HOUSE OF FINANCE

WORKING PAPER SERIES

Jasmin Gider – Simon N. M. Schmickler – Christian Westheide

High-Frequency Trading and Price Informativeness

SAFE Working Paper No. 248

SAFE I Sustainable Architecture for Finance in Europe
A cooperation of the Center for Financial Studies and Goethe University Frankfurt

House of Finance | Goethe University
Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main

Tel. +49 69 798 30080 | Fax +49 69 798 33910
info@safe-frankfurt.de | www.safe-frankfurt.de

High-Frequency Trading and Price Informativeness *

Jasmin Gider[†] Simon N. M. Schmickler[‡]
j.gider@uvt.nl snms@princeton.edu

Christian Westheide[§]
christian.westheide@univie.ac.at

July 2019

Abstract

We study how stock price informativeness changes with the presence of high-frequency trading (HFT). Our estimate is based on the staggered start of HFT participation in a panel of international exchanges. With HFT presence market prices are a less reliable predictor of future cash flows and investment, even more so for longer horizons. Further, idiosyncratic volatility decreases, mutual funds trade less actively and their holdings deviate less from the market-capitalization weighted portfolio. These findings suggest that price informativeness declines with HFT presence, consistent with theoretical models of HFTs' ability to anticipate informed order flow, reducing incentives to acquire fundamental information.

JEL classification: G10, G14
Keywords: High-Frequency Trading, Price Efficiency, Information Acquisition, Information Production

---

*We thank Mark Van Achter, Lieven Baele, Itzhak Ben-David, Tobias Berg, Dion Bongaerts, Jonathan Brogaard, Gökhan Cebiroğlu, François Degeorge, Sarah Draus, David Easley, Raphael Flore, Thomas Gehrig, Hendrik Hakenes, Nikolaus Hautsch, Frank de Jong, William Mann, Satchit Sagade, Farzad Saidi, Isabel Schnabel, Jan Schneemeier, Andriy Shkilko, Erik Theissen, Philipp Valta, Brian Weller, Gunther Wuyts, Sergey Zhuk, and seminar participants at the FIRN Market Microstructure Meeting, Belgian Financial Research Forum, FMA Annual Meeting, Annual Financial Market Liquidity Conference, University of Mannheim, University of Vienna, LMU Munich, Tilburg University, University of Bonn, and University of Hohenheim for helpful comments and discussions. Christian Westheide gratefully acknowledges research support from the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE. All remaining errors are our own.

[†]Department of Finance, Tilburg University, PO Box 90163, 5000 LE Tilburg, The Netherlands, +31 13 466 3238.

[‡]Princeton University, Department of Economics, Princeton, New Jersey 08540, USA.

[§]University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, +43 1 4277 37505.

1 Introduction

High-frequency traders (HFTs) have emerged as a new major type of participant in financial markets over the last two decades. On modern equity exchanges, HFTs nowadays account for the majority of order messages and a significant share of trading volume. In the U.S., high-frequency trading (HFT) constitutes approximately half of trading volume, in Europe about one third.[1] HFTs are characterized by short holding periods and a high degree of technological sophistication enabling rapid communication with the exchange server, thus allowing the submission of order messages with low latency.[2]

We investigate how stock price informativeness about fundamentals changes with the beginning of HFT. To that end, we estimate a generalized difference-in-differences model, using an international panel of 18 stock markets and HFT start dates that are based on pronounced increases in order-cancellation ratios and decreases in average trade sizes (see Aitken et al. (2015)). We measure price informativeness using the welfare-based measure suggested by Bai et al. (2016), which captures the variation in future cash flows and investment in the next one to five years that is predicted by current market prices. The staggered introduction across different markets reduces the likelihood that a simultaneous unrelated event drives the results. Price informativeness about future cash flows and investment, and idiosyncratic volatility decrease with the start of HFT, and institutional investors deviate less from the market portfolio. These findings suggest that information acquisition deteriorates with the start of HFT.

[1] See Deutsche Bank Research citing estimates from TABB Group for 2014.

[2] HFT is a subset of algorithmic trading. Algorithmic trading refers to the general class of trading strategies which determine order submissions and cancellations in an automated fashion based on a set of input variables stemming from market data. See Menkveld (2016) for a recent survey of the literature on HFT.

Fundamentally informative prices matter from a social welfare perspective because they lead to an efficient allocation of real resources. Prices that reveal the attractiveness of future investment opportunities enable funds to flow accordingly. Information acquisition also matters for social welfare if the information that market participants acquire feeds into real decision making, e.g., through learning or incentive channels. If market participants acquire information that is not known to decision makers at the firm, then the revelation of this information leads to more efficient investment decisions as conjectured by Hirshleifer (1971), or more recently, the market feedback loop literature (e.g., Dow et al. (2017), or Edmans et al. (2015)).

Informative prices require two conditions: first, existing information needs to be impounded into prices through the trading process. Second, new information has to be acquired by investors to begin with. Most empirical studies on HFT focus on short-horizon efficiency, the former channel, and document a positive effect. These studies examine outcomes such as how closely prices resemble a random walk, or whether HFT trade against transitory pricing errors. Several theoretical models yield predictions on information acquisition, the latter channel. Short-run efficiency might improve, while information acquisition actually deteriorates. Empirically testing this channel is difficult, because information acquisition cannot be directly observed from the researcher's perspective. Our paper contributes by studying price informativeness and thereby providing indirect evidence on information acquisition. Depending on the impact of HFT on information acquisition, the net effect of HFT on price informativeness can be either positive or negative.

Our analysis shows that the start of HFT is associated with a substantial reduction in the informativeness of prices about future cash flows, amounting to at least 50% of one standard deviation for horizons greater than or equal to two years. The economic magnitude of this decline increases further for longer horizons, reaching approximately 71% of one standard deviation for a five year horizon. The forecasting power of prices with respect to investment also decreases by at least one third of a standard deviation, and becomes more pronounced for longer horizons.

Examining the timing of changes in price informativeness around the start of HFT, we find that the wedge between markets where HFT has not (yet) been adopted coincides with the estimated start dates. The findings cannot be explained by differential exposures to the growing importance of exchange-traded funds (ETFs). Further, the findings cannot be explained by changes in the composition of markets, which might directly affect price informativeness. Moreover, we obtain comparable estimates when we account for potential differences in the precision with which the price informativeness measures are estimated at the exchange-year level. Cross-sectional tests reveal that the effect is more pronounced for firms in which HFT are known to be more active (large firms), and for firms that are more difficult to value (young and high growth firms).

Idiosyncratic volatility, the variation in stock returns that cannot be explained by asset pricing factors, measures the incorporation of firm-specific information into prices. We document that idiosyncratic volatility decreases by 15% of one standard deviation subsequent to the start of HFT, thus providing empirical support for the detrimental effect of HFT on information acquisition based on firm-level observations.

Further, we study the behavior of institutional investors in these markets. If institutional investors acquire and process less firm-specific information, we expect this to be reflected in their investment decisions. For each market, we compute the deviation of portfolio holdings of mutual funds from a market capitalization-weighted portfolio

(Active weight) and trades leading to changes in their active positions (Active trade). We find that both measures decrease with the start of HFT, by approximately 39% and 68% of one standard deviation, respectively.

Taken together, these findings lend support to the hypothesis that HFT is detrimental to information acquisition activities. Hence, we provide empirical support for the existence of a tension between the incorporation of existing information in prices[3] and incentives to acquire new information that appears to be aggravated by HFT. Our results help reconcile the opposing views of most of the existing academic literature on HFT and the opinions expressed by some institutional investors who base their investment decisions on fundamental information, and who indeed appear to be the group of market participants who are negatively affected by HFT.

In addition to the above mentioned literature on the real effects of financial markets,[4] this study complements the various strands of the literature on HFT. Our study tests predictions made in a number of recent theoretical studies investigating the effect of HFT on information acquisition by other market participants. Stiglitz (2014) voices the concern that HFTs anticipate informed order flow and appropriate the information rents that would have otherwise accrued to the investors that incurred information acquisition costs. As the rents from investing in fundamental information acquisition decrease, information production by investors decreases accordingly. As a result, less fundamental information is impounded into prices and resource allocation deteriorates, because it is based on less informative market prices. Yang and Zhu (2018) analyze this mechanism formally, by building on a two-period Kyle (1985) model, to which

[3]See, e.g., Foucault et al., 2016, Brogaard et al., 2015, Chakrabarty et al., 2018, Brogaard et al., 2014, Carrion, 2013, Riordan and Storkenmaier, 2012, Conrad et al., 2015, Boehmer et al., 2018, Zhang, 2017, as examples for literature studying this process in relation to HFT.

[4]See Bond et al. (2012) for a comprehensive survey.

they add a so-called "back-runner". Their model analyzes the strategic interaction between two types of informed traders: a trader that is fundamentally informed, and the back-runner that infers this fundamental information from observing past order flow. If this order flow signal is sufficiently accurate, the fundamentally informed trader adds noise to his trading strategies in an attempt to conceal his private information. As a result, less fundamental information is revealed in equilibrium. An extension of the model with endogeneous information acquisition shows that the fundamental trader acquires less information to begin with in the presence of a back-runner. Draus (2018), in a three-period Kyle model, considers an HFT that is either able to learn from fundamental-based order flow or to obtain a noisy signal about the fundamental investor's information irrespective of the order flow. In both cases, the fundamental investor acquires less information than in the absence of the HFT and long-term price informativeness is lower. Baldauf and Mollner (2018) model order anticipation in a fragmented market where HFTs can act both as liquidity demanders and suppliers. They find that if HFTs become faster, both information acquisition and the bid-ask spread decrease. Dugast and Foucault (2018) show that price informativeness can decline if readily available, raw, but imprecise information becomes sufficiently inexpensive such that market participants reduce their demand for more accurate, processed information.

Our analysis is also related to studies on transaction costs of institutional investors. Tong (2015) finds that HFT activities increase transaction costs in the U.S., whereas Brogaard et al. (2014), in a study of the U.K. equity market, do not find any significant effects. Van Kervel and Menkveld (2019) and Korajczyk and Murphy (2019), for the Swedish and Canadian markets, respectively, find that HFTs can apparently identify large institutional orders. Korajczyk and Murphy (2019) also show that HFT increases the costs of large informed trades.

In a contemporaneous paper that, to our knowledge, is the only empirical study investigating the implications of automated trading on information acquisition, Weller (2018) documents that algorithmic trading decreases the amount of information that is impounded into prices in the period prior to quarterly earnings announcements. His evidence supports the existence of a trade-off between the incorporation of existing and new information in prices. Our approach is complementary to that of Weller (2018) because, applying a different methodology, we study a longer-term measure of fundamental price informativeness. Our results suggest that the information that is impounded into prices with a delay far exceeds the content of a quarterly earnings announcement. Based on the reasoning by Hirshleifer (1971), longer-term information is more relevant for allocative efficiency as compared to information which is latent but will be revealed with certainty in the short-run.

The remainder of this paper is organized as follows: Section 2 describes the empirical strategy, the main measures and the data. Section 3 presents the results on the informativeness of prices about cash flows and investment. Section 4 examines more direct measures of information acquisition, and Section 5 concludes.

2 Empirical Strategy and Data

2.1 Empirical Strategy

The main idea behind our empirical strategy is to use the staggered start of HFT presence in international markets to study the effect of HFT on price informativeness. We use the estimated HFT start dates by Aitken et al. (2015) who follow two approaches.

HFT is generally considered to be associated with a large amount of order cancellations relative to trading volume and small trade sizes. Thus, using order book and trade data from Thomson Reuters Tick History (TRTH), Aitken et al. (2015) identify times with a pronounced and persistent increase in order cancellation-to-trade ratios, or a decrease in trade sizes, respectively.[5] Start dates based on order cancellation rates are not available for five markets with HFT. We use a combination of both approaches. We use the earlier of the two start dates, in case they are both available for the given market, and the trade size-based start dates for the markets without information on order cancellation ratios. Table 1 shows the HFT start dates for 13 international exchanges. The start dates based on trade size range from the early adopters (United States, Germany) in the beginning of 2003 to the late adopters in 2009 (Indian stock exchanges). The start dates based on order cancellation are broadly comparable, but occur 16 months earlier for Toronto and 24 months earlier for London.

Five exchanges (Seoul, Shanghai, Shenzhen, Singapore and Hong Kong) serve as counterfactuals in our analyses, because these markets were not exposed to HFT over our sample period. On the mainland Chinese exchanges in Shenzhen and Shanghai, it is prohibited to open and close a position in a security on the same trading day (Bian et al., 2017).[6] HFT in the Hong Kong and Korean equity markets is nearly impossible because of a financial transaction tax that is payable on each transaction even for positions that are closed by the end of the trading day without exemptions. High exchange trading fees have made HFT uneconomical in Singapore.[7]

[5]For the precise definition of the dates, see the appendix of Aitken et al. (2015). TRTH is a database developed by SIRCA, founded by Professor Michael Aitken.

[6]As is the case for exchanges in other countries, rules are different for derivatives markets.

[7]See Meyer and Guernsey (2017), https://www.hkex.com.hk/Services/Rules-and-Forms-and-Fees/Fees/Securities-(Hong-Kong)/Trading/Transaction?sc_lang=en, and http://www.nts.go.kr/eng/data/KOREANTAXATION2012.pdf.

The use of colocation, i.e., the housing of trading firms' computer servers within an exchange's data center, is closely related to HFT activity. We use these dates as a third alternative definition for HFT start dates. While colocation today is used also by other major market participants, HFTs have originally been the primary clientele of exchanges' colocation offerings. It is important to note, however, that colocation is not a necessary condition because HFTs may house their servers in close geographic proximity to exchanges without the latter offering colocation services. In fact, it is likely that exchanges begin to offer colocation as an endogenous response to the demand by HFTs. Colocation does facilitate HFT and likely results in a larger amount of HFT, even though the first HFTs might have traded on an exchange before the initiation of colocation offerings. Aitken et al. (2015) identify the dates when exchanges offered colocation for the first time and show that the start of HFT based on trade size has preceded the offering of colocation services. We recognize that these approaches to estimate the start of HFT in certain markets are noisy. In the Appendix in Tables A4, A5, and A6 we investigate the sensitivity of the results to the choice of the start date.

Based on the HFT start dates, we run a difference-in-differences analysis with multiple events using a panel of exchange-year observations and estimate

$$\mathrm{Y}_{k,m,t} = \beta_0 + \beta_1 \mathrm{HFT}_{m,t} + \delta X_{m,t} + \eta_t + \mu_m + \varepsilon_{m,t}, \tag{1}$$

where m indicates the stock exchange and t the year, Y_k represents price informativeness about future cash flows or investment for the time horizons $k = 1, ..., 5$. HFT is zero prior to the HFT start date and one for all following years. X is a vector of control variables that consists of the natural logarithm of total market size and Electronic, a dummy variable capturing the effect of the transition from floor to electronic trading based on Gorham and Singh (2009). η_t are year fixed effects, μ_m stock exchange fixed effects, and $\varepsilon_{m,t}$ is the error term. Following the same approach, we also analyze changes in idiosyncratic volatility using a panel of firm-year observations.

As indicated above, the models in both the exchange-level and the firm-level analysis include year and exchange or firm fixed effects, respectively. The former flexibly eliminates common trends. The latter eliminates the impact of time-invariant unobservable firm or stock exchange-specific characteristics. Our estimates of the coefficient of HFT are thus driven by variation within markets and within firms.

We argue that HFT adoption has likely been brought about by the presence of sophisticated investors in combination with the automation of trading platforms. Because the former likely start out trading in their home markets, differences between the populations of investors in different countries matter. The latter has been adopted on different exchanges at different points in time. The start of HFT requires certain technical and institutional preconditions: the market has to permit direct market access, or offer exchange membership to HFTs, and intraday trading has to be legal.

The key to our identification strategy is the staggered chronology of the start of HFT across international markets. Given that the start of HFT is not randomly allocated across markets, potential concerns about reverse causality or an omitted factor driving the HFT start dates need to be addressed. To address these concerns, we discuss a number of alternative interpretations of our findings.

Reverse causality could threaten a causal interpretation of our results. Whether this concern is plausible depends on whether HFTs can directly benefit from a decline in the fundamental informativeness of prices, i.e., from an increasing distance between prices and their fundamental values. One could argue that some of their strategies involve arbitrage between securities or markets, and this is why they might profit from inefficient prices. However, HFTs predominantly hold securities for short horizons, mostly intradaily with little overnight exposure. As a consequence, HFTs are unlikely to have sufficient patience to wait until prices converge to their fundamental values. Hence, it is unlikely that HFT profits are directly determined by the informativeness of prices.

The observed chronological order in this paper is inconsistent with the notion that HFTs enter informationally inefficient markets first. If anything, the markets in the U.S. and Germany, in which we observe the first start of HFT, rather rank among the more efficient markets. Also in the cross-section of stocks, the existing evidence is inconsistent with a preference for trading inefficient stocks: Brogaard et al. (2014) show that HFTs are more active in large than in small cap stocks, which suggests that they do not prefer to trade in less efficient markets, even if the potential profit, as a fraction of their trading volume, may be higher in such an environment.

A causal interpretation of the estimates in our study hinges upon the assumption that there is no unobservable confounding factor that drives both HFT and price informativeness. The staggered nature of events and the use of exchange fixed effects mitigate the concern that this assumption is violated, as any such confounding factor would have to be correlated with the chronological order of the start of HFT. Further, we analyze pre-trends and directly address potential confounders that might bias our results, by considering the impact of ETF growth, financial crisis, changes in market compositions, and the chronological order of the introduction of electronic trading platforms further below.

2.2 Measuring the Informativeness of Prices and Information Acquisition

2.2.1 Informativeness about Future Cash Flows and Investment

We measure the informativeness of prices following the approach suggested by Bai et al. (2016) and used in Kacperczyk et al. (2019). This measure captures how well the cross-section of firms' market prices in a given market at a given point in time predict the cross-section of their future cash flows or the cross-section of future investment, respectively.

Building on standard Q theory, Bai et al. (2016) consider firms that choose capital adjustments given a productivity shock and capital adjustment costs. Managers and outside investors receive signals about the firm-specific productivity shocks that are not perfectly correlated. Since the signals include information that is outside of the managers' information set, managers take into account market prices when making investment decisions. As a result, the efficiency of firms' investment decisions, and thus welfare, increases with the informativeness of market prices about the productivity shock.

Following Bai et al. (2016), we regress cash flows in the future one to five years on current market values, controlling for current cash flows and industry membership and scaling all variables by firms' total assets to obtain the measure of price informativeness. Market values are measured at the end of March following the end of the firm's fiscal year. We estimate

$$\frac{E_{i,t+k}}{A_{i,t}} = a_{m,t,k} + b_{m,t,k}\log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_{m,t,k}\left(\frac{E_{i,t}}{A_{i,t}}\right) + d^{s}_{m,t,k}\mathbf{1}^{s}_{i,t} + \varepsilon_{i,t,k} \tag{2}$$

for each market and each year, where i identifies each firm, m identifies the market, t the year, E is EBITDA, A is total assets, M is the market value of equity, $\mathbf{1}^s$ indicates the firm's first digit of the SIC code and $k = 1, ..., 5$. The informativeness of prices about cash flows (Priceinfo^{CF}) in horizon k in year t and in market m is given by the square root of the predicted variance of future cash flows using current market prices, which is the coefficient $b_{m,t,k}$ above multiplied by the standard deviation of $\log\frac{M_{i,t}}{A_{i,t}}$, the natural logarithm of scaled market prices.

We construct the informativeness of prices about investment similarly. Capital expenditures one to five years ahead are regressed on current market values, controlling for current investment, current EBITDA and industry dummies. We estimate

$$\frac{I_{i,t+k}}{A_{i,t}} = a_{m,t,k} + b_{m,t,k}\log\left(\frac{M_{i,t}}{A_{i,t}}\right) + c_{m,t,k}\left(\frac{E_{i,t}}{A_{i,t}}\right) + d_{m,t,k}\left(\frac{I_{i,t}}{A_{i,t}}\right) + e^s_{m,t,k}\mathbf{1}^s_{i,t} + \varepsilon_{i,t,k}, \quad (3)$$

where I denotes capital expenditure and the other variables are as defined above. Informativeness about investment (Priceinfo^I) with respect to horizon k in year t and in market m is given by the predicted variance of future investment based on market prices, which is the coefficient $b_{m,t,k}$ above multiplied by $log\frac{M_{i,t}}{A_{i,t}}$.

2.2.2 Idiosyncratic Volatility

Idiosyncratic, or firm-specific volatility denotes the portion of variation in stock returns that is not explained by asset pricing factors. French and Roll (1986) argue that this portion of variation captures the rate of the incorporation of private information into prices via trading. It has been used and supported as a measure of the incorporation of firm-specific information into prices by a number of articles, including Durnev et al.

(2003), Durnev et al. (2004), Chen et al. (2007), or Fernandes and Ferreira (2009).

Idiosyncratic volatility is computed as the standard deviation of the residuals obtained from a Fama-French three factor model estimated using daily returns over the last 12 months (Fama and French (1993)).[8]

2.2.3 Mutual Fund Holdings and Trades

Since investors acquire information to use it when constructing their portfolios seeking superior returns, any change in the extent of information acquisition should be reflected in their portfolio choices. In particular, if investors acquire less information their portfolio weights should be closer to that of a passive benchmark. We use mutual fund holdings data from the Thomson Reuters Global Ownership database to measure fund managers' active decisions as proxies for their information acquisition. We cannot observe the stated benchmark of all funds at all times, nor the constituents and their weighting of all indices, and it is also not clear that the official benchmark is the one actually used by the fund manager as a baseline portfolio. We therefore follow Doshi et al. (2015) who define Active weight as the deviation from the value-weighted portfolio, and show that their measure positively predicts fund performance. Active weight for fund i at time t is defined as

$$\text{Active weight}_{it} = \frac{1}{2} \sum_{j} |w_{it}^{j} - w_{it}^{jm}|, \tag{4}$$

with w being portfolio weights, j indicating stocks contained in the portfolio, and jm referring to the market capitalization-based weight of the stock in the portfolio under

[8] We thank Heiko Jacobs for providing the data used in Jacobs (2016).

consideration.

A smaller amount of active trading decisions taken by investors should also correlate with a reduced portfolio turnover as investors have less reason to adjust their portfolios. This should obviously hold true if the activeness of portfolio holdings decreases. To the extent that investors try not to reduce the amount of active positions in their portfolio, e.g., because their investors may have a preference for more active portfolios as opposed to closet indexing, they may do so by acquiring information about only a subset of stocks at a time, and consequently by replacing active positions less frequently. We therefore define an additional measure that we term Active trade, which refers to the active change in portfolio weights from one year to the next. Active trade for portfolio i at time t is defined as

$$\text{Active trade}_{it} = \frac{1}{2}\sum_{j}(w_{it}^{j} - w_{it}^{jm}) - (w_{it-1}^{j} - w_{it-1}^{jm}), \tag{5}$$

where we sum only over those stocks contained in the portfolio in both years.

The measures as defined above are defined on a fund-level. However, for our purposes, we need to obtain exchange-level rather than fund-level observations. This requires two additional steps: First, within each fund, we compute Active weight and Active trade separately for stocks listed on each exchange. Second, we aggregate the measure on an exchange-level by value-weighting the individual funds' exchange portfolios' active weights and trades.

2.3 Sample and Summary Statistics

The empirical analysis is based on annual data of an international panel of listed firms spanning the period from 1993 to 2012. We use accounting data from Compustat North America and Compustat Global and price and volume data from CRSP and Compustat Global for the U.S. and international exchanges, respectively. In cases of stock prices being available for the same firm and different exchange codes, we choose the exchange with the largest number of shares traded as the relevant one for the given firm and the given year.

Macroeconomic variables such as information on gross domestic product or trade are from the World Bank. We convert all values denominated in non-U.S. currency to U.S. dollars using exchange rates from the Federal Reserve System. We use the U.S. GDP deflator indexed to 2009 from the Federal Reserve Economic Data to turn nominal into real values.

We exclude firms with negative values of book equity and require that the current book value of total assets, current earnings, and future earnings are available.[9] For each market-year, we require at least 50 firm observations for the estimation of our informativeness measures. We select firms from 18 different stock markets. Figure 1 shows how the composition of our sample, which consists of 13 markets that exhibit the start of HFT during our sample periods and 5 counterfactual markets. The figure also shows the staggered start of HFT across these 18 markets and 20 years. The final sample consists of 330 rather than 360. This is because of missing financial statement and price data for these markets in the Compustat database, or because some firms get

[9] For robustness, we exclude financial firms, i.e., firms with a Standard Industry Classification code starting with 6, from our sample. When we omit financial firms, the results remain similar in terms of statistical significance and economic magnitude.

delisted, thus reducing the number of available data points to below 50. Further, we lose some market-year observations for longer horizons because information on some firms is not available for these longer horizons.

Table 2 shows descriptive statistics of our sample. The upper part displays statistics of our main measures of informativeness, whereas the bottom part shows other firm characteristics. Since our dataset comprises all stocks available in the major databases, the size of sample firms spans a wide range from a few million dollars to the largest global firms. The average firm is traded on an exchange where price informativeness is positive, even though there is a wide dispersion in the price informativeness. The bottom 5th percentiles of the measures are negative for all five horizons. This suggests that for some markets at certain points in time, the valuation of firms is negatively associated with future cash flows. The informativeness measures increase for longer horizons.

3 Price Informativness about Cash Flows and Investment

In this section, we analyze how the informativeness of prices changes with the start of HFT. After presenting the baseline results, we investigate pre-trends, address potentially confounding factors, and test cross-sectional implications.

3.1 Empirical Results

Panel A of Table 3 shows the results of regressions of Priceinfo^{CF} with respect to the next one to five years on the dummy variable HFT, control variables and exchange and time fixed effects. Standard errors are clustered at the year level.[10] The coefficient of HFT is negative for all five horizons, consistent with the notion that HFT decreases price informativeness. The coefficient in column 1 is negative, but with -0.35 rather small in terms of economic magnitude and fails to be statistically significant at conventional levels. The coefficient estimates for horizons 2 to 5 increases substantially and are statistically significant at least at the 5% level. Economically, the decrease amounts to approximately 50% of one standard deviation in column 2, or 88% relative to the mean value. The magnitude of this negative coefficient increases further for longer horizons, suggesting that the negative association between HFT and Priceinfo^{CF} becomes more pronounced for longer prediction horizons. For horizons 3, 4 and 5, the decrease even amounts to approximately 56%, 68% and 71%, respectively, of one standard deviation of the outcome variable.

Next, we analyze price informativeness about investment as an outcome variable. Panel B of Table 3 shows the regression results. Using Priceinfo^{I} as an outcome variable, the coefficient estimate of the dummy variable HFT is negative for all horizons and statistically significant at the 5% level for horizon 1 and at the 1% level for horizons 2 to 5. The magnitude of the negative coefficient of the HFT dummy increases with the time horizon, from -0.28 for horizon 1 to -1.48 for horizon 5. Economically, the effect of HFT is sizeable. The effect ranges from 27% (horizon 1) to 61% of one standard

[10] The results are qualitatively very similar when using two-dimensional clustering in both the year and exchange dimension, or when we use the bootrapping approach by Cameron et al. (2011) to adjust for a low number of clusters in the year and exchange dimension, respectively.

deviation in horizon 5. In sum, Priceinfo^I appears to deteriorate with the start of HFT, especially for longer horizons.

In order to illustrate the timing of the effect relative to the start of HFT, we estimate a modified version of Equation 1 in which we replace the HFT dummy variable with its interactions with event time dummies around the HFT start dates. Figures 2 and 3 plot the coefficients of the interaction terms for all five horizons. Adopting exchanges and non-adopting or later-adopting exchanges appear to evolve on similar paths in the periods prior to the start of HFT. These graphs show that the decrease in Priceinfo^{CF} and Priceinfo^I coincides with the estimated start of HFT. The decrease still persists several years after the start of HFT, suggesting a rather permanent change. We note that the confidence intervals widen substantially for later periods.

Appendix Table A2 shows the distribution of the coefficient estimates for 1,000 randomly assigned HFT start dates between 2003 and 2009. These estimates are on average close to zero, and only the left tails of the distributions are in the order of magnitude of the coefficients based on the observed HFT start dates. In Table A3 we investigate to which extent differences in the precision of the informativeness measure may confound our estimate. Differences in the precision can be caused by the the fact that the number of firm observations for each market year varies substantially, from 51 firm observations to 2,844 firm observations. Our results are comparable when we estimate weighted regressions where we use the number of observations used to compute the price informativeness measure in a given market-year as a weight. The coefficients remain very similar - in some cases they decrease very slightly - while the standard errors decrease for all 5 horizons.

We analyze the sensitivity of our results with respect to alternative HFT start dates based on trade size decreases, order-cancellation ratios and colocation offerings in Tables A4, A5 and A6 in the Appendix. The results are comparable, though order-cancellation based start dates seem to be associated with the most pronounced decrease in Priceinfo^{CF}.

3.2 Potentially Confounding Factors

Next, we directly investigate potentially confounding factors that might bias our estimates. More specifically, we consider the market-specific growth in exchange traded funds and potential changes in the composition of markets.

As ETFs and HFTs both grew substantially over the past decades, the correlation between ETF trading and HFT presence is positive. However, the direction of causality is not obvious. On the one hand, HFTs benefit from ETF trading by arbitraging between ETFs and their constituent securities. On the other hand, their activities enable a liquid ETF market. A plausible concern is that the growth in ETF trading might directly affect the informativeness of prices and that the results reported so far are confounded by this effect. There is evidence in the literature both for higher and for lower price informativeness resulting from ETFs (Israeli et al., 2017, or Glosten et al., 2017). ETF growth has been most pronounced in the U.S. If the decline in price informativeness can be explained by ETF growth, we expect weaker results when we exclude U.S. markets. The estimates in Panel A of Table 4 show that when excluding the U.S. the decline in forecasting power of prices for cash flows becomes even more pronounced for longer horizons. Panel B shows that the statistical and economic significance excluding U.S. markets is only slightly reduced compared to our baseline estimates for using price informativeness about investment as an outcome variable.

Further, to investigate whether ETF growth can explain our results, we include the natural logarithm of trading volume of the respective largest ETF replicating the performance of the exchanges' main indices as a further explanatory variable. Table 5 shows that, while the coefficients on ETF trading are negative and the size of the HFT coefficient decreases slightly, the economic and statistical magnitude of the coefficient estimate remain sizeable.

Price informativeness is also determined by the types of firms that the given market is composed of. The fundamental characteristics of firms traded on the exchanges can vary, or there can be entries or exits of firms that lead to changes in market compositions. Firms can become easier or harder to evaluate. For instance, Farboodi et al. (2018) argue that informativeness increases with firm age and firm size. Such changes in market composition may correlate with our HFT start dates and, consequently, bias our estimates. If firms became younger and smaller with the start of HFT, we would overestimate the drop in informativeness associated with HFT. Similarly, if firms became older and larger, the effect of HFT on informativeness would be underestimated. Further, the variability of cash flows or investment can change such that price informativeness decreases without changes in information acquisition. To account for these alternative explanations, we construct measures of average firm size, firm age and the standard deviation of earnings for each market and year. In Panel A of Table 6, we use these measures as outcome variables in a regression on HFT, control variables, and market and year fixed effects. The composition of markets with respect to firm size, firm age, or the variability of earnings is unchanged with the start of HFT, as suggested by the small and statistically insignificant coefficients in columns 1, 2, and 3.

In Panels B and C of Table 6 we include these characteristics as additional control variables to analyze to what extent changes in market composition affect our informativeness measures. Even if there are no overall changes in a certain direction, it could still be the case that if these characteristics change in a few markets, they are affecting our estimate. The coefficient of HFT decreases only slightly when regressing Priceinfo^{CF}, but still remains economically and statistically significant. For Priceinfo^{I} as an outcome variable, the coefficients increase slightly in magnitude. Collectively, these results reject the notion that the observed decrease in informativeness after the start of HFT can be explained by changes in the composition of firms.[11]

In the Appendix, we analyze further potentially confounding factors. The results of Tables A7 and A8 show that differential exposures to crises, or the correlation of HFT starts with the switch to electronic trading is unlikely to confound our estimates.

3.3 Cross-Sectional Tests

Next, we calculate the informativeness of stock prices for portfolios within markets. The goal of this exercise is, first, to test whether the effects are larger for firms that have greater exposure to HFT, and, second, to test whether the effects are larger for firms that are more difficult to value. To that end we form portfolios by splitting observations in each market and year at the median value of market capitalization, firm age, and Tobin's Q. We construct the price informativeness measures for each of these portfolios. The market capitalization split is motivated by the notion that HFTs are more active in stocks with large market capitalization, as supported by the findings in Brogaard et al. (2014). Information asymmetries are high in young and growth firms, as they

[11] The results are very similar when we use the linear or log-linear functional terms of these variables, or when we include the (log) mean firm size, mean firm age and the standard deviation of earnings as control variables in one single regression.

tend to have short histories of revenues and profits, and the bulk of their value consists of future investment opportunities. As a consequence, these firms are harder to value.

Table 7 shows the estimated coefficient of the HFT dummy variable for these portfolios. Panel A shows that the decrease in informativeness is more pronounced for large firms than for small firms. The difference of the coefficients for large and small firms is negative for all horizons, but is statistically weak and only significant at the 10% level for horizon 3. This result thus weakly supports the notion that stocks with a higher share of HFT activity experience larger decreases in price informativeness. Panel B shows that the decrease in Priceinfo^{CF} is greater for younger than for older firms. The difference is positive for all but horizons but horizon 2 and statistically significant at the 1% level for horizon 4 and at the 5% level for horizon 5. This finding supports the hypothesis that young firms are more difficult to value and, hence, more likely to suffer from decreased information acquisition. In Panel C, we compare the effect between firms with high and low Tobin's Q. The difference is negative for all five horizons, though it is statistically significant only at the 10% level for horizon 3. Taken together, there is some weak support that the decrease in Priceinfo^{CF} appears to be more pronounced for firms with high Tobin's Q, lending support to the notion that firms that are more difficult to value are more exposed to the effects of HFT.

4 Information Acquisition Activities

The previous analyses investigate the informativeness of prices. In this section, we turn to measures of information acquisition activities. Specifically, we study idiosyncratic return volatility as a measure of the incorporation of information into prices, and the activeness of mutual fund portfolios as a measure of the use of information acquired by investors.

4.1 Idiosyncratic Volatility

Table 8 shows the result of a regression of idiosyncratic volatility on the dummy variable HFT when controlling for different sets of variables such as firm characteristics, macroeconomic variables, and firm fixed and time fixed effects. According to the results in column 3, idiosyncratic volatility decreases by 0.23 percentage points after the start of HFT. This coefficient is statistically significant at the 1% level and also economically sizeable, as it corresponds to approximately 15% of one standard deviation and 8% of its mean value. This finding suggests that information-based trading decreases with the start of HFT.

In addition to the analysis of idiosyncratic volatility, we include a similar panel analysis using the bid-ask spread as an outcome variable in column 4 of Table 8. We find that the spread significantly decreases by approximately 10% of one standard deviation or 11% relative to its mean value after the start of HFT. This suggests that stock liquidity for trades that are sufficiently small so as to require only one execution at the best price has improved. This result is consistent with existing studies such as Boehmer et al. (2018). Since the adverse selection component forms an important part of the bid-ask spread, this result is consistent with a less informed order flow. This finding supports the notion that HFTs' ability to identify informed trading activity in the order flow enables trading by small uninformed traders at a low cost whereas large investors, who need to split their orders into small parts, face higher costs, which we cannot measure directly.

4.2 Holdings and Trades by Funds

Next we test whether a decrease in price informativeness is also reflected in holdings and trades of institutional investors. Table 9 shows that exchange-level active holdings by mutual funds decrease after the start of HFT. The coefficient is statistically significant at the 10% level only, but economically significant as it represents 39% of one standard deviation. The decrease in active trade is substantial, as suggested by the results in column 4. The coefficient is statistically significant at the 1% level and accounts for 68% of one standard deviation. Figure 4 plots the coefficient for individual years around this effect. For *Active weight* as an outcome variable, the coefficient of the HFT dummy turns negative after the start of HFT and reverts back to zero for later periods. The coefficients for the individual post periods fail to be statistically significant, which is not surprising, given that the coefficient for all post-HFT start periods is only statistically significant at the 10% level. When studying active trade as an outcome, the drop after the start of HFT appears to be slightly more persistent. These results indicate that the decrease in active holdings and active trades by institutional investors coincides with the staggered start of HFT across these markets. Institutional investors deviate less from the market portfolio in their holdings and trades, which is consistent with the notion that they acquire and process less information about individual securities. The reversal in *Active weight* some time after the start of HFT suggests that investors adjust their behavior by taking longer term active positions, consistent with the argument made earlier that a reduction in information acquisition can be associated with a lower portfolio turnover even if the activeness of the portfolio holdings is not reduced.

5 Conclusion

The two principal functions of financial markets are risk-sharing and efficient resource allocation. Accordingly, market quality is generally defined as consisting of two dimensions: liquidity and price discovery. While these two dimensions are naturally interlinked, this paper addresses the latter. As pointed out by Hirshleifer (1971), the efficiency of prices depends on two different types of activities, the incorporation of existing information into prices and the acquisition of new information.

The previous literature on high-frequency trading primarily studies the former. This paper examines the influence of HFT on stock price informativeness, related to cash flows and investment realized years into the future, and thus speaks to the latter.

The empirical evidence in this paper suggests that the informativeness of prices declines with the presence of HFT. With HFT, market valuations predict future cash flows and investment less precisely. This decrease becomes even more pronounced for longer prediction horizons. At the firm level, bid-ask spreads decrease, and idiosyncratic volatility, which captures the process of impounding firm-specific information into prices, also decreases. Institutional investors appear to take less active investment decisions after the start of HFT. In sum, our results provide empirical support for the the arguments of Stiglitz (2014), modeled theoretically by, e.g., Yang and Zhu (2018). The findings are consistent with the hypothesis that HFT reduces the gains from information for institutional investors through order anticipation, i.e., the ability to use past order flow to predict future order flow by institutional investors in the same direction, making the execution of large informed trades more expensive. Hence, institutional investors acquire less information and, as a consequence, market prices reflect less fundamental information. Thus, HFT distorts the basis for resource allocation. This effect of HFT unambiguously decreases total welfare, while the aggregate effect of HFT on welfare would have to consider the trade-off with effects on risk-sharing, which is facilitated by higher liquidity for small trades, as have been reported in the existing literature. Since different trading strategies are involved in beneficial liquidity provision and aggressive exploitation of order anticipation, market operators or regulators may reasonably consider potential mechanisms to rein in aggressive HFT.

References

Aitken, M., D. Cumming, and F. Zhan (2015). High Frequency Trading and End-of-Day Price Dislocation. *Journal of Banking & Finance 59*, 330–349.

Bai, J., T. Philippon, and A. Savov (2016). Have Financial Markets Become More Informative? *Journal of Financial Economics 122*, 625–654.

Baldauf, M. and J. Mollner (2018). High-Frequency Trading and Market Performance. Working Paper.

Bian, J., T. Su, and J. Wang (2017). Non-Marketability and One-Day Selling Lockup. Working Paper.

Boehmer, E., K. Y. Fong, and J. J. Wu (2018). Algorithmic Trading and Market Quality: International Evidence. Working Paper.

Bond, P., A. Edmans, and I. Goldstein (2012). The Real Effects of Financial Markets. *Annu. Rev. Financ. Econ. 4*(1), 339–360.

Brogaard, J., B. Hagströmer, and R. Riordan (2015). Trading Fast and Slow: Colocation and Liquidity. *Review of Financial Studies 28*(12), 3407–3443.

Brogaard, J., T. Hendershott, S. Hunt, and C. Ysusi (2014). High-Frequency Trading and the Execution Costs of Institutional Investors. *Financial Review 49*(2), 345–369.

Brogaard, J., T. Hendershott, and R. Riordan (2014). High Frequency Trading and Price Discovery. *Review of Financial Studies 27*(8), 2267–2306.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust Inference with Multiway Clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Carrion, A. (2013). Very Fast Money: High-Frequency Trading on the NASDAQ. *Journal of Financial Markets 16*(4), 680–711.

Chakrabarty, B., P. C. Moulton, and X. F. Wang (2018). Attention Effects in a High-Frequency World. Working Paper.

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

Conrad, J., S. Wahal, and J. Xiang (2015). High-Frequency Quoting, Trading, and the Efficiency of Prices. *Journal of Financial Economics 116*(2), 271–291.

Doshi, H., R. Elkamhi, and M. Simutin (2015). Managerial activeness and mutual fund performance. *The Review of Asset Pricing Studies 5*(2), 156–184.

Dow, J., I. Goldstein, and A. Guembel (2017). Incentives for Information Production in Markets where Price Affect Real Investment. *Journal of the European Economic Association 15*, 877–909.

Draus, S. (2018). High Frequency Trading and the Dynamics of Price Informativeness. Working Paper.

Dugast, J. and T. Foucault (2018). Data Abundance and Asset Price Informativeness. *Journal of Financial Economics 130*, 367–391.

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

Durnev, A., R. Morck, B. Yeung, and P. Zarowin (2003). Does Greater Firm-Specific

Return Variation Mean More or Less Informed Stock Pricing. *Journal of Accounting Research 41*, 797–836.

Edmans, A., I. Goldstein, and W. Jiang (2015). Feedback Effects, Asymmetric Trading, and the Limits to Arbitrage. *American Economic Review 105*, 3766–3797.

Fama, E. F. and K. R. French (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics 33*, 3–56.

Farboodi, M., A. Matray, and L. Veldkamp (2018). Where Has All the Big Data Gone? Working Paper.

Fernandes, N. and M. A. Ferreira (2009). Insider Trading Laws and Stock Price Informativeness. *Review of Financial Studies 22*(5), 845–1887.

Foucault, T., J. Hombert, and I. Roşu (2016). News Trading and Speed. *Journal of Finance 71*, 335–382.

French, K. and R. Roll (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics 17*, 5–26.

Glosten, L. R., S. Nallareddy, and Y. Zou (2017). ETF Activity and Information Efficiency of Underlying Securities. Working Paper.

Gorham, M. and N. Singh (2009). *Electronic Exchanges: The Global Transformation from Pits to Bits.* Elsevier and IIT Stuart Center for Financial Markets press. Amsterdam and Chicago: Elsevier and Stuart School of Business Center for Finacial Markets.

Hirshleifer, J. (1971). The Private and Social Value of Information and the Reward to Inventive Activity. *American Economic Review 61*(4), 561–574.

Israeli, D., C. Lee, and S. Sridharan (2017). Is there a Dark Side to Exchange Traded Funds? An Information perspective. *Review of Accounting Studies 22*, 1048–1083.

Jacobs, H. (2016). Market Maturity and Mispiricing. *Journal of Financial Economics 122*, 270–287.

Kacperczyk, M., S. Sundaresan, and T. Wang (2019). Do Foreign Investors Improve Market Efficiency? Working Paper.

Korajczyk, R. A. and D. Murphy (2019). High Frequency Market Making to Large Institutional Trades. *Review of Financial Studies 32*, 1034–1067.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica: Journal of the Econometric Society 53*, 1315–1335.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics 8*, 1–24.

Meyer, D. R. and G. Guernsey (2017). Hong Kong and Singapore Exchanges Confront High Frequency trading. *Asia Pacific Business Review 23*(1), 63–89.

Riordan, R. and A. Storkenmaier (2012). Latency, Liquidity and Price Discovery. *Journal of Financial Markets 15*(4), 416–437.

Stiglitz, J. E. (2014). Tapping the Brakes: Are Less Active Markets Safer and Better for the Economy? Speech at the Federal Reserve Bank of Atlanta 2014 Financial Markets Conference.

Tong, L. (2015). A Blessing or a Curse? The Impact of High Frequency Trading on Institutional Investors. Working Paper.

Van Kervel, V. and A. Menkveld (2019). High-Frequency Trading around Large Institutional Orders. *Journal of Finance*, forthcoming.

Weller, B. M. (2018). Efficient Prices at Any Cost: Does Algorithmic Trading Deter Information Acquisition? *Review of Financial Studies 31*, 2184–2226.

Yang, L. and H. Zhu (2018). Back-Running: Seeking and Hiding Fundamental Information in Order Flows. *Review of Financial Studies*, forthcoming.

Zhang, S. S. (2017). Need for Speed: An Empirical Analysis of Hard and Soft Information in a High Frequency World. *Journal of Futures Markets 38*, 3–21.

Figure 1: Staggered start of HFT across markets

This graph shows the HFT start dates across the markets in our sample.

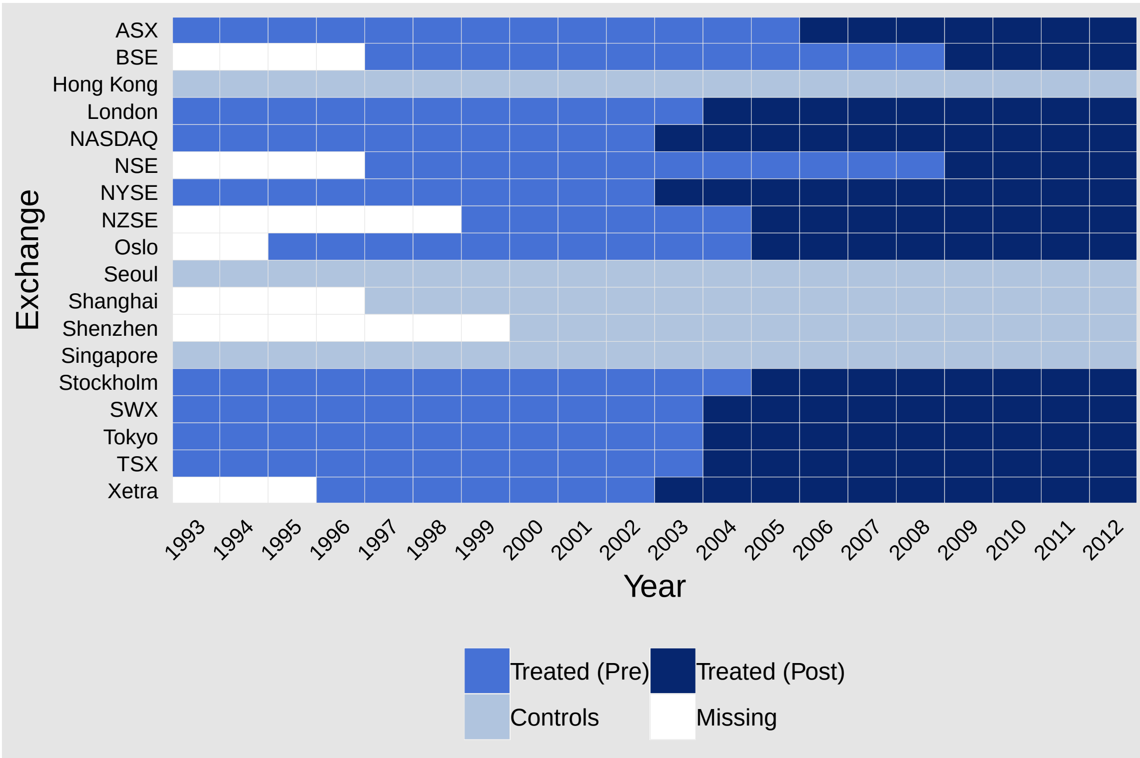

Figure 2: Informativeness about cash flows in event time

This graph shows the coefficient estimates and 90% confidence intervals of the HFT dummy interacted with event time dummy variables around the HFT start from a regression of Priceinfo^{CF} from horizon 1 to horizon 5 on exchange controls, time fixed effects and exchange fixed effects. The event time dummy variable indicates the number of years before or respectively after the start of HFT for the respective exchange.

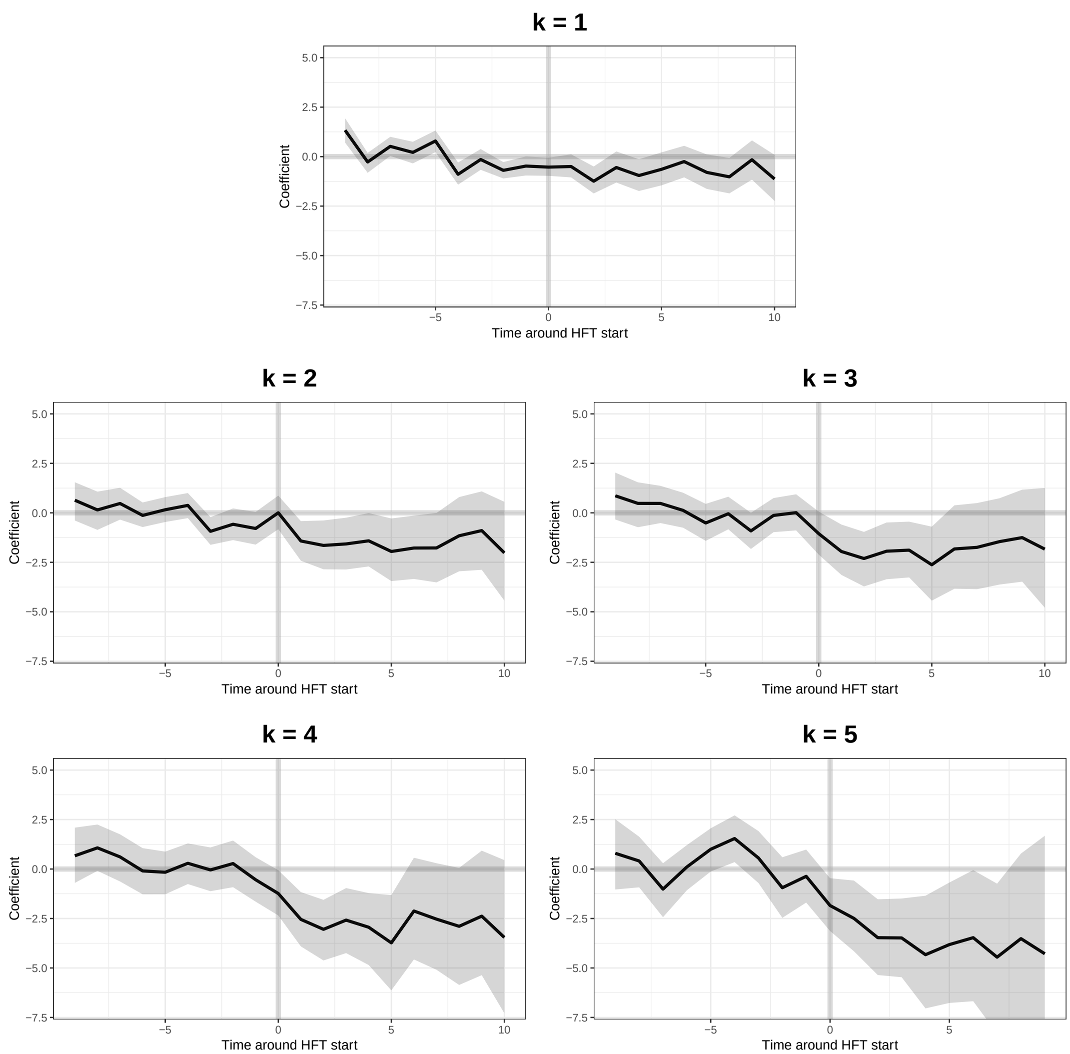

Figure 3: Informativeness about investment in event time

This graph shows the coefficient estimates and 90% confidence intervals of the HFT dummy interacted with event time dummy variables around the HFT start from a regression of Priceinfo^{I} over horizon 1 to horizon 5 on exchange controls, time fixed effects and exchange fixed effects. The event time dummy variable indicates the number of years before or respectively after the start of HFT for the respective exchange.

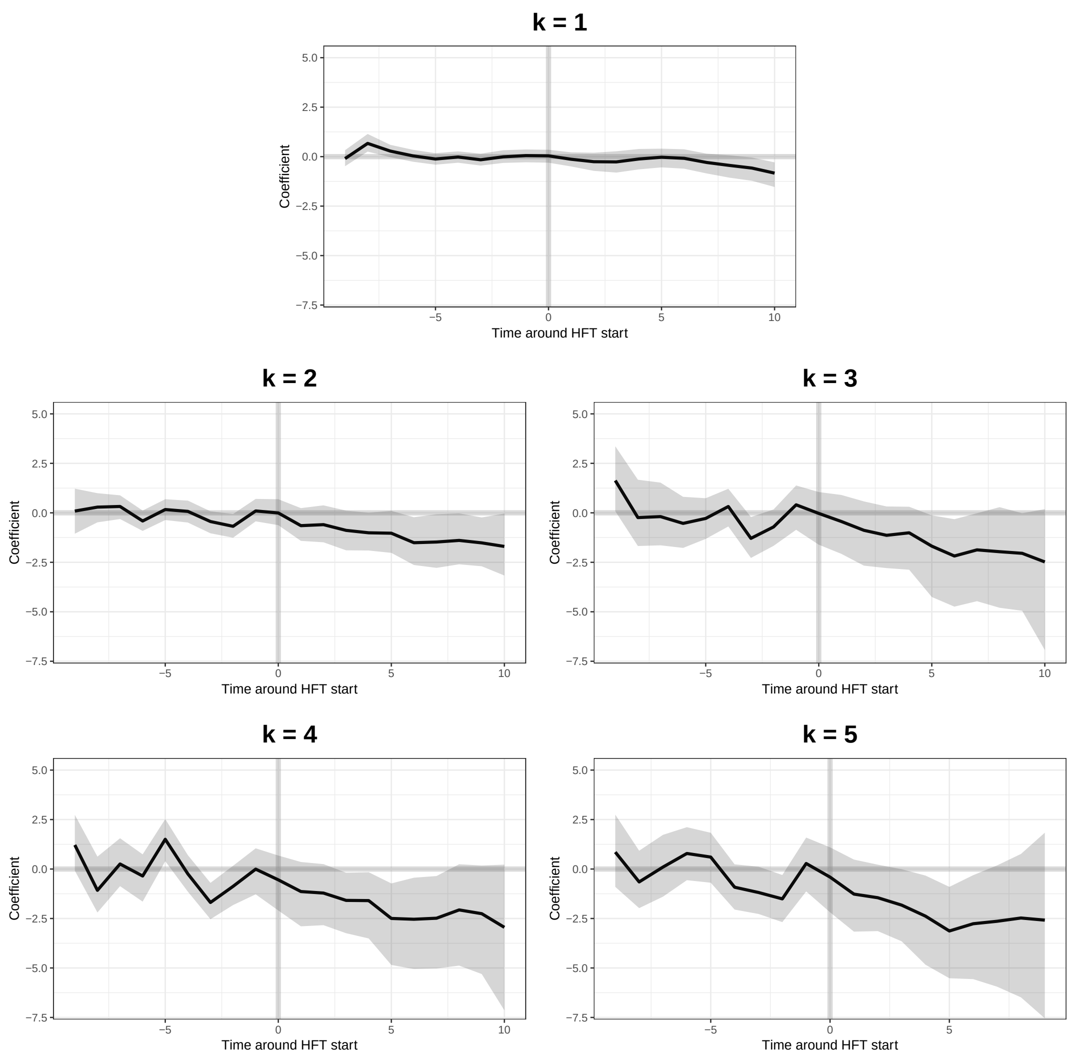

Figure 4: Holdings and trades by institutional investors

This graph shows the coefficient estimates and 90% confidence intervals of the HFT dummy interacted with event time dummy variables around the HFT start from a regression of Active weight and Active trade over horizon 1 to horizon 5 on exchange controls, time fixed effects and exchange fixed effects. The event time dummy variable indicates the number of years before or respectively after the start of HFT for the respective exchange.

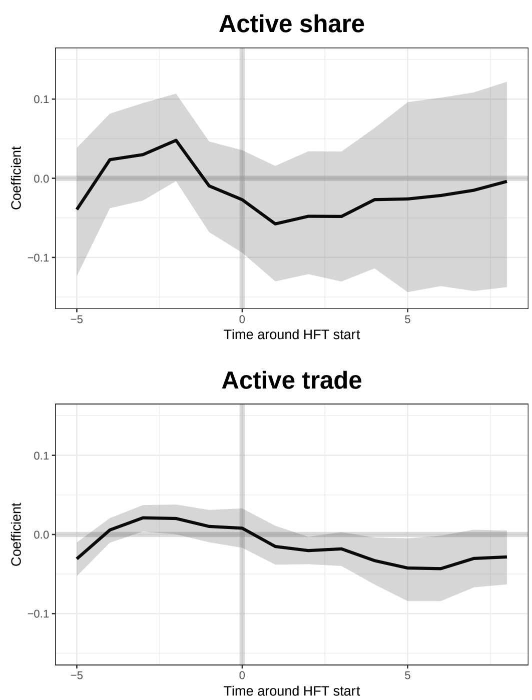

Table 1: Estimated HFT start dates across markets

This table shows HFT start dates based on trade size, order cancellation rates and colocation offerings (see Aitken et al. (2015)). We combine the stocks listed on the two Indian exchange, the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE), because NSE is the larger market in terms of trading volume for large firms that generally trade on both exchanges, many other stocks trade only on BSE.

Market	Country	HFT start date		
		trade size	order cancel	colocation
Nasdaq Stock Market	United States	Jan-03		Mar-07
Xetra	Germany	Jan-03		Aug-06
New York Stock Exchange	United States	May-03	Jul-03	Aug-08
SIX Swiss Exchange	Switzerland	Jan-04		Apr-12
New Zealand Stock Exchange	New Zealand	Nov-04		
Oslo Stock Exchange	Norway	Apr-05	Feb-05	Apr-10
Stockholm Stock Exchange	Sweden	Apr-05		Mar-11
Tokyo Stock Exchange	Japan	May-05	Apr-04	Jan-10
Toronto Stock Exchange	Canada	May-05	Jan-04	Apr-08
Australian Stock Exchange	Australia	Apr-06	Jun-06	Oct-08
London Stock Exchange	United Kingdom	Feb-06	Feb-04	Sep-09
National Stock Exchange	India	May-09	May-09	Jan-10
Bombay Stock Exchange	India	May-09	May-09	Jan-10
Counterfactuals				
Korea Exchange	South Korea			
Shanghai Stock Exchange	China			
Shenzhen Stock Exchange	China			
Singapore Exchange	Singapore			Jul-11
Stock Exchange of Hong Kong	Hong Kong			Oct-12

Table 2: Descriptive statistics

This table shows summary statistics for our sample spanning annual data from 1993 to 2012. Variable definitions are provided in Table A1.

Variable	Lower 5%	Median	Mean	Upper 5%	S.D.
Informativeness measures					
Priceinfo^{CF} (k=1)	-1.84	0.93	0.91	3.24	1.49
Priceinfo^{CF} (k=2)	-2.71	1.21	1.17	4.62	2.08
Priceinfo^{CF} (k=3)	-2.95	1.64	1.64	5.93	2.51
Priceinfo^{CF} (k=4)	-2.29	2.07	2.20	6.63	2.77
Priceinfo^{CF} (k=5)	-1.75	2.62	2.98	8.16	3.12
Priceinfo^{I} (k=1)	-0.12	0.83	0.99	3.05	1.01
Priceinfo^{I} (k=2)	-0.27	1.14	1.57	4.85	1.77
Priceinfo^{I} (k=3)	-0.67	1.41	1.71	5.45	1.90
Priceinfo^{I} (k=4)	-0.83	1.45	1.89	6.17	2.24
Priceinfo^{I} (k=5)	-0.70	1.77	2.12	6.84	2.40
Idiosyncratic volatility	1.08	2.39	2.81	6.05	1.60
Active weight	0.11	0.30	0.30	0.45	0.10
Active trade	0.04	0.12	0.12	0.18	0.04
Firm controls					
Market capitalization (USD million)	7	192	2,116	7487	11458
Book value of total assets (USD million)	9	287	5,887	14,317	48,786
Tobin's Q	0.63	1.20	1.75	4.69	1.66
Log(marketcap/assets)	-2.33	-0.35	-0.39	1.45	1.15
Cash/assets	0.01	0.11	0.18	0.63	0.20
Long-term debt/assets	0.00	0.06	0.12	0.44	0.15
EBITDA/assets	-0.23	0.08	0.06	0.25	0.17
Capex/assets	0.00	0.04	0.06	0.21	0.07
Firm age	1.00	6.00	7.03	17.00	4.98
Bid-ask spread	0.25	1.03	1.47	4.02	1.64

Table 3: Price informativeness about cash flows and investment

This table shows the results of a regression of price informativeness about cash flows (Panel A) and about investment (Panel B) of horizon k on the dummy variable HFT, a set of control variables and year and stock exchange fixed effects. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.348	-1.034***	-1.395**	-1.893***	-2.220***
	(0.306)	(0.341)	(0.543)	(0.425)	(0.595)
Electronic	0.478	1.178***	0.965*	-0.221	-0.058
	(0.309)	(0.383)	(0.532)	(0.516)	(0.729)
Log market size	-0.153	-0.348**	-0.551**	-0.572**	-0.529**
	(0.119)	(0.156)	(0.239)	(0.246)	(0.238)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.273	0.456	0.406	0.411	0.339
Obs	330	325	324	322	304

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.277**	-0.754***	-0.887***	-1.143***	-1.475***
	(0.114)	(0.193)	(0.303)	(0.380)	(0.366)
Electronic	-0.223	0.521	0.227	0.351	0.469
	(0.256)	(0.576)	(0.549)	(0.663)	(0.748)
Log market size	0.222*	0.180	0.212	0.300	0.159
	(0.111)	(0.163)	(0.132)	(0.182)	(0.280)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.492	0.527	0.410	0.337	0.381
Obs	326	321	320	319	301

Table 4: Excluding U.S. markets

This table shows the results of a regression of price informativeness about cash flows and investment of horizon k on the dummy variable HFT, a set of control variables and year and stock exchange fixed effects. Observations from U.S. exchanges are excluded. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.283	-0.950**	-1.403**	-2.104***	-2.560***
	(0.362)	(0.406)	(0.635)	(0.488)	(0.609)
Electronic	0.586	1.117*	0.820	-1.190	-1.187
	(0.450)	(0.548)	(0.802)	(0.918)	(1.226)
Log market size	-0.161	-0.342**	-0.550**	-0.508**	-0.505**
	(0.125)	(0.150)	(0.234)	(0.226)	(0.204)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.230	0.386	0.321	0.329	0.233
Obs	290	285	284	282	266

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.168	-0.613***	-0.532**	-0.812**	-1.153***
	(0.109)	(0.179)	(0.198)	(0.367)	(0.347)
Electronic	-0.328	1.045	0.257	0.406	0.617
	(0.430)	(0.940)	(0.943)	(1.076)	(1.033)
Log market size	0.239*	0.111	0.262	0.336*	0.181
	(0.116)	(0.188)	(0.162)	(0.175)	(0.277)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.518	0.552	0.462	0.344	0.382
Obs	286	281	280	279	263

Table 5: Controlling for ETF trading

This table shows the results of a regression of price informativeness about cash flows and investment of horizon k on the dummy variable HFT, the natural logarithm of the average monthly trading volume of the market's main exchange-traded fund, a set of control variables and year and stock exchange fixed effects. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.335	-0.948**	-1.280**	-1.674***	-1.796**
	(0.324)	(0.351)	(0.571)	(0.486)	(0.637)
Log ETF volume	-0.007	-0.043	-0.056	-0.106	-0.179*
	(0.031)	(0.030)	(0.045)	(0.071)	(0.100)
Electronic	0.478	1.181***	0.967*	-0.235	-0.062
	(0.310)	(0.382)	(0.531)	(0.505)	(0.745)
Log market size	-0.155	-0.351**	-0.556**	-0.579**	-0.581**
	(0.116)	(0.155)	(0.240)	(0.254)	(0.248)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.271	0.457	0.408	0.419	0.357
Obs	330	325	324	322	304

Panel B: price informativeness about investment

Dep. var.: Inv. predict	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.309**	-0.776***	-0.906**	-1.067**	-1.374***
	(0.123)	(0.210)	(0.322)	(0.410)	(0.401)
Log ETF volume	0.016	0.011	0.009	-0.036	-0.042
	(0.012)	(0.026)	(0.037)	(0.051)	(0.053)
Electronic	-0.221	0.522	0.228	0.342	0.464
	(0.256)	(0.576)	(0.550)	(0.659)	(0.745)
Log market size	0.225*	0.180	0.212	0.302	0.152
	(0.111)	(0.161)	(0.130)	(0.191)	(0.290)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.492	0.525	0.408	0.337	0.380
Obs	326	321	320	319	301

Table 6: Market composition

Panel A of this table shows the results of a regression of the natural logarithm of mean firm size (column 1), mean firm age (column 2) or the standard deviation of earnings (column 3) on the dummy variable HFT, a set of control variables, and year and stock exchange fixed effects. Panel B and C show the results of a regression of price informatinvess about cash flows and about investment as an outcome variable, respectively. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: Changes in market composition

Dep. var.:	Log size	Age	SD earnings
HFT	-0.233	-0.057	0.005
	(0.136)	(0.054)	(0.006)
Electronic	0.377***	-0.006	-0.025***
	(0.080)	(0.042)	(0.005)
Log market size	0.254***	0.097***	-0.006***
	(0.033)	(0.016)	(0.002)
Year FE	yes	yes	yes
Exchange FE	yes	yes	yes
Adjusted R2	0.925	0.850	0.857
Obs	330	330	330

Panel B: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5	(6) k = 1	(7) k= 2	(8) k = 3	(9) k = 4	(10) k = 5
HFT	-0.304	-0.953***	-1.337**	-1.858***	-2.094***	-0.324	-1.041***	-1.350**	-1.944***	-2.238***
	(0.301)	(0.323)	(0.545)	(0.413)	(0.570)	(0.313)	(0.342)	(0.547)	(0.427)	(0.590)
Electronic	0.409	1.048**	0.869	-0.273	-0.292	0.481	1.178***	0.970*	-0.245	-0.066
	(0.301)	(0.380)	(0.556)	(0.549)	(0.829)	(0.305)	(0.386)	(0.530)	(0.505)	(0.738)
Log market size	-0.200*	-0.445***	-0.621***	-0.614**	-0.714**	-0.194*	-0.339**	-0.616**	-0.491*	-0.510*
	(0.108)	(0.143)	(0.205)	(0.248)	(0.306)	(0.107)	(0.153)	(0.265)	(0.277)	(0.263)
Log firm size	0.185	0.356*	0.263	0.157	0.692					
	(0.166)	(0.197)	(0.338)	(0.382)	(0.542)					
Firm age						0.421	-0.118	0.762	-0.964	-0.239
						(0.642)	(0.591)	(0.830)	(0.767)	(1.162)
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R2	0.273	0.458	0.406	0.409	0.343	0.273	0.454	0.407	0.412	0.336
Obs	330	325	324	322	304	330	325	324	322	304

Panel B: price informativeness about cash flows (continued)

Dep. var.: Priceinfo^{CF}	(11) k = 1	(12) k= 2	(13) k = 3	(14) k = 4	(15) k = 5
HFT	-0.312	-0.996***	-1.364**	-1.886***	-2.240***
	(0.286)	(0.320)	(0.539)	(0.437)	(0.632)
Electronic	0.304	0.969**	0.787	-0.261	0.116
	(0.281)	(0.353)	(0.492)	(0.471)	(0.575)
Log market size	-0.196	-0.384**	-0.584**	-0.580**	-0.492**
	(0.114)	(0.158)	(0.236)	(0.234)	(0.222)
SD earnings	-6.991	-7.635*	-6.487	-1.596	7.386
	(4.258)	(4.053)	(5.251)	(5.891)	(11.020)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.285	0.463	0.409	0.409	0.340
Obs	330	325	324	322	304

Panel C: price informativeness about investment

Dep. var.: Priceinfo^I	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5	(6) k = 1	(7) k= 2	(8) k = 3	(9) k = 4	(10) k = 5
HFT	-0.348***	-0.769***	-1.078***	-1.256***	-1.603***	-0.295**	-0.756***	-0.914***	-1.166***	-1.516***
	(0.119)	(0.213)	(0.353)	(0.396)	(0.352)	(0.125)	(0.190)	(0.318)	(0.400)	(0.383)
Electronic	-0.109	0.545	0.541	0.521	0.702	-0.246	0.519	0.193	0.312	0.425
	(0.209)	(0.511)	(0.400)	(0.598)	(0.685)	(0.259)	(0.589)	(0.559)	(0.686)	(0.783)
Log market size	0.295**	0.197	0.430**	0.432**	0.336	0.277**	0.185	0.279*	0.365**	0.223
	(0.107)	(0.166)	(0.194)	(0.157)	(0.240)	(0.110)	(0.174)	(0.145)	(0.169)	(0.283)
Log firm size	-0.293	-0.063	-0.830	-0.492	-0.667*					
	(0.181)	(0.282)	(0.533)	(0.397)	(0.339)					
Firm age						-0.563*	-0.061	-0.798	-0.791	-0.797
						(0.324)	(0.605)	(0.508)	(0.904)	(1.075)
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R2	0.502	0.525	0.433	0.341	0.389	0.498	0.525	0.412	0.338	0.381
Obs	326	321	320	319	301	326	321	320	319	301

Panel C: price informativeness about investment (continued)

Dep. var.: Priceinfo^I	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.231**	-0.711***	-0.811**	-1.077***	-1.397***
	(0.103)	(0.188)	(0.295)	(0.356)	(0.338)
Electronic	-0.008	0.700	0.607	0.707	0.757
	(0.203)	(0.495)	(0.462)	(0.625)	(0.739)
Log market size	0.232**	0.142	0.245*	0.225	0.121
	(0.105)	(0.136)	(0.135)	(0.158)	(0.260)
Log SD earnings	0.873**	0.480	1.484*	1.239**	1.062
	(0.364)	(0.528)	(0.792)	(0.590)	(0.786)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.499	0.530	0.408	0.336	0.383
Obs	345	340	339	338	319

Table 7: Cross-sectional tests

This table shows the results of a regression of price informativeness about cash flows of horizon k on the dummy variable HFT, a set of control variables and year and stock exchange fixed effects for several portfolios. In Panel A, price informativeness is constructed based on all observations which are above or below the median market capitalization for the given market and year. In Panel B, the measure is constructed based on all observations which are above or below the median firm age for the given market and year. In Panel C, the measure is constructed based on all observations which are above or below the median value of Tobin's Q in a given market and year. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Dep. var.: Priceinfo^{CF}	(11) k = 1	(12) k= 2	(13) k = 3	(14) k = 4	(15) k = 5
Panel A: firm size					
HFT × large	-0.865***	-1.674***	-2.230***	-2.461***	-2.663***
	(0.233)	(0.412)	(0.645)	(0.657)	(0.698)
HFT × small	-0.422	-0.999*	-0.842	-1.790**	-1.974*
	(0.435)	(0.504)	(0.733)	(0.774)	(1.023)
Difference	-0.411	-0.731	-1.321*	-0.872	-0.706
	(0.358)	(0.554)	(0.738)	(1.126)	(0.787)
Panel B: firm age					
HFT × old	-0.510*	-1.309***	-1.383***	-1.216**	-1.311
	(0.251)	(0.373)	(0.471)	(0.542)	(0.899)
HFT × young	-0.518	-1.127**	-1.676**	-2.650***	-2.969***
	(0.405)	(0.403)	(0.587)	(0.467)	(0.755)
Difference	0.063	-0.229	0.420	1.570***	1.915**
	(0.425)	(0.466)	(0.401)	(0.498)	(0.908)
Panel C: Tobin's Q					
HFT × high Q	-0.865***	-1.673***	-2.229***	-2.459***	-2.661***
	(0.233)	(0.412)	(0.645)	(0.658)	(0.698)
HFT × low Q	-0.422	-1.000*	-0.843	-1.792**	-1.976*
	(0.435)	(0.504)	(0.733)	(0.774)	(1.023)
Difference	-0.410	-0.729	-1.318*	-0.868	-0.702
	(0.358)	(0.554)	(0.739)	(1.127)	(0.787)

Table 8: Idiosyncratic volatility

This table shows the results of a regression of idiosyncratic volatility and the bid-ask spread (both multiplied by 100) on the dummy variable HFT, a set of firm and market-varying control variables, year fixed effects and firm fixed effects. All unscaled values are converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the firm level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	(1)	(2)	(3)	(4)
Dep. var.:	Ivol	Ivol	Ivol	Spread
HFT	-0.289***	-0.578***	-0.234***	-0.163***
	(0.013)	(0.017)	(0.018)	(0.026)
Log price		-0.201***	-0.271***	-0.345***
		(0.006)	(0.012)	(0.018)
Log market cap		-0.350***	-0.212***	-0.330***
		(0.005)	(0.012)	(0.017)
Leverage		0.159***	0.175***	0.013
		(0.044)	(0.052)	(0.067)
Tobin's Q		0.295***	0.170***	0.098***
		(0.005)	(0.005)	(0.006)
Log GDP		0.454***	0.189***	-0.227***
		(0.011)	(0.048)	(0.065)
GDP growth		-0.007***	-0.030***	-0.026***
		(0.002)	(0.002)	(0.002)
Inflation		0.144***	-0.026***	-0.008
		(0.004)	(0.004)	(0.005)
Log trade		-0.486***	-0.272***	-0.398***
		(0.014)	(0.065)	(0.093)
Year FE	no	yes	yes	yes
Firm FE	no	no	yes	yes
Adjusted R2	0.008	0.451	0.739	0.670
Obs	157,469	157,469	157,469	157,469

Table 9: Holdings and trades by institutional investors

This table shows the results of a regression of active weight and active trade on the on the dummy variable HFT, a set of firm and market-varying control variables, year fixed effects and firm fixed effects. All unscaled values are converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the firm level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

	(1)	(2)	(3)	(4)
Dep. var.:	Active weight	Active trade	Active weight	Active trade
HFT	-0.060***	-0.039***	-0.039*	-0.027***
	(0.018)	(0.008)	(0.019)	(0.007)
Electronic			0.047***	0.003
			(0.015)	(0.007)
Log market size			0.031***	0.018***
			(0.008)	(0.004)
Year FE	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes
Adjusted R2	0.488	0.465	0.526	0.524
Obs	287	269	287	269

Appendix

Table A1: Definition of variables

This table defines the variables used in this paper.

Variable	Definition
Informativeness measures	
Priceinfo^{CF}	The coefficient of the natural logarithm of the market value of equity scaled by the book value of assets when regressing EBITDA in the future one to five years scaled by total assets on contemporaneous EBITDA scaled by assets on the natural logarithm of market value of equity and an industry dummy variable, multiplied by the standard deviation of the natural logarithm of the market value of equity scaled by the book value of total assets.
Priceinfo^{I}	The coefficient of the natural logarithm of the market value of equity scaled by the book value of assets when regressing capital expenditures in the future one to five years scaled by total assets on contemporaneous capital expenditures scaled by total assets, contemporaneous EBITDA scaled by assets on the natural logarithm of market value of equity and an industry dummy variable, multiplied by the standard deviation of the natural logarithm of the market value of equity scaled by the book value of total assets.
Idiosyncratic volatility	The standard deviation of the residual from a regression of daily excess returns on the Fama-French three factor model, based on the daily excess returns of the 12 months in the past fiscal year.
Active weight	Exchange-year specific deviation of actual holdings by mutual funds from those implied by the relative market capitalization of the firms, as specified in Equation 4.
Active trade	Exchange-year specific active changes in positions held by mutual funds, as specified in Equation 5.

Definition of variables (continued)

Variable	Definition
HFT measures	
HFT (trade size)	This dummy variable is set to 1 if HFT is assumed to have started on this market according to a pronounced decrease in trade size, and to 0 otherwise (see Aitken et al. (2015)).
HFT (order cancellation)	This dummy variable is set to 1 if HFT is assumed to have started on this market according to a pronounced increase in order-to-trade cancellation ratios, and to 0 otherwise (see Aitken et al. (2015)).
HFT	This dummy variable is set to 1 if HFT is assumed to have started on the given market either according to the order-to-trade cancellation ratio or according to trade size, and to 0 otherwise.
Coloation	This dummy variable is set to 1 if the market has launched colocation offerings, and to 0 otherwise (see Aitken et al. (2015)).
Firm characteristics	
Market capitalization (USD million)	Share price at the end of December of the given fiscal year multiplied by the number of shares outstanding.
Total assets (USD million)	Book value of total assets.
Tobin's Q	The market value of total assets (computed as the market value of equity plus total assets minus the book value of equity) scaled by the book value of total assets.
Cash/assets	Cash and cash equivalents divided by total assets.
Long-term debt/assets	Long-term debt scaled by total assets.
EBITDA/assets	EBITDA scaled by total assets.
Capex/assets	Capital expenditures scaled by total assets.
R&D/assets	Research and development expenditures scaled by total assets.
Firm age	The number of years since the firm has been first covered by Compustat.
Spread	Annual average of the bid-ask spread measured in percent of the stock price.

Definition of variables (continued)

Variable	Definition
Exchange-level factors	
Log GDP	The country's gross domestic product from the World Bank's World Development Indicators (WDI).
GDP growth	Annual growth of the country's gross domestic product.
Inflation	CPI inflation from WDI.
Log trade	The natural logarithm of the value of exports plus imports scaled by the gross domestic product from WDI.
ETF volume	Average monthly trading volume of the market's main exchange traded fund based on trading volume given by Datastream.
Crisis	This dummy variable is set to 1 if the annual return over the main stock market index of this market is less than minus 5%, and to 0 otherwise.
Electronic	This dummy variable is set 1 if the given market has switched to electronic trading, and to 0 otherwise (see Gorham and Singh (2009)).

Robustness

We randomize HFT start dates between 2003 and 2009, the actual first and last start date in our sample, to derive the distribution of the estimates under the null hypothesis. To do so, we randomly assign the HFT start dates and repeat the regressions shown in Panels A and B of Table 3 1,000 times. Panels A and B of Table A2 show that the estimated coefficients based on randomized start dates are typically close to zero. Only the bottom 1% yield estimates that are economically similar to the estimates based on the observed HFT start dates.

We recognize that there is substantial heterogeneity in the number of firms traded on the different exchanges. This condition might imply different degrees of precision with which we measure price informativeness. We employ weighted regressions, where we use the number of observations used to estimate price informativeness for a given exchange-year as the weight, to potentially improve the precision of our results. The findings in Panels A and B of Table A3 show that the coefficient estimates are quantitatively very close to our baseline regression results in Table 3.

Table A2: Random start dates

This table shows the distribution of the estimated coefficients of 1,000 random HFT start dates between 2003 and 2009 based on the regressions shown in Panel A and B of Table 3.

Panel A: price informativeness about cash flows

Estimate of beta	(1) k=1	(2) k=2	(3) k=3	(4) k=4	(5) k=5
Mean	0.012	0.017	0.023	0.025	0.024
S.D.	0.315	0.436	0.601	0.714	0.807
1st percentile	-0.681	-0.954	-1.365	-1.474	-1.771
5th percentile	-0.517	-0.703	-0.925	-1.115	-1.236
25th percentile	-0.192	-0.275	-0.374	-0.467	-1.236
Median	0.010	0.002	0.000	-0.023	0.006
75th percentile	0.218	0.328	0.434	0.530	0.603
95th percentile	0.535	0.729	1.036	1.242	1.343
99th percentile	0.721	1.044	1.479	1.736	1.842

Panel B: price informativeness about investment

Estimate of beta	(1) k=1	(2) k=2	(3) k=3	(4) k=4	(5) k=5
Mean	0.006	0.009	0.015	0.015	0.000
S.D.	0.205	0.379	0.524	0.603	0.647
1st percentile	-0.440	-0.858	-1.192	-1.375	-1.515
5th percentile	-0.318	-0.607	-0.821	-0.930	-1.045
25th percentile	-0.140	-0.249	-0.343	-0.404	-1.045
Median	0.001	0.013	0.003	0.001	-0.003
75th percentile	0.151	0.264	0.383	0.434	0.436
95th percentile	0.348	0.647	0.877	1.024	1.046
99th percentile	0.468	0.885	1.199	1.400	1.491

Table A3: Weighted regressions

This table shows the results of a weighted regression of Priceinfo^{CF} (Panel A) and Priceinfo^{I} (Panel B) of horizon k on the dummy variable HFT, a set of control variables and year and stock exchange fixed effects, where the exchange-year observations are weighted according to the number of firm observations in the given exchange-year that are used to estimate the price informativeness measures. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.360	-1.027***	-1.384***	-1.841***	-2.173***
	(0.280)	(0.331)	(0.480)	(0.399)	(0.565)
Electronic	0.429	1.105***	0.907*	-0.029	0.267
	(0.264)	(0.345)	(0.441)	(0.442)	(0.599)
Log market size	-0.181	-0.375**	-0.560**	-0.575**	-0.532*
	(0.115)	(0.169)	(0.228)	(0.241)	(0.257)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.322	0.493	0.451	0.443	0.384
Obs	330	325	324	322	304

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.294**	-0.704***	-0.895***	-1.091***	-1.453***
	(0.105)	(0.195)	(0.292)	(0.306)	(0.319)
Electronic	-0.172	0.115	0.295	0.505	0.644
	(0.188)	(0.378)	(0.369)	(0.415)	(0.549)
Log market size	0.201*	0.235	0.236	0.381**	0.255
	(0.103)	(0.150)	(0.137)	(0.162)	(0.262)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.567	0.565	0.476	0.433	0.459
Obs	326	321	320	319	301

Alternative HFT Start Dates

We investigate the sensitivity of the results with respect to alternative approaches to identify the entry of HFT in these markets. Table A4 shows the regression results when our estimated HFT start dates are determined based on trade size. These results are very similar to the main findings shown in Table 3, with the coefficients being slightly smaller for the predictability of cash flows. Next, we use an increase in order cancellation rates as an alternative indicator. Panel A of Table A5 shows that the results are similar under this alternative definition. The number of observations decreases to 256 in column 1, because the order cancellation-based start dates are not available for all 13 markets with HFT participation. For cash flow predictability, the coefficients are greater in terms of economic magnitude for all horizons. The analysis of investment predictability in Panel B shows that the coefficient estimates are slightly smaller, but close to the estimates in Panel B of Table 3.

Next, we investigate the launch of colocation offerings based on Aitken et al. (2015) as a further alternative. The coefficient of the dummy variable Colocation in the first five columns of Panel A Table A6 is negative for all five horizons. The economic magnitude for horizon 1 (40% of one standard deviation) is larger as compared to the baseline case in Panel A of Table 3 (23% of one standard deviation), but substantially smaller for horizons 2 and 5. When controlling both for the trade size-based HFT start date and colocation, the coefficient of colocation is mostly negative and statistically significant at the 10% level, but fails to be statistically significant at conventional levels for all horizons but the first. The economic significance of the coefficient estimate of HFT in columns 6 to 10 differ only slightly when compared to the baseline case in Table

3. These findings suggest that there is a small additional detrimental effect on price informativeness after colocation starts.

Using Priceinfo^I as the outcome, the coefficient of Colocation is positive in all horizons but horizon 5. For horizon 1 the coefficient is even statistically significant at the 5% level, suggesting that Priceinfo^I increases with the start of colocation. For horizon 5, the coefficient is negative and statistically significant at the 10% level. When controlling for both the HFT dummy and the Colocation dummy variable, the coefficient is positive for all horizons but horizon 5. The economic magnitude of the coefficient estimate of HFT is again almost the same as in our baseline case. Taken together, these results suggest that the decline in informativeness coincides with the estimated start dates based on increases in order cancellation and drops in trade size.[12]

[12]We note that the colocation offering dates in Aitken et al. (2015) differ from the ones given in Boehmer et al. (2018). The results in Table A6 are similar when we use the colocation dates indicated in the latter study.

Table A4: Start dates based on trade size

This table shows the results of a regression of price informativeness about cash flows (Panel A) and about investment (Panel B) for horizon k on the dummy variable HFT based on pronounced drops in trade sizes, a set of control variables and year and stock exchange fixed effects. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.352	-0.936**	-1.252**	-1.782***	-2.086***
	(0.311)	(0.385)	(0.570)	(0.459)	(0.631)
Electronic	0.475	1.167***	0.950*	-0.239	-0.080
	(0.306)	(0.383)	(0.529)	(0.517)	(0.728)
Log market size	-0.154	-0.336**	-0.533**	-0.559**	-0.510**
	(0.121)	(0.158)	(0.237)	(0.245)	(0.235)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.273	0.453	0.402	0.407	0.334
Obs	330	325	324	322	304

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.332**	-0.816***	-0.948***	-1.181**	-1.545***
	(0.128)	(0.210)	(0.317)	(0.426)	(0.399)
Electronic	-0.221	0.522	0.228	0.353	0.473
	(0.258)	(0.575)	(0.554)	(0.668)	(0.757)
Log market size	0.215*	0.170	0.202	0.293	0.148
	(0.112)	(0.163)	(0.131)	(0.183)	(0.278)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.494	0.528	0.411	0.338	0.382
Obs	326	321	320	319	301

Table A5: Start dates based on order cancellation

This table shows the results of a regression of the predictability of cash flows (Panel A) and of the predictability of investment (Panel B) of horizon k on the dummy variable HFT based on order cancellation rates, a set of control variables and year and stock exchange fixed effects. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT (order cancellation)	-0.448	-1.055***	-1.409**	-2.275***	-3.080***
	(0.302)	(0.362)	(0.532)	(0.443)	(0.659)
Electronic	0.469	0.894*	0.318	-0.756	-0.270
	(0.303)	(0.461)	(0.654)	(0.660)	(0.938)
Log market size	-0.132	-0.312*	-0.465*	-0.550**	-0.652**
	(0.134)	(0.160)	(0.244)	(0.257)	(0.275)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.397	0.508	0.452	0.470	0.402
Obs	256	253	252	251	238

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT (order cancellation)	-0.231	-0.617**	-0.780**	-1.109**	-1.431***
	(0.136)	(0.229)	(0.358)	(0.396)	(0.420)
Electronic	-0.185	0.943	0.386	0.868	0.880
	(0.342)	(0.725)	(0.758)	(0.725)	(0.796)
Log market size	0.201	0.134	0.245	0.412**	0.272
	(0.121)	(0.197)	(0.148)	(0.194)	(0.317)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.503	0.547	0.452	0.429	0.479
Obs	253	250	249	248	235

Table A6: Start dates based on colocation

This table shows the results of a regression of the predictability of cash flows (Panel A) and of the predictability of investment (Panel B) of horizon k on the dummy variable HFT based on colocation offerings, a set of control variables and year and stock exchange fixed effects. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5	(6) k = 1	(7) k= 2	(8) k = 3	(9) k = 4	(10) k = 5
Colocation	-0.605**	-0.370	-0.448	-0.479	-1.167**	-0.537*	-0.054	-0.016	0.114	-0.557
	(0.278)	(0.271)	(0.347)	(0.358)	(0.413)	(0.304)	(0.278)	(0.328)	(0.370)	(0.427)
HFT						-0.222	-1.021**	-1.391**	-1.920***	-2.111***
						(0.319)	(0.361)	(0.561)	(0.476)	(0.610)
Electronic	0.552	1.215***	1.003*	-0.213	0.025	0.548*	1.186***	0.967*	-0.236	0.014
	(0.321)	(0.396)	(0.576)	(0.566)	(0.796)	(0.316)	(0.381)	(0.535)	(0.530)	(0.737)
Log market size	-0.143	-0.228	-0.385*	-0.340	-0.273	-0.168	-0.350**	-0.551**	-0.568**	-0.545**
	(0.106)	(0.139)	(0.214)	(0.199)	(0.164)	(0.120)	(0.156)	(0.240)	(0.242)	(0.232)
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R2	0.278	0.440	0.386	0.378	0.309	0.277	0.454	0.404	0.409	0.338
Obs	330	325	324	322	304	330	325	324.000	322	304

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5	(6) k = 1	(7) k= 2	(8) k = 3	(9) k = 4	(10) k = 5
Colocation	0.275**	0.038	0.084	0.028	-0.727*	0.386***	0.287	0.381	0.402	-0.327
	(0.123)	(0.251)	(0.277)	(0.365)	(0.362)	(0.130)	(0.284)	(0.317)	(0.441)	(0.420)
HFT						-0.366***	-0.820***	-0.975***	-1.235***	-1.412***
						(0.109)	(0.201)	(0.310)	(0.426)	(0.380)
Electronic	-0.272	0.492	0.185	0.299	0.509	-0.275	0.479	0.173	0.295	0.512
	(0.256)	(0.609)	(0.542)	(0.686)	(0.763)	(0.258)	(0.605)	(0.559)	(0.692)	(0.777)
Log market size	0.272**	0.285*	0.339**	0.457***	0.328	0.233**	0.190	0.225	0.315	0.150
	(0.109)	(0.163)	(0.124)	(0.147)	(0.244)	(0.111)	(0.170)	(0.140)	(0.187)	(0.282)
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R2	0.491	0.514	0.394	0.318	0.358	0.498	0.526	0.410	0.337	0.379
Obs	326	321	320	319	301	326	321	320	319	301

Further Potentially Confounding Factors

In this section, we investigate further potentially confounding factors: differential exposures to crises, and the switch to electronic trading.

If price informativeness decreased in times of market crisis and the latter were correlated with the staggered start of HFT, the observed effect might be falsely attributed to HFT. To address this concern, we calculate yearly market returns based on the main national index of the given market and use those returns to create the dummy variable *crisis* which is set to 1 if the yearly observation is in the bottom quintile in terms of annual market return for the given exchange and the market return is lower than -5%[13]. We find that there is no significant correlation between market returns and the start of HFT. The correlation between HFT and market returns is -0.025; that between the HFT dummy and the crisis dummy is -0.04 and statistically indistinguishable from zero.

We re-run our main analysis where we exclude observations in which the crisis dummy variable is equal to 1 (Table A7, Panel A for cash flow and Panel B for investment predictability) and find similar results.[14] In Panels C and D, we interact the dummy variables *crisis* and *HFT* to investigate how HFT affects price informativeness in crisis versus normal times. The lack of statistical significance of the interaction terms except for the predictability of investment at horizons 3, and, more importantly, the retained significance of the coefficient estimate of the HFT dummy variable, reject the

[13]The results are not sensitive to the selection of this benchmark. We find very similar results when we use a benchmark of -10%, or -15%, or a benchmark of the return being in the bottom quintile for the given exchange, or the overall sample.

[14]We additionally run regressions where we exclude the years immediately before and after a financial crisis. The results show that the coefficient of the HFT dummy variable does not change substantially. If anything, the magnitude of the coefficient increases. The results are also comparable when we exclude observations from the financial crisis period between 2007 and 2009.

concern that our results can be explained by differential exposure to financial crises.

The introduction of electronic trading platforms is an alternative type of staggered event that affects financial markets and may have occurred in a similar sequencing. There is a substantial gap between the transition to electronic markets and the start of HFT. While the former happened mostly during the 1990s, the latter mainly occurred during the last decade. One might be concerned that the dummy variable Electronic explains the drop in the dependent variable, but if the two variables are highly correlated, this effect could be falsely attributed to HFT. The two dummy variables are in fact correlated: The raw correlation between Electronic and HFT is relatively large with a value of 0.38 and statistically significant at the 1% level. We choose to include the dummy Electronic as a control variable in our main analyses to assure that our results for HFT are not driven by Electronic. Here, we directly investigate the effect of Electronic on price informativeness. We exclude observations for which HFT = 1 and begin our sample period in 1990 because the introduction of electronic trading generally occurred several years before the start of HFT participation. Table A8 shows the results. In Panel A, the coefficient is even positive, and statistically significant at the shorter horizons. In Panel B, the coefficient estimate of Electronic is insignificant and its sign varies. Based on these results, we can reject the objection that we misattribute a potential impact of Electronic to the introduction of HFT.

Table A7: Controlling for financial crisis

This table shows the results of the predictability of cash flows and investment of horizon k on the dummy variable HFT, a set of control variables and year and stock exchange fixed effects. Panel A and B exclude all observations, for which the dummy variable crisis is equal to one. Panel C and D include the interaction term between the dummy variable HFT and the crisis indicator. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows excluding crisis years

Dep. var.: CF predict.	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.312	-1.089**	-1.342**	-1.972***	-2.321***
	(0.363)	(0.404)	(0.626)	(0.475)	(0.769)
Electronic	0.534	1.209***	1.072	-0.165	0.058
	(0.344)	(0.410)	(0.636)	(0.571)	(0.981)
Log market size	-0.122	-0.376	-0.690*	-0.717*	-0.527
	(0.201)	(0.237)	(0.367)	(0.380)	(0.361)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.262	0.437	0.400	0.416	0.321
Obs	275	271	269	268	251

Panel B: price informativeness about investment excluding crisis years

Dep. var.: Priceinfo^I	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.329**	-0.823***	-1.242***	-1.380***	-1.819***
	(0.147)	(0.202)	(0.400)	(0.480)	(0.469)
Electronic	-0.375	0.500	0.209	0.146	0.652
	(0.309)	(0.695)	(0.646)	(0.827)	(0.979)
Log market size	0.332***	0.440**	0.334	0.387	0.077
	(0.099)	(0.158)	(0.246)	(0.303)	(0.344)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.486	0.525	0.426	0.343	0.351
Obs	261	257	255	255	238

Table A7: Controlling for financial crisis (continued)

Panel C: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.453	-1.108**	-1.341**	-1.926***	-2.173***
	(0.326)	(0.388)	(0.598)	(0.451)	(0.640)
Crisis	-0.271*	-0.295	-0.075	-0.234	-0.104
	(0.134)	(0.215)	(0.268)	(0.425)	(0.330)
HFT × crisis	0.638	0.400	-0.395	0.137	-0.311
	(0.493)	(0.356)	(0.480)	(0.385)	(0.560)
Electronic	0.482	1.173***	0.949*	-0.232	-0.073
	(0.306)	(0.379)	(0.536)	(0.509)	(0.740)
Log market size	-0.164	-0.368**	-0.563**	-0.590**	-0.546**
	(0.119)	(0.155)	(0.251)	(0.256)	(0.258)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.274	0.454	0.403	0.407	0.334
Obs	330	325	324	322	304

Panel D: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
HFT	-0.264**	-0.777***	-0.920**	-1.136**	-1.519***
	(0.119)	(0.214)	(0.325)	(0.416)	(0.373)
Crisis	-0.078	-0.274	-0.292	-0.340	-0.041
	(0.126)	(0.213)	(0.293)	(0.474)	(0.337)
HFT × crisis	-0.109	0.072	0.131	-0.152	0.265
	(0.141)	(0.142)	(0.328)	(0.667)	(0.546)
Electronic	-0.232	0.505	0.214	0.324	0.474
	(0.257)	(0.586)	(0.548)	(0.666)	(0.757)
Log market size	0.215*	0.156	0.186	0.267	0.159
	(0.117)	(0.182)	(0.146)	(0.216)	(0.303)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.490	0.525	0.408	0.335	0.376
Obs	326	321	320	319	301

Table A8: Price informativeness and electronic trading

This table shows the results of a regression of the predictability of cash flows (Panel A) and of the predictability of investment (Panel B) of horizon k on the dummy variable Electronic, a set of control variables and year and stock exchange fixed effects for the period from 1990 to 2012. Market-year observations, in which the dummy variable HFT is equal to 1, are excluded. All values are expressed in real terms and converted to U.S. dollars. All variables are defined as in Table A1. The table reports point estimates. Standard errors, clustered at the year level, are given in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Panel A: price informativeness about cash flows

Dep. var.: Priceinfo^{CF}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
Electronic	0.628**	1.224***	1.048*	0.376	0.193
	(0.287)	(0.393)	(0.593)	(0.449)	(0.822)
Log market size	-0.094	-0.262	-0.627**	-0.531*	-0.598**
	(0.140)	(0.157)	(0.297)	(0.274)	(0.224)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.187	0.333	0.278	0.378	0.319
Obs	255	250	249	246	240

Panel B: price informativeness about investment

Dep. var.: Priceinfo^{I}	(1) k = 1	(2) k= 2	(3) k = 3	(4) k = 4	(5) k = 5
Electronic	-0.241	1.078	0.898	0.912	1.039
	(0.365)	(0.666)	(0.637)	(0.852)	(1.008)
Log market size	0.194	0.019	-0.105	-0.075	-0.234
	(0.144)	(0.211)	(0.181)	(0.218)	(0.284)
Year FE	yes	yes	yes	yes	yes
Exchange FE	yes	yes	yes	yes	yes
Adjusted R2	0.325	0.456	0.381	0.328	0.407
Obs	241	236	234	233	228

Recent Issues

No. 247	Mario Bellia, Loriana Pelizzon, Marti G. Subrahmanyam, Jun Uno, Draya Yuferova	Paying for Market Liquidity: Competition and Incentives
No. 246	Reint Gropp, Felix Noth, Ulrich Schüwer	What Drives Banks' Geographic Expansion? The Role of Locally Non-Diversifiable Risk
No. 245	Charline Uhr, Steffen Meyer, Andreas Hackethal	Smoking Hot Portfolios? Self-Control and Investor Decisions
No. 244	Mauro Bernardi, Michele Costola	High-Dimensional Sparse Financial Networks through a Regularised Regression Model
No. 243	Nicoletta Berardi, Marie Lalanne, Paul Seabright	Professional Networks and their Coevolution with Executive Careers: Evidence from North America and Europe
No. 242	Ester Faia, Vincenzo Pezone	Monetary Policy and the Cost of Wage Rigidity: Evidence from the Stock Market
No. 241	Martin Götz	Financial Constraints and Corporate EnvironmentalResponsibility
No. 240	Irina Gemmo, Martin Götz	Life Insurance and Demographic Change: An Empirical Analysis of Surrender Decisions Based on Panel Data
No. 239	Paul Gortner, Baptiste Massenot	Macroprudential Policy in the Lab
No. 238	Joost Driessen, Theo E. Nijman, Zorka Simon	Much Ado About Nothing: A Study of Differential Pricing and Liquidity of Short and Long Term Bonds
No. 237	Nathanael Vellekoop	Explaining Intra-Monthly Consumption Patterns: The Timing of Income or the Timing of Consumption Commitments?
No. 236	Aleksey Kolokolov, Giulia Livieri, Davide Pirino	Statistical Inferences for Price Staleness