies that HFT activity encouraging more clustering in individual returns during periods of extreme market movement. If β_3 is negative (indicating herding under general market conditions) and β_4 is positive (indicating anti-herding under extreme market movements), it suggests that HFT activity promoting a higher diversity in individual returns during periods of extreme market movement.

To investigate whether potential HFT herding is noise- or fundamental-driven, we decompose CSAD to its fundamental and non-fundamental components. Specifically, this is carried out by regressing CSAD against the Fama and French three factors:

$$CSAD_t = \beta_0 + \beta_1 \left(r_{mkt,t} - r_{f,t}\right) + \beta_2 HML_t + \beta_3 SMB_t + \beta_4 MOM_t + e_t \qquad (4)$$

where *HML* is the high (book-to-market ratio) minus low value return factor, *SMB* is the small (market capitalization) minus big size return factor, and *MOM* is the monthly momentum factor. From (4), the residual e_t captures the component of CSAD due to non-fundamental factors i.e.

$$CSAD_{NONFUND,t} = e_t \qquad (5)$$

The component of the CSAD due to the fundamental factors is therefore the difference between CSAD in the baseline model and $CSAD_{NONFUND}$:

$$CSAD_{FUND,t} = CSAD_t - CSAD_{NONFUND,t} \qquad (6)$$

We subsequently re-estimate (3) separately for $CSAD_{FUND}$ and $CSAD_{NONFUND}$:

$$CSAD_{FUND,t} = \beta_0 + \beta_1 \left|R_{m,t}\right| + \beta_2 R_{m,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t \qquad (7)$$

$$CSAD_{NOFUND,t} = \beta_0 + \beta_1 |R_{m,t}| + \beta_2 R_{m,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t \quad (8)$$

A statistically significant and negative (positive) β_4 in (7) or (8) would imply that HFT is associated with herding (anti-herding) due to fundamental or non-fundamental information. Finally, employing a sample of international stocks, we test if herding is associated with the colocation start date and therefore with HFT activities in the following manner:

$$CSAD_{i,t} = \beta_0 + \beta_1 |R_{m,i,t}| + \beta_2 R_{m,i,t}^2 + \beta_3 Col_{start_{i,t}} + \beta_4 Col\_start_{i,t} R_{m,i,t}^2 + \varepsilon_{i,t} \quad (9)$$

where $CSAD_{i,t}$ is the cross-sectional absolute deviation of exchange *i* on day *t*, $R_{m,i,t}$ is the market return of exchange *i* on day *t*, $Col\_start_{i,t}$ denotes a colocation dummy variable in exchange *i* which equals zero before the colocation date and switches to one after the colocation date. Notably, Aitken et al. (2015) argue that HFT firms might have physically located themselves next to the exchange in order to obtain time advantages long before colocation was officially offered by exchanges. Therefore, it is possible that the colocation start date does not accurately proxy for the effective HFT date. In separate regressions, we therefore regress CSAD against a proxy of HFT effective start date (*HFT_start*) as follows:

$$CSAD_{i,t} = \beta_0 + \beta_1 |R_{m,i,t}| + \beta_2 R_{m,i,t}^2 + \beta_3 HFT_{start_{i,t}} + \beta_4 HFT\_start_{i,t} R_{m,i,t}^2 + \varepsilon_{i,t} \quad (10)$$

where $HFT\_start_{i,t}$ is a dummy variable which equals zero before HFT effective date and one after that date. We follow the literature (see Aitken, Cumming and Zhan, 2015 and 2023) and control for year and exchange fixed effects in (9) and (10). Additionally, all continuous variables are winsorized at the 1st and 99th percentile.

Table 2 reports descriptive statistics for the U.S. and international sample. For the U.S. sample, (Table 2, Panel A), CSAD ranges from 0.0034 to 0.0304. The average daily CSAD ranges from a mean of 0.0076 in 2015 to 0.0068 in 2017. The average number of quote updates decreases from around 55 per minute in 2015 to around 38 in 2017. The average figure for *HFT_trades* also decreases from -43 in 2015 to -50 in 2017. In Table 2, Panel B, we present the descriptive statistics for the international sample covering the sample period from 3 years before the HFT effective start date to 3 years after the colocation start date (see Figure 1). The average daily CSAD ranges (Panel B) from a 0.0198 to 0.013 across the 10 exchanges in the sample.

[Insert Table 2 here]

4. Analysis of HFT's impact on stock herding

4.1 Do HFT activities induce herding?

In Table 3, Column 1, we report the results from (1). As expected, β_1 is positive and significant at the 1% level. Also, β_2 is positive and insignificant. This result is consistent with earlier studies that there is no herding in the U.S. market (see Christie and Huang, 1995; Chang et al., 2000; Gleason, Mathur and Peterson, 2004; Chiang and Zheng, 2010). In Table 3, Columns 2 and 3, we test Hypothesis 1 that HFT activities induce herding and report the results from (3). When we proxy HFT activity using the number of quote updates (*QuoteUpdates*), the coefficient for β_4 is negative and significant at 10% level (-5.868 with a t-statistic of -1.83) indicating that HFT activities induce herding under extreme market conditions. Moreover, stronger evidence is garnered when we use *HFT_trades* i.e., β_4 is -6.595 with a t-statistic of -3.69 and significant at 1% level. These provide evidence to support Hypothesis 1. The statistically significant and positive β_3 in Columns 2 and 3 suggests that during times of high HFT activity (when *QuoteUpdates* and *HFT_trades* equals one), the cross-sectional absolute deviation (CSAD) of returns tends to be higher than during periods of lower HFT activity (when HFT proxies equals zero). This could be interpreted as HFT activity enhancing anti-herding behavior in normal market conditions. These results might suggest that the impact of HFT on market behavior is complex and context-dependent, with HFT possibly contributing to market stability under normal conditions but amplifying collective responses in more extreme situations.

Interestingly, in the estimation of (3), β_2 is shown positive and significant (7.764 with a t-statistic of 1.74 in Column 2 and 6.862 with a t-statistic of 3.25 in Column 3). A positive value for β_2 indicates that cross-sectional return dispersion during period of large market movements is higher, not lower, than conventional asset pricing models predictions (see Gebka and Wohar, 2013 and Christie and Huang, 1995). In this case, market participants largely ignore information that reflect the market consensus. Given that in Table 3, we have essentially decomposed herding into two categories (i.e., herding induced by more active HFT, and herding induced by less active HFT), the results indicate more active HFT induced herding (statistically significant and negative β_4), suggesting high frequency traders tend to align their trading decisions with the broader market consensus. Instead, less active HFT inspired anti-herding (statistically significant and positive β_2) in the market, implies market participants rely more on others or localized sources of information.

[Insert Table 3 here]

Next, we investigate if herding is conditional on firm size given HFTs are more likely to participate in trading large stocks (see Brogaard et al., 2018). To operationalise, for each sample year, we select all stocks that belong to the top 30% according to their market capitalisation at the end of that calendar year. We then re-estimate (1) and (3) for the sample of large stocks only, with the results presented in Table 4.

Table 4, Column 1, provides the results from (1). The statistically significant and positive β_1 (-0.275 with a t-statistic of -1.75) and positive and significant coefficient for β_2 (28.931 with a t-statistic of 3.09) indicates anti-herding effects for large stocks. We then decompose trading activity to more active HFT and less active HFT (Table 4, $\beta_4\ and\ \beta_2$ from Columns 2 and 3). The results indicate that, in line with the results in Table 3, more HFT activities induce herding for larger market-cap stocks under extreme market conditions but induce anti-herding with less active HFT. Indeed, we find statistically significant and positive β_3 in Columns 2 and 3. β_4 in Column 2 and Column 3 is negative and significant at 1% level (-28.869 with a t-statistic of -6.29 in Column 2 and -21.945 with a t-statistic of -4.31 in Column 3). Also, our results indicate the presence of stronger herding effects amongst the sample of large-cap stocks compared to all stocks, which provide evidence to support the hypothesis two. When considering *QuoteUpdates* (*HFT_trades*) as the HFT activity measure, β_4 is 5 (3.3) times larger for large-cap stocks than for all stocks.

[Insert Table 4 here]

Given the body of empirical evidence indicating HFT activity intensifies during periods of heightened price volatility (Jarnecic and Snape, 2014), we dissect how variations in volatility might moderate the relationship between HFT and herding. If HFTs induce herding, then we would expect herding to be stronger in period of high volatility than in periods of low volatility. Below, we investigate this hypothesis. To this end, we employ the VIX index to split our sample to low and high volatility days. We define a day as a low (high) volatility day when

VIX is in the bottom (top) 25% of its distribution. We then estimate (1) and (3) again, separately for low-volatility and high-volatility days. We report the results in Table 5.

In Table 5, Column 1, we report the baseline results for low volatility days where there is no evidence of herding (β_2 is insignificant). In Column 2 and Column 3, we report the results from (3). Unlike earlier results, there is no evidence of HFT herding (the β_4s are insignificant). But there is some limited anti-herding evidence with less active HFT, when the number of quotes (*QuoteUpdates*) is intensive (β_2 is 28.052 with a t-statistic of 2.02 in Column 2). In Table 5, Column 4, we report the results for the high-volatility days. The baseline results again indicate no herding (β_2 is insignificant). However, when we decompose trading activity to more active HFT and less active HFT, the results strikingly indicate that more HFT activities induce herding during high volatility days (β_4 is negative and significant in both Columns 5 and 6). Also, relative to the results in Table 3, the association between HFT activities and herding is stronger during periods of higher volatility.

[Insert Table 5 here]

The results suggest that HFT traders contribute to market herding during high-volatility periods. This indicates that in turbulent times, when market stability could already be at risk, the presence of more active HFT may further intensify market swings by promoting herd-like trading behavior. This could potentially lead to more pronounced market fluctuations and increase systemic risk. In low volatility conditions, less active HFT shows some degree of anti-herding behavior in the market when the number of quotes is high. This suggests that in less turbulent market conditions, market participants may take a contrarian position to the market consensus, potentially aiding in the process of price discovery and adding to market stability.

4.2 Analysis of the mechanism and channel by which HFT induces herding

Moving on, we aim to illuminate the various means by which HFT induces herding and anti-herding behavior within financial markets. This exploration is motivated by the revelation that evidence of herding is more pronounced among large-cap stocks, thereby hinting at a complex interplay between HFT and market behavior in this segment. Our exploration is two-fold. Firstly, we delve into the mechanism that underpin HFT's influence - liquidity. We examine the interplay between liquidity conditions, a market aspect that HFT notably impacts, and herding behavior (Galariotis, Krokida and Spyrou, 2016). Secondly, we turn our attention to the channel through which HFT's influence is transmitted - return comovements (Malceniece et al., 2019). We investigate how HFT impacts return comovements among large-cap stocks and how these comovements subsequently influence herding and anti-herding behavior. By investigating these mechanism and channels, we seek to provide a more textured understanding of how HFT shapes investor behavior, particularly within the large-cap stock arena, across diverse market conditions.

Liquidity is an integral factor that governs trading decisions and market behavior, and its relationship with HFT is well-documented (e.g., Hendershott, Jones and Menkveld, 2011, Brogaard and Garriott, 2019). As such, understanding how liquidity conditions interact with herding behavior under the influence of HFT is a logical next step in our investigation. This will not only broaden our understanding of HFT's impact on market dynamics, but also provide a more comprehensive view of the mechanism through which HFT shapes investor behavior. Liquidity has many dimensions and many different measures, with each liquidity measure designed to capture a different aspect of liquidity. To investigate liquidity's role, we employ a modified version of the Amihud (2002) measure - a widely recognized tool to assess illiquidity taking into account price response elasticity to trading volume. We adapt this measure in line with Karolyi et al. (2012), allowing it to measure liquidity rather than illiquidity. This tailored variable enables us to explore how price responses interact with trading volumes, effectively reflecting liquidity conditions.[13]

In order to further test Hypothesis 2, Table 6 presents our analysis of the role liquidity plays in herding behavior. In Column 1, we present the baseline results for low liquidity days, which show evidence of anti-herding (β_1 is negative and β_2 is positive, both statistically significant). In Columns 2 and 3, we extend the analysis to include HFT activity. Similar to results in Tables 3 and 4, evidence suggests more active HFT induces herding (β_4 is negative and significant). Furthermore, we observe anti-herding when HFT is less active, as seen in the positive and significant β_2 in both Columns 2 and 3. In contrast, for high liquidity days (Column 4), baseline results do not indicate herding (β_2 is insignificant). However, a more detailed picture unfolds when we distinguish between the effects of when HFT are particularly active or not particularly active. We find that high liquidity days are associated with HFT-induced herding (β_4 is negative and significant in both Columns 5 and 6). In fact, the link between HFT activity and herding intensifies during high liquidity periods, being about twice as strong compared to low liquidity days. Moreover, the consistently positive and significant β_3s indicate that in normal market conditions, more active HFT tends to enhance anti-herding in the market.

[Insert Table 6 here]

It is interesting to note the distinct behavior between more active HFT and less active HFT under varying liquidity conditions. Firstly, these observations suggest that more active HFT, potentially due to the capability to process information and execute trades quickly, may be better equipped to capitalize on opportunities arising in high liquidity conditions and thus

---

[13] We first measure liquidity as $Liq_{i,s} = -\log\left(1 + \frac{|R_{i,s}|}{P_{i,s}Vol_{i,s}}\right)$, where $R_{i,s}$, $P_{i,s}$, and $Vol_{i,s}$ is the return, price, and trading volume for stock *i* on trade *s*. We then calculate the average liquidity ($Liq_{i,t}$) for each large stock *i* on day *t*. The market liquidity is measured across stocks as $Liq_{m,t} = \frac{1}{N}\sum_1^N Liq_{i,t}$.

tend to herd more during these periods. Secondly, the results show a dynamic interaction between HFT, liquidity, and herding, with implications for market stability and efficiency. Periods of high liquidity, associated with increased HFT-induced herding, could potentially exacerbate price deviations from fundamental values and contribute to the creation of speculative bubbles. On the other hand, in low liquidity conditions, the anti-herding behavior inspired by less active HFT might act as a stabilizing force.

Building upon Sarr and Lybek's (2002) assertion, we propose that the percentage spread is a more fitting measure for cross-stock comparisons. This is primarily due to its ability to account for the fact that higher priced stocks tend to have less costly spreads. Therefore, to enrich our analysis and validate our previous findings, we incorporated a direct liquidity measure (the quoted bid-ask spread) as an additional robustness check.[14]

[Insert Table 7 here]

The subsequent results, presented in Table 7, adhere to the same sample period as Table 6. The table reveals an anti-herding pattern in the baseline regression under extreme liquidity conditions. Evidence of HFT-related herding are qualitatively similar to the results reported with the Amihud measure during higher liquidity days (β_4 is significantly negative in both Columns 5 and 6). A notable divergence finding from table 6 is the detected anti-herding evidence in high liquidity stocks from the baseline regression in Column 4 (statistically significant and positive β_2). In contrast to the results presented in Table 6, Table 7 did not reveal any evidence of HFT-related herding during periods of lower liquidity. This suggests

[14] The percentage bid-ask spread for large stock i on quote q is calculated as $\%spread_{i,q} = \frac{ask_{i,q} - bid_{i,q}}{(\frac{ask_{i,q} - bid_{i,q}}{2})}$. We estimate average spread for each large stock i on day t as the average spread across all quotes. The market spread on day t is measured across stocks as $\%spread_{m,t} = \frac{1}{N}\sum_{1}^{N} \%spread_{i,t}$.

additional layers of complexity in the relationship between liquidity conditions and herding behaviour, influenced by HFT activities, which worth further exploration.[15]

In the context of our analysis on the channels of HFT, we employ return comovement as a key measure and differentiate HFT activity through two proxies. When we take into account number of quotes (*QuoteUpdates*) as a proxy of HFT activity, we find evidence of HFT-induced herding during periods of lower return comovement (β_4 is -13.664 with a t-statistic of -2.16). This suggests that when there is intensive activity in quote updates, this HFT activity leads to herding in market periods characterized by lower return comovements. This could indicate that in situations where the best prices are rapidly changing, HFTs might play a more influential role, especially when market prices are not moving very closely together. On the other hand, when considering *HFT_trades* as the proxy for HFT activity, we find evidence of HFT-induced herding in periods of higher return comovements (β_4 is -8.119 with a t-statistic of -3.37). This result implies that when trading volumes increase relative to the number of messages, HFTs could contribute to herding in periods where market prices tend to move more closely together. Interestingly, no evidence of herding is observed for HFT trades (*HFT_trades*) during periods of lower return comovements, nor for number of quote updates (*QuoteUpdates*) under periods of higher return comovements. This potentially underscores the varied roles and impacts of HFT activities under different market conditions.

[Insert Table 8 here]

The implications of these findings are twofold. First, they underscore the complexity role of HFT activities in influencing market dynamics under various market conditions, further emphasizing the need for differentiated and context-aware regulatory policies. Secondly, they

[15] The observed differences might be due to the unique aspects of the two different liquidity measures or could be suggestive of other underlying factors impacting these relationships. These differences are valuable as they can lead to further investigations to refine and expand our understanding of these dynamics.

highlight the heterogeneity within HFT activities, with each HFT proxy (*QuoteUpdates* and *HFT_trades*) displaying unique patterns of association with herding under different market conditions.

4.3 Is HFT related to fundamental or non-fundamental herding?

Below we investigate hypothesis three whether herding induced by HFTs is intentional (i.e., non-fundamental) or spurious (i.e., fundamental). In particular, we decompose $CSAD_t$ to $CSAD_{NONFUND,t}$ and $CSAD_{FUND,t}$ as described earlier in subsection 3.3, estimating (7) and (8). We report the results in Table 9. When observing the baseline results for $CSAD_{NONFUND,t}$ (Table 9, Column 1), β_2 is positive and insignificant, indicating the absence of non-fundamental herding. However, when we decompose trading activity to more active HFT and less active HFT (β_4 *and* β_2 from Columns 2 and 3), there is a clear indication that more HFT activities are associated with non-fundamental herding (β_4 is -6.18 with a t-statistic of -1.93 in Column 2 and -6.741 with a t-statistic of -3.55 in Column 3). Relatedly, less active HFT is associated with anti-herding (β_2 is positive and highly significant at 1% level). Equally, when looking at fundamental herding (Column 4), β_2 is insignificant. However, there is no herding evidence that HFT activities are associated with fundamental herding (β_4 is insignificant in Columns 5 and 6).

[Insert Table 9 here]

In order to investigate this further, in Table 10, we focus on large stocks. When looking at $CSAD_{NONFUND,t}$, the baseline results indicate anti-herding (β_2 is 27.616 with t-statistic of 3.01 in Column 1). We further decompose trading activity to more active HFT and less active HFT in Columns 2 and 3. More active HFT is clearly associated with non-fundamental driven herding for large stocks herding (β_4 is negative and highly significant at 1% level for both *QuoteUpdates* and *HFT_trades*). The size of β_4 also indicate that for large stocks HFT activities are strongly associated with intentional (i.e. non-fundamental), where β_4 of *QuoteUpdates* (*HFT_trades*) is 4.7 (3.3) times larger for large cap stocks than for all stocks. In contrast, when looking at $CSAD_{FUND,t}$, β_4 is insignificant in Columns 5 and 6 and indicates no herding evidence that HFT activities induce fundamental-driven herding for large stocks.

Overall, the results in Tables 9 and 10 provide evidence to hypothesis three and indicate that HFT activities are strongly associated with non-fundamental herding, especially for large stocks. This finding is consistent with the view that HFT trading strategies are highly correlated (see Chaboud et al., 2014; Boehmer et al., 2018) and that market making strategies account for the majority of HFT volume (see Hagströmer and Nordén, 2013; Menkveld, 2013). Relatedly, Malceniece et al. (2019) show that whilst HFTs are faster at incorporating market-wide information, two-thirds of the increase in commonality in returns is associated with correlated trading strategies of HFTs. Indeed, our results show that HFTs engage in non-fundamental, highly correlated trading, that is induced by the HFT arms race.

[Insert Table 10 here]

4.3 Evidence on the International Herding

Finally, in this section, we use the colocation dates as an exogenous shock in HFT activities in order to investigate the effect of HFT trading on herding. To this end, we estimate (9) and (10) using our sample of international stocks from 10 exchanges in nine countries (see Table 1). All regressions include year and exchange fixed effects.

In Table 11, we compare the estimation results from (1) and (10). In Column 1, we report the baseline herding equation and β_2 is found positive and significant at the 1% level (1.529 with a t-statistic of 5.13). In Column 2, we use the HFT effective date dummy to proxy for the HFT start date. β_4 is negative and significant at 1% (-8.084 with a t-statistic of -7.14), indicating that HFT is strongly associated with herding after the HFT start date. This finding is consistent with the view that the HFT effective date results in intensive HFT activity (see Aitken et al., 2015 and 2017). On the other hand, non-HFT activities seem to explain the negative stock herding coefficient picked up in the baseline equation. We investigate these findings further in Table 12.

In Table 12, we employ the colocation date as a proxy for the HFT start date (see Table 1) and estimate (1) and (9).[16] In line with the results from Table 11, the results in Table 12 indicate that HFT activities induce herding (β_4 is negative and significant, -1.32 with a t-statistic of -3.15). Importantly, β_4 in Table 11 is 6.12 times larger than in Table 12 (-8.084 compared to -1.32), implying that the HFT effective date has a stronger power in explaining the presence of herding than the colocation start date. This finding supports the work by Aitken et al. (2015) that colocation services are the result of HFT provision.

[Insert Table 11 here]

[Insert Table 12 here]

Lastly, in Table 13, we report the estimation results for (1), (9) and (10), separately for "up days" and "down days" (see Galariotis et al., 2015). "Down days" refers to days when market return is in bottom 5% of its distribution and "up days" to days when market return is at the top 5% of its distribution. In Table 13, Column 2 and 3, we decompose herding to HFT and non-HFT. The results for the "up days" indicate that the HFT start date as well as the colocation start date induce HFT herding. As expected, the coefficient for HFT_start is larger than for Col_start (-8.901 vs -8.621). In Table 13, Panel B, we report the results for "down days". HFT_start is negative but insignificant and Col_start is positive and significant.

[16] The baseline results in Table 11 are little different from the baseline results in Table 10. As Figure 2 and Table 1 show, the HFT effective dates are around one year to four years earlier than the Colocation start dates. It also covers the period of post-colocation start date. Thus, the baseline results without considering the effect of HFT are not much different.

[Insert Table 13 here]

5. Conclusion

In this paper, we examine the role of HFT in herding. To this end, we employ TAQ data for the U.S. market and daily price data for 10 exchanges in nine countries. We posit that HFT activities induce stock herding and add to the literature on how market microstructure affects market efficiency (see Easley, López de Prado, O'Hara and Zhang, 2021 and Hirschey, 2021). Our starting point is that HFT activities are strongly correlated and increase commonality in liquidity (Malceniece et al., 2019), indicating that HFTs, either knowingly or unknowingly, induce stock herding. Indeed, we show consistent results for the U.S. market and for 10 exchanges around the world that HFT activities are positively associated with herding and that such herding is stronger for larger stocks and during more volatile periods. Notably, our results suggest that HFT activities are strongly associated with non-fundamental herding, and anti-herding, pointing towards the role of HFT in fostering information cascades and resultant price inefficiencies. This indicates that the positive effects on price discovery and liquidity induced by HFT are, at least, partially offset by an increase in price inefficiencies.

These findings significantly contribute to the literature on HFT and behavioral finance, while offering valuable insights for both policymakers and investors. Budish et al. (2015) argues that the high-frequency arms race for speed is "socially wasteful" and Aquilina et al. (2022) estimate that the race is worth $5 billion annually. It is not surprising therefore that HFT activities have been a concern to policy makers. Considering the revenue generation potential of HFTs for exchanges worldwide, the potential downside of increased price inefficiencies and market fragility brought by HFT-induced herding calls for considered policymaking. For investors, understanding the existence of anti-herding could illuminate the necessity for diversification strategies to mitigate systematic risk. Future research should continue to investigate these dynamics to ensure market stability and efficiency.

References

Aitken, M., Cumming, D. and Zhan, F., 2015. High frequency trading and end-of-day price dislocation. *Journal of Banking & Finance*, 59, pp.330-349.

Aitken, M., Cumming, D. and Zhan, F., 2017. Trade size, high-frequency trading, and colocation around the world. *The European Journal of Finance*, 23(7-9), pp.781-801.

Aitken, M., Cumming, D. and Zhan, F., 2023. Algorithmic Trading and Market Quality: International Evidence of the Impact of Errors in Colocation Dates. *Journal of Banking and Finance, 151, 106843.*

Andrikopoulos, P., Kallinterakis, V., Leite Ferreira, M. and Verousis, T., 2017. Intraday herding on a cross-border exchange. *International Review of Financial Analysis*, 53, pp.25-36.

Aquilina, M., Budish, E. and O'Neill, P., 2022. Quantifying the High-Frequency Trading "Arms Race." *The Quarterly Journal of Economics*, 137(1), pp.493-564.

Banerjee, A., 1992. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3), pp.797-817.

Baron, M., Brogaard, J., Hagströmer, B. and Kirilenko, A., 2019. Risk and Return in High-Frequency Trading. *Journal of Financial and Quantitative Analysis*, 54(3), pp.993-1024.

Benkraiem, R., Bouattour, M., Galariotis, E. and Miloudi, A., 2021. Do investors in SMEs herd? Evidence from French and UK equity markets. *Small Business Economics*, 56(4), pp.1619-1637.

Bennett, J. R., Sias, R. and Starks, L., 2003. Greener pastures and the impact of dynamic institutional preferences. *Review of Financial Studies*, 16, pp.1203–1238.

Benos, E. and Sagade, S., 2012. High-Frequency Trading Behaviour and Its Impact on Market Quality: Evidence from the UK Equity Market. *SSRN Electronic Journal*.

Bernales, A., Verousis, T. and Voukelatos, N., 2020. Do investors follow the herd in option markets?. *Journal of Banking & Finance*, 119, p.104899.

Biais, B., Foucault, T. and Moinas, S., 2015. Equilibrium fast trading. *Journal of Financial Economics*, 116(2), pp.292-313.

Bikhchandani, S. and Sharma, S., 2000. Herd behavior in financial markets. *IMF Staff papers*, 47(3), pp.279-310.

Bikhchandani, S., Hirshleifer, D. and Welch, I., 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5), pp.992-1026.

Boehmer, E., Fong, K. and Wu, J., 2020. Algorithmic Trading and Market Quality: International Evidence. *Journal of Financial and Quantitative Analysis*, pp.1-30.

Boehmer, E., Li, D. and Saar, G., 2018. The Competitive Landscape of High-Frequency Trading Firms. *The Review of Financial Studies*, 31(6), pp.2227-2276.

Bohl, M., Branger, N. and Trede, M., 2017. The case for herding is stronger than you think. *Journal of Banking & Finance*, 85, pp.30-40.

Breckenfelder, J., 2020. Competition Among High-Frequency Traders, and Market Quality. *VoxEU*, org, 17.

Brogaard, J. and Garriott, C. (2019) ‘High-frequency trading competition’, *Journal of Financial and Quantitative Analysis*, 54(4), pp. 1469–1497.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A. and Sokolov, K., 2018. High frequency trading and extreme price movements. *Journal of Financial Economics*, 128(2), pp.253-265.

Brogaard, J., Hendershott, T. and Riordan, R., 2014. High Frequency Trading and Price Discovery. *The Review of Financial Studies*, 27, pp.2267-2306.

Budish, E., Cramton, P. and Shim, J., 2015. The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response*. *The Quarterly Journal of Economics*, 130(4), pp.1547-1621.

Çelen, B. and Kariv, S., 2004. Distinguishing Informational Cascades from Herd Behavior in the Laboratory. *American Economic Review*, 94(3), pp.484-498.

Chaboud, A., Chiquoine, B., Hjalmarsson, E. and Vega, C., 2014. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 69(5), pp.2045-2084.

Chakrabarty, B. and Pascual, R., 2022. Stock liquidity and algorithmic market making during the COVID-19 crisis. *Journal of Banking & Finance*, p.106415.

Chang, E.C., Cheng, J.W. and Khorana, A., 2000. An examination of herd behavior in equity markets: an international perspective. *Journal of Banking & Finance,* 24 (10), pp.1651–1679.

Chiang, T. and Zheng, D., 2010. An empirical analysis of herd behaviour in global stock markets? *Journal of Banking & Finance,* 34, pp.1911–1921.

Chong, O., Bany- Ariffin, A., Matemilola, B. and McGowan, C., 2020. Can China's cross-sectional dispersion of stock returns influence the herding behaviour of traders in other local markets and China’s trading partners?. *Journal of International Financial Markets, Institutions and Money*, 65, p.101168.

Christie, W. G. and Huang, R. D., 1995. Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal,* 51, pp.31–37.

Conrad, J., Wahal, S. and Xiang, J., 2015. High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, 116(2), pp.271-291.

Cui, Y., Gebka, B. and Kallinterakis, V., 2019. Do closed-end fund investors herd?. *Journal of Banking & Finance*, 105, pp.194-206.

Duygun, M., Tunaru, R. and Vioto, D., 2021. Herding by corporates in the US and the Eurozone through different market conditions. *Journal of International Money and Finance*, 110, p.102311.

Easley, D., López de Prado, M., O’Hara, M. and Zhang, Z., 2021. Microstructure in the Machine Age. *The Review of Financial Studies*, 34(7), pp.3316-3363.

Frino, A., Mollica, V. and Webb, R., 2014. The Impact of Co-Location of Securities Exchanges' and Traders' Computer Servers on Market Liquidity. *Journal of Futures Markets*, 34(1), pp.20-33.

Galariotis, E.C., Krokida, S.-I. and Spyrou, S.I. (2016) ‘Herd behavior and equity market liquidity: Evidence from major markets’, *International Review of Financial Analysis*, 48, pp. 140–149.

Galariotis, E.C., Rong, W. and Spyrou, S.I., 2015. Herding on fundamental information: A comparative study. *Journal of Banking & Finance*, 50, pp.589-598.

Gavriilidis, K., Kallinterakis, V. and Ferreira, M.P.L., 2013. Institutional industry herding: intentional or spurious?. *Journal of International Financial Markets, Institutions and Money*, 26, pp.192-214.

Gębka, B. and Wohar, M., 2013. International herding: Does it differ across sectors?. *Journal of International Financial Markets, Institutions and Money*, 23, pp.55-84.

Gider, J., Schmickler, S. and Westheide, C., 2019. High-Frequency Trading and Price Informativeness. *SSRN Electronic Journal*.

Gleason, K.C., Mathur, I. and Peterson, M.A., 2004. Analysis of intraday herding behavior among the sector ETFs. *Journal of Empirical Finance*, 11(5), pp.681-694.

Hagströmer, B. and Nordén, L., 2013. The diversity of high-frequency traders. *Journal of Financial Markets*, 16(4), pp.741-770.

Hasbrouck, J. and Saar, G., 2013. Low-Latency Trading. *Journal of Financial Markets*, 16(4), pp.646-679.

Hendershott, T., Jones, C. and Menkveld, A.J., 2011. Does Algorithmic Trading Improve Liquidity?. *The Journal of Finance*, 66(1), pp.1-33.

Hirschey, N., 2021. Do High-Frequency Traders Anticipate Buying and Selling Pressure?. *Management Science*, 67(6), pp.3321-3345.

Hirshleifer, D. and Hong Teoh, S., 2003. Herd Behaviour and Cascading in Capital Markets: a Review and Synthesis. *European Financial Management*, 9(1), pp.25-66.

Jarnecic, E. and Snape, M. (2014) 'The provision of liquidity by high-frequency participants', *Financial Review*, 49(2), pp. 371–394.

Jarnecic, E. and Snape, M., 2014. The Provision of Liquidity by High-Frequency Participants. *Financial Review*, 49(2), pp.371-394.

Jarrow, R. and Protter, P., 2012. A Dysfunctional Role of High Frequency Trading in Electronic Markets. *Journal of Theoretical and Applied Finance*, 15, pp.219-249.

Jørgensen, K., Skjeltorp, J. and Ødegaard, B., 2018. Throttling hyperactive robots – Order-to-trade ratios at the Oslo Stock Exchange. *Journal of Financial Markets*, 37, pp.1-16.

Kellard, N.M., Millo, Y., Simon, J. and Engel, O., 2017. Close communications: Hedge funds, brokers and the emergence of herding. British Journal of Management, 28, pp.84-101.

Li, M., Yin, X. and Zhao, J., 2020. Does program trading contribute to excess comovement of stock returns?. *Journal of Empirical Finance*, 59, pp.257-277.

Malceniece, L., Malcenieks, K. and Putniņš, T., 2019. High frequency trading and comovement in financial markets. *Journal of Financial Economics*, 134(2), pp.381-399.

Menkveld, A., 2013. High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4), pp.712-740.

MiFID II Review Report., 2021. *European Securities and Markets Authority*, ESMA70-156-4572.

O'Hara, M., 2010. What Is a Quote?. *The Journal of Trading*, 5(2), pp.10-16.

Shleifer, A. and Summers, L., 1990. The Noise Trader Approach to Finance. *Journal of Economic Perspectives*, 4(2), pp.19-33.

Sibande, X. *et al.* (2021) 'Investor sentiment and (anti) herding in the currency market: Evidence from Twitter Feed Data', *Journal of Behavioral Finance*, 24(1), pp. 56–72.

Tedeschi, G., Iori, G. and Gallegati, M., 2012. Herding effects in order driven markets: The rise and fall of gurus. *Journal of Economic Behavior &Organization*, 81(1), pp.82-96.

Van Kervel, V. and Menkveld, A., 2019. High-Frequency Trading around Large Institutional Orders. *The Journal of Finance*, 74(3), pp.1091-1137.

Voukelatos, N. and Verousis, T., 2019. Option-implied information and stock herding. *International Journal of Finance and Economics*, 24(4), pp.1429-1442.

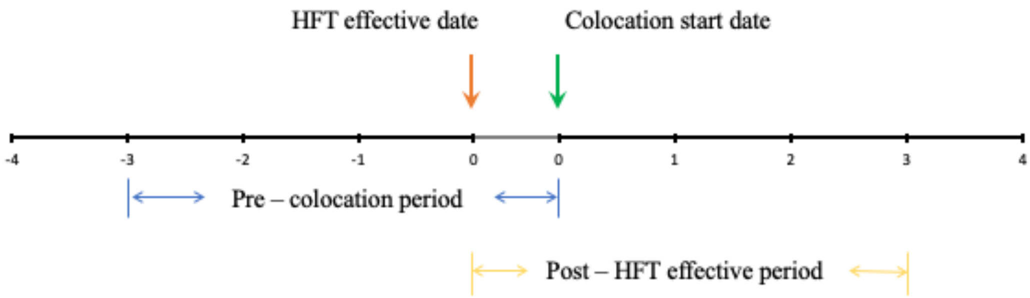
Weller, B., 2018. Does Algorithmic Trading Reduce Information Acquisition?. *The Review of Financial Studies*, 31(6), pp.2184-2226.

Yang, L. and Zhu, H., 2019. Back-Running: Seeking and Hiding Fundamental Information in Order Flows. *The Review of Financial Studies*, 33(4), pp.1484-1533.

Ye, M., Yao, C. and Gai, J., 2013. The externalities of high-frequency trading. *WBS Finance Group Research Paper*, (180).

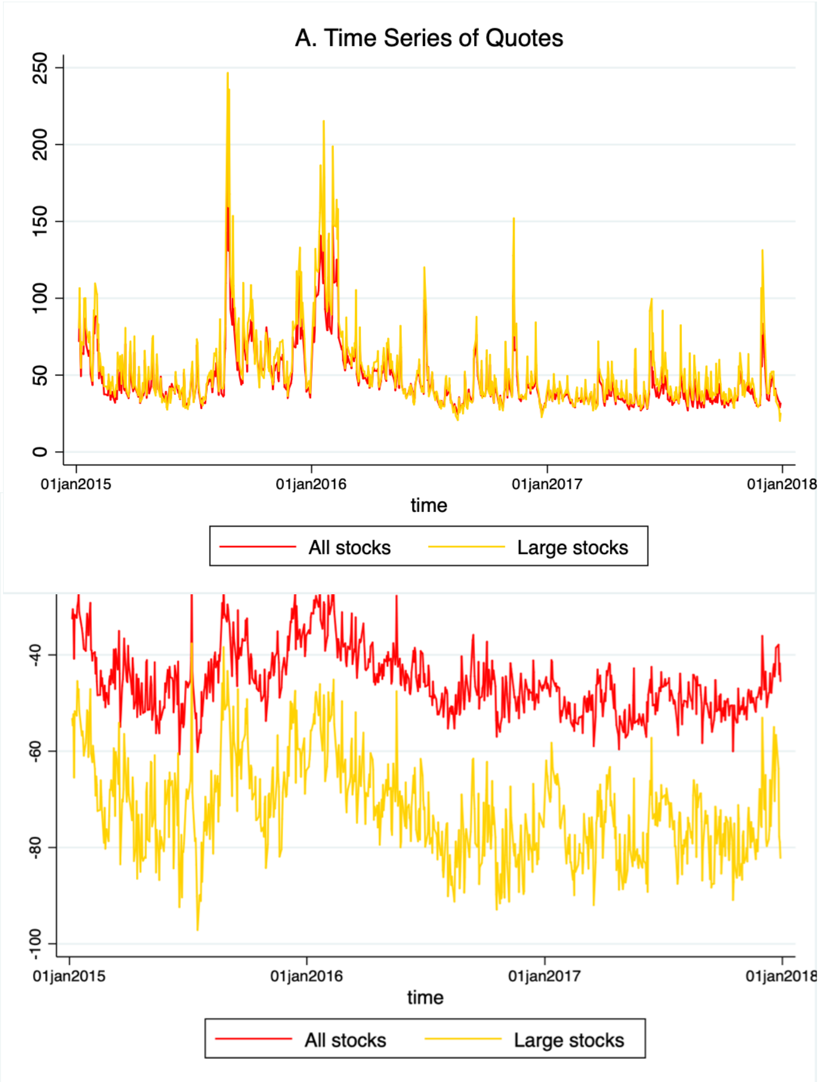
Figures and Tables

Figure 1: Timeline of the Colocation Event



Notes: This figure shows the timeline of the colocation event for each exchange. We have in total 10 exchanges in nine countries as the full sample. The HFT effective date is at least eight months earlier than the colocation start date for each exchange. We collect three years data before the HFT effective date and three years data after the colocation start date for each exchange.

Figure 2: Quote Updates and HFT Trades Measure



Notes: This figure presents the time series of average quote updates and HFT trades measure per minute across days for all stocks and large stocks. The sample covers the period between 2015 and 2017. All stocks include constituents of S&P 100, while large stocks include the top 30 stocks from the full sample. Panel A plots the time series of quote updates per minute, which proxy the HFT quotation and activity. We count all quotes that indicate changes in the price or size of the best quotes for each stock per minute and report equally weighted average of shares across days. Panel B graphs the time series of HFT trades, which defines as the negative of trading volume (in $100) divided by the number of messages. This is a proxy for HFT trades.

Table 1: HFT Effective Date and Colocation Start Date

Country	Exchange Name	HFT Effective Date	Colocation Start Date	Data Period
Australia	Australia stock exchange	Apr2006	Nov2008	01May2003-31Oct2011
Canada	Toronto stock exchange	May2005	Nov2008	01Jun2002-31Oct2011
Germany	XETRA Germany	Jan2003	Q42006	01Feb2000-30Sep2009
India	Bombay stock exchange	May2009	15Nov2010	01Jun2006-14Nov2013
India	NSE India	May2009	Aug2009	01Jun2006-31Jul2012
Japan	Tokyo stock exchange	May2005	May2009	01Jun2002-30Apr2012
United Kingdom	London stock exchange	Feb2006	Sep2008	01Mar2003-31Aug2011
United State	NASDAQ	Jan2003	Apr2005	01Feb2000-31Mar2008
Sweden	Stockholm stock exchange	Apr2005	25Jun2008	01May2002-24Jun2011
Switzerland	Swiss stock exchange	Jan2004	24Jun2008	01Feb2001-23Jun2011

This table includes the list of 10 exchanges and the corresponding country. HFT effective dates are from Aitken et al. (2015), and colocation start dates are from Boehmer et al. (2020).

Table 2: Descriptive Statistics

Panel A: Descriptive statistics for the U.S. market

	$CSAD_t$			$R_{m,t}$		
Year	2015	2016	2017	2015	2016	2017
Minimum	0.0039	0.0034	0.0038	-0.0343	-0.0364	-0.0177
Maximum	0.0304	0.0258	0.0162	0.0293	0.0238	0.0126
Mean	0.0076	0.0081	0.0068	-0.0001	0.0004	0.0007
Std.dev	0.0025	0.0031	0.0019	0.0095	0.0082	0.004
Skewness	3.3303	2.0263	1.6187	-0.3062	-0.5249	-0.3365
Kurtosis	28.535	9.5848	7.4073	4.1171	5.2358	5.7171
No. (# of assets)	100	100	100	100	100	100
Observations	251	251	251	251	251	251
	$QuoteUpdates_t$			$HFTtrades_t$		
Year	2015	2016	2017	2015	2016	2017
Minimum	28.462	23.935	26.981	-60.642	-57.007	-60.039
Maximum	206.966	156.189	83.398	-25.249	-25.913	-36.037
Mean	54.627	52.481	37.788	-42.907	-43.469	-49.681
Std.dev	22.151	24.487	7.7763	7.0643	6.7302	3.9755
Skewness	2.5725	1.9055	2.0589	0.1652	0.6003	0.2528
Kurtosis	14.559	6.5809	9.4326	2.6118	2.826	3.2425
No. (# of assets)	100	100	100	100	100	100
Observations	251	251	251	251	251	251

Panel B: Descriptive statistics for the international sample

Name of exchange	ASX		BSE		LSE		NASDAQ		NSE	
Country	Australia		India		United Kingdom		United State		India	
	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$
Minimum	0	-0.1849	0	-0.1212	0	-0.1024	0	-0.1721	0	-0.145
Maximum	1.2436	0.1219	1.0069	0.1736	1.088	0.1171	1.0072	0.0969	0.8392	0.1914
Mean	0.0151	0.0003	0.0171	0.0001	0.013	0.0003	0.0198	0.0001	0.0164	0.0002
Std.dev	0.0411	0.0178	0.0366	0.0196	0.035	0.016	0.0332	0.0173	0.0368	0.0208
Skewness	26.345	-1.446	23.324	-0.156	24.29	-0.297	20.82	-0.303	19.498	0.054
Kurtosis	744.913	18.113	579.18	9.059	639.8	10.947	515.07	9.523	400.81	10.51
No. (# of assets)	100		100		100		100		50	
Observations	2214		1941		2212		2125		1604	
Name of exchange	NASDAQ Stockholm		SIX Swiss		TYO		TSX		XETRA	
Country	Sweden		Switzerland		Japan		Canada		Germany	
	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$	$CSAD_t$	$R_{m,t}$
Minimum	0	-0.1063	0	-0.0759	0	-0.1025	0	-0.1165	0	-0.637
Maximum	0.0393	0.1283	1.5645	0.0899	0.873	0.1161	0.8493	0.097	1.265	0.7232
Mean	0.013	0.0003	0.0135	0.0002	0.013	0.0001	0.014	0.0004	0.018	0.0001
Std.dev	0.0059	0.0204	0.0413	0.0137	0.029	0.0151	0.0248	0.0151	0.0527	0.0275
Skewness	1.348	0.082	31.286	-0.232	23.56	-0.217	29.796	-0.633	18.893	2.184
Kurtosis	6.278	8.038	1054.9	8.392	603.5	7.952	957.46	11.399	371.13	324.91
No. (# of assets)	60		50		100		60		90	
Observations	1686		2706		2580		2192		2520	

Notes: This table reports descriptive statistics for the US sample (S&P 100) in Panel A, and for the international sample in Panel B. The CSAD is defined as $CSAD_t = \frac{1}{N}\sum_{i=1}^{N} |R_{i,t} - R_{m,t}|$, where $R_{m,t}$ is the market return which is measured as the equally weighted average of individual stock returns ($R_{i,t}$) and is the first difference of the log price series. The cross-sectional absolute deviation (CSAD) is measured for the US market (S&P100) in Panel A for the period between 2015 to 2017, while Panel B reports the results of CSAD for the international sample from 10 exchanges in 9 countries from 3 years before HFT effective start date to 3 years after the colocation start date (see Figure 1). No. (# of assets) reports the number of constituents of the main index for each exchange.

Table 3: Herding for All Stocks

	(1) Baseline	(2) Intensive Quote Updates	(3) Intensive HFT trades
$\lvert R_{m,t} \rvert$	0.107***	-0.006	0.037
	(2.17)	(-0.11)	(0.9)
$R_{m,t}^2$	2.688	7.764*	6.862***
	(1.11)	(1.74)	(3.25)
$HFT\_activity_t$		0.003***	0.002***
		(6.33)	(5.68)
$HFT\_activity_t R_{m,t}^2$		-5.868*	-6.595***
		(-1.83)	(-3.69)
Observations	753	753	753
R^2	0.12	0.23	0.17

Note: This table presents the results of herding on all stocks for the period 2015-2017. Column (1) shows the results of the basic herding specification from regression $CSAD_t = \beta_0 + \beta_1 \lvert R_{m,t} \rvert + \beta_2 R_{m,t}^2 + \varepsilon_t$. Column (2) and (3) reports the results of herding including the effects of HFT activity from regression $CSAD_t = \beta_0 + \beta_1 \lvert R_{m,t} \rvert + \beta_2 R_{m,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t$, where $CSAD_t$ is the cross-sectional absolute deviation on day t, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 4: Herding for Large Stocks

	(1) Baseline	(2) Intensive Quote Updates	(3) Intensive HFT trades
$\lvert R_{m,t} \rvert$	-0.275*	-0.314***	-0.266***
	(-1.75)	(-4.91)	(-2.91)
$R_{m,t}^2$	28.931***	39.320***	33.639***
	(3.09)	(7.41)	(4.88)
$HFT\_activity_t$		0.005***	0.003***
		(6.86)	(5.59)
$HFT\_activity_t R_{m,t}^2$		-28.869***	-21.945***
		(-6.29)	(-4.31)
Observations	753	753	753
R^2	0.4	0.56	0.47

Note: This table shows the herding results for large stocks from 2015 to 2017. Large stocks are 30 stocks with the higher market capitalization. The results of regression $CSAD_t = \beta_0 + \beta_1 \lvert R_{m,t} \rvert + \beta_2 R_{m,t}^2 + \varepsilon_t$ are presented in Column (1). The results of regression $CSAD_t = \beta_0 + \beta_1 \lvert R_{m,t} \rvert + \beta_2 R_{m,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t$ are presented in Column (2) and (3), where $CSAD_t$ is the cross-sectional absolute deviation on day t, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 5: HFT herding conditional on volatility

	Low volatility days			High volatility days		
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Intensive Quote Updates	Intensive HFT trades	Baseline	Intensive Quote Updates	Intensive HFT trades
$\lvert R_{m,t}\rvert$	-0.07	-0.149	-0.085	0.013	-0.083	-0.022
	(-0.48)	(-1.15)	(-0.58)	(0.17)	(-1.16)	(-0.27)
$R_{m,t}^2$	19.142	28.052**	19.876	3.613	8.515***	5.815
	(1.27)	(2.02)	(1.31)	(1.03)	(2.76)	(1.56)
$HFT\_activity_t$		0.002***	-0.0002		0.006***	0.003***
		(4.03)	(-0.49)		(6.74)	(3.94)
$HFT\_activity_t R_{m,t}^2$		-21.703	21.091		-10.101***	-5.989***
		(-1.17)	(1.67)		(-5.1)	(-2.99)
Observations	188	188	188	188	188	188
R^2	0.02	0.21	0.02	0.11	0.39	0.2

Note: This table shows the results of basic herding specification and results of herding including HFT effects on volatile days for all stocks for the period 2015-2017, following the regressions: $CSAD_t = \beta_0 + \beta_1\lvert R_{m,t}\rvert + \beta_2 R_{m,t}^2 + \varepsilon_t$ and $CSAD_t = \beta_0 + \beta_1\lvert R_{m,t}\rvert + \beta_2 R_{m,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t$, where $CSAD_t$ is the cross-sectional absolute deviation on day t, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Column (1) to (3) show the results from sub-sample of low volatile days, which defines as lower 25% of Volatility Index from Chicago Board Options Exchange's (CBOE) for period 2015-2017. Column (4) to (6) report the results for high volatile days from upper 25% of Volatility Index. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 6 Herding under extreme market liquidity of large stocks

	Low Liquid			High Liquid		
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Intensive Quote Updates	Intensive HFT trades	Baseline	Intensive Quote Updates	Intensive HFT trades
$\lvert R_{m,t}\rvert$	-0.449***	-0.409***	-0.429***	0.041	-0.475*	-0.128
	(-4.76)	(-6.99)	(-4.95)	(0.13)	(-1.97)	(-0.33)
$R^2_{m,t}$	37.939***	39.653***	38.232***	15.360	65.622***	34.833
	(8.75)	(27.74)	(10.70)	(0.76)	(3.47)	(1.18)
$HFT\_activity_t$		0.005***	0.003***		0.005***	0.003**
		(4.79)	(5.34)		(3.64)	(2.51)
$HFT\_activity_t R^2_{m,t}$		-23.329***	-17.800***		-51.798***	-30.633*
		(-9.62)	(-4.36)		(-3.78)	(-1.72)
Observations	188	188	188	188	188	188
R^2	0.829	0.896	0.846	0.249	0.664	0.437

Notes: To incorporate the concept of liquidity elasticity, we adopt the methodology introduced by Galariotis et al. (2016) to calculate the Amihud liquidity measure for stocks with large market capitalization as follows: $Liq_{i,s} = -\log\left(1 + \frac{|R_{i,s}|}{P_{i,s}Vol_{i,s}}\right)$, where $R_{i,s}$, $P_{i,s}$, and $Vol_{i,s}$ is the return, price, and trading volume for stock *i* on trade *s*. Within our sample, we identify 30 stocks classified as "large stocks," representing the top 30% of market capitalization for each year in the sample period. To determine the average liquidity ($Liq_{i,t}$) for each individual large stock *i* on day *t*. We conduct the market liquidity across stocks as $Liq_{m,t} = \frac{1}{N}\sum_1^N Liq_{i,t}$. Then we use market liquidity ($Liq_{m,t}$) to define low liquidity (lower 25% of its distribution) and high liquidity (upper 25% of its distribution). The results of regression $CSAD_t = \beta_0 + \beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + \varepsilon_t$ and $CSAD_t = \alpha + \beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R^2_{m,t} + \varepsilon_t$ are reported in this table. In Column (1) and Column (4), the results show baseline herding specification of large stocks during lower liquidity period and higher liquidity period. The findings presented in Column (2) and Column (3) (or Column (5) and Column (6)) demonstrate the presence of herding behavior under different liquidity conditions, induced by high-frequency trading. To account for this, we introduce dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$) in the regression analysis. These dummy variables take a value of one if they fall within the top 10% of their respective distributions, and zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 7 HFT-related herding and liquidity on large stocks – robustness tests

	Low Liquid			High Liquid		
	(1) Baseline	(2) Intensive Quote Updates	(3) Intensive HFT trades	(4) Baseline	(5) Intensive Quote Updates	(6) Intensive HFT trades
$\lvert R_{m,t} \rvert$	-0.790	-0.947*	-0.799	-0.690**	-0.655***	-0.675***
	(-1.41)	(-1.65)	(-1.39)	(-2.57)	(-5.95)	(-4.04)
$R^2_{m,t}$	125.236*	151.975**	125.392*	34.685***	40.659***	38.552***
	(1.79)	(2.00)	(1.76)	(2.95)	(8.06)	(4.80)
$HFT\_activity_t$		0.000	-0.000		0.008***	0.004***
		(0.52)	(-0.99)		(4.82)	(3.98)
$HFT\_activity_t R^2_{m,t}$		-57.463	-6.600		-26.668***	-20.513***
		(-1.63)	(-0.17)		(-7.90)	(-4.75)
Observations	188	188	188	188	188	188
R^2	0.393	0.453	0.395	0.522	0.706	0.625

Notes: The table mirrors the findings presented in Table 6, albeit with a notable change in our approach - we measure liquidity using the percentage bid-ask spread as $\%spread_{i,q} = \frac{ask_{i,q} - bid_{i,q}}{(\frac{ask_{i,q} - bid_{i,q}}{2})}$ for each large stock i on quote q. The spread for stock i on day t measures as the average spread across all quotes. The market spread on day t is measured across large-cap stocks as $\%spread_{m,t} = \frac{1}{N}\sum_1^N \%spread_{i,t}$. See also notes to Table 6.

Table 8 Herding on comovement in returns of large stocks

	Lower Returns Comovement		Higher Returns Comovement	
	(1)	(2)	(3)	(4)
	Intensive Quote Updates	Intensive HFT trades	Intensive Quote Updates	Intensive HFT trades
$\lvert R_{m,t}\rvert$	0.124	0.202*	0.052	0.053
	(1.21)	(1.70)	(0.64)	(0.77)
$R^2_{m,t}$	-0.207	-3.803	4.468	7.574**
	(-0.04)	(-0.61)	(0.92)	(2.40)
$HFT\_activity_t$	0.007***	0.001	0.004***	0.003***
	(5.24)	(1.08)	(4.19)	(2.85)
$HFT\_activity_t R^2_{m,t}$	-13.664**	-0.038	-5.365	-8.119***
	(-2.16)	(-0.01)	(-1.44)	(-3.37)
Observations	188	188	188	188
R^2	0.398	0.062	0.372	0.315

Notes: To generate a time series of return comovement for large stocks, we employ the weighted average of principal components. Initially, we standardize the returns of the 30 large stocks, which are selected based on being in the top 30% of market capitalization as of December 31st in each sample year. Subsequently, we calculate the daily proportion of variance communality using Principal Component Analysis (PCA). This table presents the results from regression $CSAD_t = \alpha + \beta_1 \lvert R_{m,t}\rvert + \beta_2 R^2_{m,t} + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R^2_{m,t} + \varepsilon_t$ for lower 25% return comovement in Column (1) and Column (2), and for higher 25% return comovement in Column (3) and Column (4), where $CSAD_t$ is the Cross-Sectional Absolute Deviation, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 9: Intentional and Spurious Herding for All Stocks

	Non-fundamental driven CSAD ($CSAD_{NONFUND,t}$)			Fundamental driven CSAD ($CSAD_{FUND,t}$)		
	(1) Baseline	(2) Intensive Quote Updates	(3) Intensive HFT trades	(4) Baseline	(5) Intensive Quote Updates	(6) Intensive HFT trades
$\lvert R_{m,t}\rvert$	0.089*	-0.028	0.019	0.018*	0.021	0.017
	(1.72)	(-0.45)	(0.44)	(1.71)	(1.49)	(1.65)
$R_{m,t}^2$	3.256	8.887**	7.533***	-0.567	-1.123	-0.671
	(1.24)	(1.97)	(3.18)	(-1.2)	(-1.31)	(-1.25)
$HFT\_activity_t$		0.003***	0.002***		0.0002	0.0001
		(6.29)	(5.49)		(1.22)	(0.67)
$HFT\_activity_t R_{m,t}^2$		-6.18*	-6.741***		0.318	0.146
		(-1.93)	(-3.55)		(0.53)	(0.31)
Observations	753	753	753	753	753	753
R^2	0.12	0.22	0.17	0.05	0.02	0.01

Note: This table reports results for basic herding specification for all stocks from regression $CSAD_{NONFUND,t} = \beta_0 + \beta_1 \lvert R_{m,t}\rvert + \beta_2 R_{m,t}^2 + \varepsilon_t$ in Column (1) and regression $CSAD_{FUND,t} = \beta_0 + \beta_1 \lvert R_{m,t}\rvert + \beta_2 R_{m,t}^2 + \varepsilon_t$ in Column (4). $CSAD_{NONFUND,t} = e_t$ from regression $CSAD_t = \beta_0 + \beta_1 (R_{mkt,t} - R_f) + \beta_2 HML_t + \beta_3 SMB_t + \beta_4 MOM_t + e_t$; $CSAD_{FUND,t} = CSAD_t - CSAD_{NONFUND,t}$. Column (2), (3), (5), and (6) decompose $CSAD_t$ to deviations due to non-fundamental and fundamental factors and presents results from the regressions: $CSAD_{FUND,t} = \beta_0 + \beta_1 \lvert R_{mkt,t}\rvert + \beta_2 R_{mkt,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t$, and $CSAD_{NONFUND,t} = \beta_0 + \beta_1 \lvert R_{mkt,t}\rvert + \beta_2 R_{mkt,t}^2 + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R_{m,t}^2 + \varepsilon_t$, where $CSAD_t$ is the cross-sectional absolute deviation on day t, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 10: Intentional and Spurious Herding for Large Stocks

	Non-fundamental driven CSAD ($CSAD_{NONFUND,t}$)			Fundamental driven CSAD ($CSAD_{FUND,t}$)		
	(1) Baseline	(2) Intensive Quote Updates	(3) Intensive HFT trades	(4) Baseline	(5) Intensive Quote Updates	(6) Intensive HFT trades
$\lvert R_{m,t} \rvert$	-0.252*	-0.289***	-0.243***	-0.024*	-0.030**	-0.026**
	(-1.65)	(-4.33)	(-2.76)	(-1.93)	(-2.38)	(-2.07)
$R^2_{m,t}$	27.616***	38.083***	32.404***	1.337**	1.401**	1.343**
	(3.01)	(6.80)	(4.73)	(2.53)	(2.42)	(2.51)
$HFT\_activity_t$		0.005***	0.003***		0.000**	0.000
		(6.95)	(5.57)		(2.22)	(0.83)
$HFT\_activity_t R^2_{m,t}$		-29.012***	-22.272***		-0.166	0.086
		(-6.03)	(-4.29)		(-0.24)	(0.10)
Observations	753	753	753	753	753	753
R^2	0.38	0.54	0.46	0.03	0.03	0.02

Note: Column (1) to (3) report results for non-fundamental herding for large stocks. $CSAD_{NONFUND,t}$ is the error term from regression $CSAD_t = \beta_0 + \beta_1(R_{mkt,t} - R_f) + \beta_2 HML_t + \beta_3 SMB_t + \beta_4 MOM_t + e_t$, while $CSAD_{FUND,t} = CSAD_t - CSAD_{NONFUND,t}$. Column (1) and (4) show baseline results from $CSAD_{NONFUND,t}$ and $CSAD_{FUND,t}$, respectively. Column (2) and (3) report results of HFT activity reacting to non-fundamental information from regression $CSAD_{NONFUND,t} = \beta_0 + \beta_1 \lvert R_{mkt,t} \rvert + \beta_2 R^2_{mkt,t} + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R^2_{m,t} + \varepsilon_t$, while results of HFT activity to fundamental information is regressed as $CSAD_{FUND,t} = \beta_0 + \beta_1 \lvert R_{mkt,t} \rvert + \beta_2 R^2_{mkt,t} + \beta_3 HFT\_activity_t + \beta_4 HFT\_activity_t R^2_{m,t} + \varepsilon_t$ in Column (5) and Column (6), where $CSAD_t$ is the cross-sectional absolute deviation on day t, $R_{m,t}$ is the market return, $HFT\_activity_t$ are dummy variables for number of quotes ($QuoteUpdates_t$) and HFT trades ($HFTtrades_t$), which takes value of one on its top 10% distribution and takes value of zero otherwise. Standard errors are robust to heteroscedasticity and autocorrelation. T-statistics in parentheses.

Table 11: Herding for the Post-HFT Effective Start Date

	(1) Baseline	(2) HFT start date
$\|R_{m,t}\|$	0.601***	0.449**
	(3.5)	(3.12)
$R^2_{m,t}$	1.529***	9.88***
	(5.13)	(9.19)
HFT_start_t		0.002**
		(2.95)
$HFT\_start_t R^2_{m,t}$		-8.084***
		(-7.14)
Number of exchanges	10	10
R^2	0.22	0.24
Year FEs	Yes	Yes
Exchange FEs	Yes	Yes

Notes: This table presents the results of herding on 10 exchanges from the full sample. Column (1) shows the results of the basic herding specification for period of the post-HFT effective start date from regression $CSAD_{i,t} = \beta_0 + \beta_1|R_{m,i,t}| + \beta_2 R^2_{m,i,t} + \varepsilon_{i,t}$. Column (2) reports the results of herding including the effect of HFT start date from regression $CSAD_{i,t} = \beta_0 + \beta_1|R_{m,i,t}| + \beta_2 R^2_{m,i,t} + \beta_3 HFT\_start_{i,t} + \beta_4 HFT\_start_{i,t} R^2_{m,i,t} + \varepsilon_{i,t}$, where $CSAD_{i,t}$ is the cross-sectional absolute deviation for exchange i on day t, $R_{m,i,t}$ is the market return, $HFT\_start_{i,t}$ is a dummy variable for HFT effective date which takes value of one after this date on exchange i and takes value of zero before this date. Year fixed effects and exchange fixed effects are included in both regressions. Standard errors are clustered on the exchange level. T-statistics in parentheses.

Table 12: Herding for the Post-Colocation Start Date

	(1)	(2)
	Baseline	Colocation start date
$\lvert R_{m,t}\rvert$	0.509**	0.618**
	(2.64)	(3.17)
$R_{m,t}^2$	1.453***	2.591***
	(5.37)	(4.2)
Col_start_t		-0.001*
		(-1.98)
$Col\_start_t R_{m,t}^2$		-1.32**
		(-3.15)
Number of exchanges	10	10
R^2	0.19	0.21
Year FEs	Yes	Yes
Exchange FEs	Yes	Yes

Notes: This table presents the results of herding on 10 exchanges from the full sample. Column (1) shows the results of the basic herding specification from regression $CSAD_{i,t} = \beta_0 + \beta_1\lvert R_{m,i,t}\rvert + \beta_2 R_{m,i,t}^2 + \varepsilon_{i,t}$ during period of the post-colocation start date. Column (2) reports the results of herding after the introduction of colocation services from exchanges, which follows regression $CSAD_{i,t} = \beta_0 + \beta_1\lvert R_{m,i,t}\rvert + \beta_2 R_{m,i,t}^2 + \beta_3 Col\_start_{i,t} + \beta_4 Col\_start_{i,t} R_{m,i,t}^2 + \varepsilon_{i,t}$, where $CSAD_{i,t}$ is the cross-sectional absolute deviation for exchange i on day t, $R_{m,i,t}$ is the market return, $Col\_start_{i,t}$ is a dummy variable of colocation services which takes value of one after exchange i offers colocation services and takes value of zero otherwise. Year fixed effects and exchange fixed effects are included in both regressions. Standard errors are clustered on the exchange level. T-statistics in parentheses.

Table 13: Herding during Extreme Market Returns

	Panel A: Up days			Panel B: Down days		
	(1) Baseline	(2) HFT start date	(3) Colocation start date	(1) Baseline	(2) HFT start date	(3) Colocation start date
$\lvert R_{m,t}\rvert$	1.686*	0.697	0.301	2.195**	2.037**	1.831**
	(1.95)	(1.39)	(0.69)	(3.06)	(2.58)	(2.81)
$R^2_{m,t}$	-0.131	9.811***	10.045***	-0.355	3.769	0.026
	(-0.11)	(6.48)	(5.32)	(-0.29)	(0.35)	(0.03)
HFT_start_t		0.021***			0.03	
		(5.47)			(0.48)	
$HFT\_start_t R^2_{m,t}$		-8.901***			-3.896	
		(-9.95)			(-0.39)	
Col_start_t			0.017*			-0.027*
			(2.06)			(-2.01)
$Col\_start_t R^2_{m,t}$			-8.621***			3.246*
			(-5.14)			(2.13)
Number of exchanges	10	10	10	10	10	10
R^2	0.24	0.33	0.32	0.28	0.29	0.29
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Exchange FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results of herding during extreme market returns for the full sample, following regressions: $CSAD_{i,t} = \beta_0 + \beta_1\lvert R_{m,i,t}\rvert + \beta_2 R^2_{m,i,t} + \varepsilon_{i,t}$, where $CSAD_{i,t}$ is the cross-sectional absolute deviation for exchange i on day t, $R_{m,i,t}$ is the market return; $CSAD_{i,t} = \beta_0 + \beta_1\lvert R_{m,i,t}\rvert + \beta_2 R^2_{m,i,t} + \beta_3 HFTstart_{i,t} + \beta_4 HFTstart_{i,t} R^2_{m,i,t} + \varepsilon_{i,t}$, where $HFTstart_{i,t}$ is dummy variable for HFT effective date which takes value of one after this date on exchange i and takes value of zero before this date; $CSAD_{i,t} = \beta_0 + \beta_1\lvert R_{m,i,t}\rvert + \beta_2 R^2_{m,i,t} + \beta_3 Col_{i,t} + \beta_4 Col_{i,t} R^2_{m,i,t} + \varepsilon_{i,t}$, where $Col_{i,t}$ is dummy variable for the colocation start date which takes value of one after this date on exchange i and takes value of zero before this date. Panel A shows the results of sub-sample of market return on its upper 5% distribution and Panel B shows the results of sub-sample of market return on its lower 5% distribution. The regressions include year FEs and exchange FEs. *** (**) (*) Significance at the 1% (5%) (10%) two-tailed level. Standard errors are clustered on the exchange level. T-statistics in parentheses.