High-Frequency Trading and Market Quality: Evidence from Account-Level Futures Data*

John Coughlan[†]
U.S. Commodity Futures Trading Commission

Alexei G. Orlov[‡]
U.S. Commodity Futures Trading Commission

This draft: July 29, 2022
The most recent draft is at https://ssrn.com/abstract=4069573.

Abstract

We use rich regulatory data on intraday transactions and end-of-day positions to study how high-frequency traders (HFTs) affect futures market quality. Panel estimation evidence shows that greater participation by HFTs is strongly associated with improvements in market quality (as measured by traded bid-ask spreads and the Amihud price impact), although higher rates of aggressive trading, such as those observed when HFTs trade directionally to reduce their positions, produce an adverse effect on market quality. Our results suggest that the market quality improvements brought about by HFTs' market-making outweigh the negative effects of HFTs' aggressive directional trading. We also find that while futures contracts are sensitive to market uncertainty (as measured by VIX), they are even more sensitive to own price volatility. We take advantage of the 2015 change in CME's daily settlement methodology for agricultural commodities to address potential endogeneity using a fixed-effects difference-in-difference setup. Our results are robust to relying on alternative estimation techniques, using overly conservative (clustered) standard errors, modeling various forms of cross-sectional and temporal dependence, and studying each market separately.

Keywords: Derivatives, futures markets, high-frequency traders, market quality, intraday transactions data, panel estimation.

JEL classification: G10, G12, G13, G23, Q02, C23.

---

*For helpful comments we thank Andrei Kirilenko, Wenwei Lin (discussant), Esen Onur, Michael Penick, Michel Robe, Pradeep Yadav (discussant), session participants at the 2022 Financial Markets and Corporate Governance conference, the 2022 Commodity and Energy Markets Association conference, and participants of the CFTC's Office of the Chief Economist research seminar. The authors are solely responsible for any errors or omissions. The research presented in this paper is coauthored by Commodity Futures Trading Commission (CFTC) economists in their official capacities with the CFTC. The analyses and conclusions expressed in this paper are those of the authors and do not reflect the views of other Commission staff, the Office of the Chief Economist, or the Commission.

[†]U.S. Commodity Futures Trading Commission, 1155 21st Street NW, Washington, DC 20581, USA; jcoughlan@cftc.gov.

[‡]Corresponding author. U.S. Commodity Futures Trading Commission, 1155 21st Street NW, Washington, DC 20581, USA; aorlov@cftc.gov.

1 Introduction

This paper takes advantage of rich regulatory data on intraday futures transactions and end-of-day futures positions for the period 2012 through 2021 to study the link between market quality and the extent of high-frequency traders' (HFTs') participation across markets and over time. The paper aims to add to the growing literature on HFTs (e.g., Menkveld, 2016; Brogaard, Carrion, Moyaert, Riordan, Shkilko and Sokolov, 2018; Brogaard, Hendershott and Riordan, 2019; Van Kervel and Menkveld, 2019) and, in particular, the subset of this literature that examines how HFTs impact market quality and, by extension, social welfare (e.g., Carrion, 2013; Menkveld, 2013, 2016).[1] We believe that the evidence from the commodity markets presented in this paper helps to further inform the ongoing debate over whether—and under which circumstances—HFTs augment or diminish market quality. A focus on commodity futures, including agricultural and energy, is particularly useful in light of the recent geopolitical and climate events that severely impacted these markets.

Most of the existing research on the link between high-frequency trading and market quality draws on equity order-book data. Although much of this research points to a positive association between HFTs and market quality, the debate on whether HFTs help improve social welfare has not yet been resolved. Hoffmann (2014) argues theoretically that HFT participation results in social welfare reduction due to rent extraction at the expense of slower traders as well as due to "arms races" among HFTs. Two recent studies confirm these results empirically: Goldstein, Kwan and Philip (2021) find that HFTs can crowd out profitable limit orders and exacerbate order imbalance, and Aquilina, Budish and O'Neill (2022) show that the negative effects of arms races can be substantial. Han, Khapko and Kyle (2014) and Budish, Cramton and Shim (2015) build models that show that fast market-makers push trading costs onto slower, fundamental investors. Brogaard, Hendershott and Riordan (2017) show empirically that HFTs had an adverse impact on market liquidity during the 2008 short-sale ban. In a recent study, Goldstein et al. (2021) find that HFTs can crowd out

[1]The rest of a representative (but not exhaustive) list of existing research includes Hendershott, Jones and Menkveld (2011), Hagströmer and Nordén (2013), Brogaard, Hendershott and Riordan (2014), Chaboud, Chiquoine, Hjalmarsson and Vega (2014), Kirilenko, Kyle, Samadi and Tuzun (2017), Menkveld and Zoican (2017), Korajczyk and Murphy (2019), and Malceniece, Malcenieks and Putniņš (2019). Biais and Foucault (2014) provide a survey of issues that the literature on HFT and market quality is attempting to address.

profitable limit orders and exacerbate order imbalance.

On the other side of this debate, many studies report evidence that HFT market participation does not adversely affect liquidity or market quality in general. For example, Carrion (2013) finds that HFTs are more likely to place limit orders into the market, thereby supplying liquidity, when spreads are wide. Menkveld (2013) reports that spreads tightened after HFTs entered Dutch stock trading on Chi-X Europe. Similarly, Conrad, Wahal and Xiang (2015) find that high-frequency trading activity does not reduce market quality and is associated with improvements in price discovery and trading costs. O'Hara (2015) notes that there is "general, but not universal, agreement that HFT market making enhances market quality by reducing spreads and enhancing informational efficiency" (p. 259). However, the extent and timing of HFTs' market-making activities remains an open question.

Our paper extends the existing literature in at least four dimensions. First, this paper is one of the first to study market quality and high-frequency trading using account-level regulatory data, and it is the first, to the best of our knowledge, to produce a comprehensive analysis using account-level regulatory data on such a scale in terms of both the number of contracts and the sample period. Raman, Robe and Yadav (2020) study the effect of increased participation of non-commercial institutional financial traders on market quality of the WTI crude oil futures market, which is one of the markets under consideration in our paper. Our paper is the first to study the impact of HFTs on multiple financial and non-financial commodity futures markets simultaneously. There is some research that uses account-level regulatory data to analyze market liquidity without focusing on the role of HFTs (e.g., Baker, McPhail and Tuckman, 2020). There is also a small subset of the literature that uses account-level data to study the behavior of HFTs without considering implications for market quality. For example, Kirilenko et al. (2017) study HFT behavior during the 2010 equity market flash crash, and Brogaard et al. (2019) examine the contribution to price discovery by different order types placed by HFTs and non-HFTs. Our rich regulatory data allow us to identify and classify individual trading accounts based on observed trading behavior over time as well as construct a novel set of control variables to more accurately isolate the effects of HFTs on market quality.

Second, one of the advantages of using the commodity futures setting is that each of these markets trades on a single exchange.[2] Unlike in the case of equities, which have been extensively studied by the existing literature, we can be sure there are no unobserved and potentially offsetting effects on market quality in off-exchange dark pools, nor are there any complications—with measuring the impact of HFTs' participation or otherwise—related to fragmented trading across multiple equity exchanges in the equity markets. The futures contracts studied in our paper are large, liquid markets, each traded on its own central limit order book (CLOB), and currently there are no analogues of dark pools in the futures markets. Indeed, some retail traders and other non-HFT participants in the equity markets may have decided to flee the public stock exchanges in response to HFTs' expanded market presence for the fear that they (non-HFTs) would be taken advantage of by the more informed traders with superior trading technology. Any such effects may not be captured by the existing studies of the equities markets.

Third, we thoroughly exploit both time-series and cross-sectional dimensions of our data in a panel setting using a variety of alternative econometric techniques. On the time-series side, our dataset spans 10 years' worth of intraday data which we aggregate at the daily level. Cross-sectionally, our dataset includes nine commodity futures markets that span diverse sectors—including agricultural markets, energy, metals, and financial markets—that feature various degrees and patterns of HFTs' involvement at any point in time across markets and over time within each market.

Fourth, we employ two complementary measures of market quality—the traded bid-ask spread and the Amihud price impact—and conduct a battery of robustness exercises that include, among others, several alternative ways of computing these measures of market quality. Traded bid-ask spreads provide a metric that is grounded in actual prices at which transactions occur, which complements the extensive existing research that relies on quoted spreads. The Amihud measure is often used as a measure of liquidity. We argue that the Amihud measure can also be construed as a market quality proxy that reflects several aspects of order book data, including depth of the book and price resiliency (see, for example, Amihud, 2002; Coën and de La Bruslerie, 2019). Through these measures, therefore, our analysis captures the transaction costs, resiliency (or reversion of pricing errors) and the

[2]In fact, all of the futures in our sample are traded on the Chicago Mercantile Exchange (CME) Group exchanges.

depth dimentions of market quality.[3]

In line with much of the equity literature, our main results suggest a strong, positive link between HFTs' activity in commodity markets, as measured by their daily market share, and improvements in market quality. In particular, our empirical models consistently show that HFTs' market share has a negative and highly significant coefficient on both measures of market quality, implying that greater HFT participation translates into greater liquidity (or lower price impact) and narrower traded bid-ask spreads.

Our unique dataset also allows us to measure HFTs' aggressive trading, which consumes liquidity and is often viewed as a temporary divergence from market-making (or liquidity provision) activity. The results show that aggressive trading by HFTs has a positive and significant coefficient on both measures of market quality, which suggests that liquidity dwindles and traded bid-ask spreads widen when HFTs trade more aggressively. As for economic significance, the coefficient on the HFT market share is about two to three times larger than the coefficient on the aggressive trading proxy. Taken as a whole, therefore, our results show that HFTs have an overall positive effect on market quality, even after acknowledging and controlling for HFTs' temporary diversions from market-making activity.

Our results are consistent with the mechanisms discussed in Carrion (2013), Menkveld (2013), and Brogaard et al. (2019), among others. Following Brogaard et al. (2019), we identify HFT accounts in the data as large, short-term speculators that trade frequently throughout the day and do not hold large intraday or overnight positions. Thus, it is reasonable to assume that the majority of HFTs' trading activity relates to market-making. We also find that the majority of HFT account trades are executed passively, in that those trades provide liquidity that is consumed by non-HFT traders after temporarily resting in the order book.

The next section provides the relevant market and institutional background. Section 3 describes the data, account classification, and variable construction including the market quality metrics. Section 4 reports our results and robustness checks with respect to al-

[3]Price efficiency is another important dimension. It has been extensively studied by the equity markets literature (e.g., Hendershott and Jones, 2005; Fotak, Raman and Yadav, 2014; Foley and Putniņš, 2016), and there is existing research on futures markets as well (e.g., Nie, 2019). Price efficiency can be estimated by the magnitude of price deviations from an informationally efficient price, athough pinpointing an informationally efficient price might not be trivial or incontrovertible.

ternative econometric techniques, alternative methods of constructing the market quality metrics, as well as studying the individual markets separately. Section 5 addresses potential endogeneity, and Section 6 concludes.

2 Background

2.1 Institutional details

Many financial markets exhibit high concentration driven by a relatively small number of large firms, some of which are high-frequency trading firms (HFTs), that may account for a sizable share of the overall activity in certain markets and during certain periods (see Figure 1). This is often due to the economies of scale that large financial firms can realize and the critical role that market-makers play in providing liquidity to other market participants. In recent years, HFTs have taken over the market-maker role in many futures markets. This market dynamics is of special interest to regulators and researchers since electronically traded markets have come to rely on HFTs for liquidity, and unlike banks or brokers, HFTs are not directly regulated.

[Insert Figure 1 here]

2.2 Open-outcry and electronic trading: a simplistic comparison

In many ways, the current market structure is the digital reincarnation of trading in open-outcry futures markets. In open-outcry trading (often referred to as "pit" trading), there were two main types of traders: floor brokers and market-makers (also known as "locals" at the time). Floor brokers brought client and other outside orders to the market, which were filled by market-makers, who were members of the exchange. Proprietary firms trading their own capital were also present in the pits.

Nowadays, futures markets trade electronically. CME launched its Globex platform in 1992, and by 2007 open-outcry volumes had fallen to less than 9% of the total trading volume.[4] The CME Group opened its Aurora, Illinois data center to clients for colocation in 2012, which provided low-latency market access to all products on the Globex platform.[5] Most clients still gain access to the market through an intermediary broker, but exactly how their orders get matched and filled is determined by the central limit order book (CLOB) and its associated matching engine. In the current system, client orders can be filled by any entity with a matching order on the CLOB, which is increasingly done by HFTs that employ specialized, low-latency (high-frequency) market-making algorithms.

Two important differences between these market paradigms are speed and anonymity. Obviously, the current electronic markets trade much faster than open-outcry pits did. In addition to the increased speed, electronic markets are also largely anonymous, which is in contrast to the pits trading, where brokers and locals often knew each other and incorporated that information into their trading decisions.

3 Data and main methodology

3.1 Data overview

Our main data source is a rich regulatory dataset provided to the CFTC by futures exchanges. The dataset contains records for every consummated trade in each futures market and identifies the associated trading accounts.[6] The data allow us to track each account individually and categorize them based on their trading behavior. Our sample period is January 1, 2012, through August 31, 2021. The time-series dimension consists of 2,399 trading days after including only days where all contracts in our sample are traded (not all contracts observe the same holiday schedules) to ensure a strongly balanced panel and, consequently, clear comparisons. On each trading day, all market-facing, outright trades in the most actively traded expiration for each contract (as measured by traded volume) are included. The data are aggregated by account for each trading day.

[4] "Ending an era, CME Group to shutter most futures pits," February 4, 2015, https://www.reuters.com/article/us-cme-group-closure/ending-an-era-cme-group-to-shutter-most-futures-pits-idUSKBN0L82QJ20150205.

[5] "CME colocation facility," https://www.advantagefutures.com/services/server-hosting/cme-colocation-facility/.

[6] These data are collected under 17 CFR Part 16.02 – Daily trade and supporting data reports, and is often referred to as "Trade Capture Reporting" or "TCR".

In order to capture a diverse sample of trading behavior, nine large, highly liquid futures contracts across different sectors are included in this study. Listed by sector, the contracts included are: agricultural markets—corn, soybeans, and wheat; energy and metals markets—natural gas, WTI crude oil, and gold; financial markets—Euro FX, E-mini S&P 500, and 10-Year Treasury Note.[7]

3.2 Account classification

As previously mentioned, our regulatory dataset includes the account information associated with each trade. We adopt the HFT classification methodology used by Brogaard et al. (2019). Every account is classified as belonging to one of three categories: high-frequency traders (HFTs), position traders, and other traders. Accounts are classified based on their observed trading behavior, not the assumptions made about the business model, or declaration thereof, of each trading firm. Accounts are classified as HFTs if they trade more than 0.25% of total volume in a calendar year, have an end-of-day inventory less than 20% of their trading volume, and never hold more than 30% of their daily trading volume at any point in time within the trading day. Accounts are classified as position traders if they end each day with a position of 20% or more of their trading volume, on average. Accounts are classified as "other" if they do not meet the criteria for either HFTs or position traders. Total volumes for each trading day is used to classify all accounts in each calendar year. Classifications are allowed to shift from year to year to account for changes in firm behavior and market participation over time.

Although there is no single agreed-upon method to identify HFTs, this approach captures the salient feature of HFTs—namely, that they are large short-term speculators (Brogaard et al., 2019). Some firms that are commonly thought of as HFTs might not meet the criteria used here—for example, if they hold substantial positions overnight or do not meet the traded volume threshold. Any misclassified accounts could attenuate the results, but the benefits of a standardized, comparable method applied across markets outweigh the potential costs of minor misclassifications.

[7] All contracts in this study trade on CME Group exchanges. Specific contract details are available at https://www.cmegroup.com/markets/products.html?redirect=/trading/products/.

3.3 Secular increases and variation in HFTs' market share

During our sample period of 2012 to 2021, HFT firms increased their market share in most of the major markets under consideration. Figure 1 and Table 1 show that HFT market share grew by over 25% in all but one of the nine contracts that we study in this paper, and several markets saw increases over 100%. In addition to this overall expansion of HFTs' market share, HFTs have increased their presence in various markets at various points in time, as is evident from Figure 1, and there is significant variation in HFTs' presence over time in any particular market. The secular increase along with significant variation across time and markets allow us to observe market quality under dramatically changing levels of HFTs' involvement and, thus, to be able to estimate more precisely the link between HFTs' market share and market quality.

[Insert Table 1 here]

Since HFTs have been well-established in equity and bond markets during our sample, it is especially useful to analyze the effects that increasing HFT activity may have on market quality in futures markets. The benefits of using futures data vis-à-vis equity data are particularly apparent in light of the evidence presented in Table 1 and Figure 1. It is worth noting that the general trend of increasing HFT involvement in financial markets was well under way by the time our regulatory dataset begins in 2012 (e.g., O'Hara, 2015), and HFT firms' decisions regarding which markets to enter and when are outside of the scope of this paper. The extant research takes as given the decision of HFTs to change the extent of their presence, and we follow the literature in this regard.

Since HFTs were present in all contracts included in our study from the beginning of the sample, their varied market share may raise the endogeneity question: does the level of HFT market share impact market quality, or was HFT participation changing in response to changes in market quality? This potential issue is similar to the observation by Conrad et al. (2015), who note that for HFTs, quote updates and prices are endogenous and jointly determined, and HFTs could be incentivized to trade more during times of greater liquidity. Our use of the traded bid-ask spreads addresses this issue, at least partially, because HFTs are not likely to be attracted to markets with narrower spreads, ceteris paribus.

There were several events that could have influenced HFT involvement during the course of our sample. In 2012, the CME changed the daily settlement methodology for agricultural products to be a blend of pit-traded activity and electronically-traded activity.[8] HFT involvement during the period of blended settlement was likely dampened since the HFT speed advantage is not as prominent when pit trades are included in the daily settlement calculations. This should be particularly apparent since settlement occurs at the the end of the trading day when HFTs need to close out their positions for the trading session. Blended settlement calculations ended in 2015 when CME closed its futures pits.[9] This blended settlement could be part of the reason that HFT involvement in the agricultural sector was the lowest in our sample's starting year of 2012 (see Table 1). All three agricultural contracts in our sample that were affected by the settlement rule change saw increases of 100% or more in HFT activity during the remainder of the sample period. Thus, the closure of the pits in 2015 can be used for robust identification to further help alleviate the endogeneity concerns. To this end, Section 5 leverages these market developments to conduct a fixed-effects difference-in-difference exercise.

Another development that may have encouraged greater HFT involvement was the CME's switch to its Market by Order data feed in 2017.[10] For each price level, these data furnish individual queue position, full depth of book, and individual order sizes.[11] This more detailed data feed likely provides more certainty for HFTs to manage their inventory and respond to intraday price movements. Since this change occurred in the second half of our sample when HFT involvement was already well established, it is unclear what impact this had, if any, on the overall HFT market share. These considerations suggest that in addition to the advantages of using futures regulatory data that were mentioned in Section 1, our dataset uniquely positions us to be able to address the potential issue of endogeneity that is likely prevalent in the literature on HFTs and market quality.

[8] "New settlement methodology for CBOT agricultural futures," June 8, 2012, https://www.cmegroup.com/tools-information/lookups/advisories/market-regulation/SER-6245R.html.

[9] "CME Group delays closure of open outcry futures trading in Chicago and New York," June 23, 2015, https://www.cmegroup.com/media-room/press-releases/2015/6/23/cme_group_delaysclosureofopenoutcryfuturestradinginchicagoandnew.html.

[10] CME Market by Order (MBO) data description, https://www.cmegroup.com/education/market-by-order-mbo.html.

[11] Ibid.

3.4 Market quality metrics and other variables

Our analysis uses two market quality metrics calculated from the regulatory transactions data: the Amihud price impact and the traded bid-ask spread. These metrics provide insights into market quality by using the actual trades that occurred in the market, as opposed to orders that may never be matched for a trade (as can be the case with quoted spreads in the existing equity literature). Therefore, these metrics reflect the actual trading activity and capture the changes in market quality associated with changes in the HFT activity, rather than summarizing the liquidity environment that may or may not be used by traders in the market.

3.4.1 Amihud price impact

The Amihud (2002) measure of price impact is generally viewed as a measure of market illiquidity, with higher values reflecting lower liquidity. More generally, the Amihud measure can be construed as a proxy for characteristics of the state of the order book, including bid-ask spread and, in particular, depth and resiliency (e.g., Amihud, 2002; Coën and de La Bruslerie, 2019). In this study, the Amihud measure is calculated using trade-by-trade price changes during U.S. daytime trading hours for each trading day. The heavy trading volumes observed in each of the markets in our sample provide a rich input to the granular, trade-by-trade calculations which result in a useful metric for the analysis. The Amihud measure is computed as follows:

$$Amihud = \frac{1}{N}\sum_{i=1}^{N}\frac{|\%\Delta P_i|}{volume_i},$$

where N is the total number of trades minus one, trades are indexed by i, and P_i is the realized price.

Volume is measured by the number of contracts traded, so that the Amihud metric measures price impact per contract traded.[12] Trading times include core daytime trading hours, which vary slightly by contract but roughly correspond to U.S. business hours. Since this metric requires a trade-by-trade analysis, the near-continuous trading observed during

[12]This is analogous to price impact per share traded commonly used in equity market analyses.

the core trading times is the most informative. Futures contracts trade nearly 23 hours per day, but significantly less volume is traded in overnight trading sessions where substantial time gaps between trades can occur. We therefore exclude overnight sessions from our analysis to capture the most continuous trading behavior. The daily settlement period is included with the rest of daily trading in our main analysis, and we check the sensitivity of our results to the exclusion of the settlement period, as well as to focusing solely on the settlement period, in Subsection 4.3.

Outliers are defined as any minute of each day that traded more than 2% of the total volume that day. Any minutes outside of the settlement period that meet this criterion are excluded from the daily calculation. The Amihud measure requires the trades to be ordered sequentially, so the calculations may be sensitive to heavily traded periods, which is why we flag the outliers according to the volume trade and do not simply winsorize (or otherwise trim) the final Amihud values themselves. This procedure removes from our sample non-representative trading periods which are often heavily directional, such as anomalous price spikes and popular market news announcements. Subsection 4.3 checks the sensitivity of our results with respect to the inclusion of outliers.

Figure 2 plots the Amihud price impact over time for each contract in our sample. Depicted are daily values along with 21-day moving averages to highlight the underlying trend. The Amihud measure in Figure 2 is the same measure that is used in our main analysis below and is defined as the volume-weighted average percentage change in price per 1,000 contracts. Since the sample covers the COVID-induced market turmoil, WTI crude oil and financial contracts all contain spikes during the period. The WTI spike is a significant outlier particularly in light of the negative futures price event that occurred on April 20, 2020.[13]

[Insert Figure 2 here]

3.4.2 Traded bid-ask spread

We also use the traded (or trade-implied) bid-ask spread (TBAS) as a measure of market quality. We follow Stoll (2000) in using this measure, which aims to capture the bid-ask

[13] "US oil prices turn negative as demand dries up," April 21, 2020, https://www.bbc.com/news/business-52350082.

spread implied by the observed trades when quote data are incomplete, unavailable, or less informative than the actual transactions data.[14] The TBAS is calculated by subtracting the average price of trades on the bid side from the average price of trades on the ask side and taking the average for the day (all averages are volume-weighted). In the words of Stoll (2000), "the quoted spread is a measure of total friction—the sum of real and informational frictions. The traded spread is a measure of real friction because it reflects real earnings for suppliers of immediacy" (p. 1488).

Our regulatory transactions dataset has an aggressor field which allows us to identify the bid and ask side of each trade. If for a particular transaction the buyer of a trade is the aggressor, the price for that transaction is assumed to be at the best ask. Conversely, if the seller is the aggressor, the price for the transaction is assumed to be at the best bid. The TBAS calculations defined in Stoll (2000) are done with averages across an entire trading day. In order to get a more accurate picture of trading activity, we calculate the TBAS for each minute of the trading day and then take the volume-weighted average of all aggregated minutes. Since the TBAS formula can produce negative spread values, any minutes with negative traded spreads are flagged as outliers and excluded from the daily average. Similar to the outliers flagged in the Amihud metric, minutes with negative spreads often show abnormal volatility and heavily directional trading, thereby making them less representative of the liquidity environment of the trading day. We check the robustness of our results with respect to the inclusion of outliers in Subsection 4.3.

The TBAS is computed as follows:

$$TBAS = \frac{1}{N}\sum_{i=1}^{N}\left[\left(P_j^{bid} - P_j^{ask}\right) \times n_j\right],$$

where

$$P_j^{bid} = \frac{1}{n}\sum_{i=1}^{n}\left(trades\ at\ bid\ price\ \times\ bid\ volume\right),$$

$$P_j^{ask} = \frac{1}{n}\sum_{i=1}^{n}\left(trades\ at\ ask\ price\ \times\ ask\ volume\right),$$

n is the volume traded in each minute, N is the total volume traded in the entire trading day, and i and j index minutes and trades, respectively.

[14]Our TBAS measure corresponds to twice the effective spread, which is frequenty used in the microstructure literature.

Figure 3 plots trade-implied bid-ask spreads over time for each contract in our sample. The bid-ask spreads are measured in traded price differences for each contract. The subplots show daily values as well as 21-day moving averages to highlight the underlying trend. Similar to the Amihud price impact graphs in Figure 2, the effect of the COVID-related market turmoil in early 2020 is clearly evident in many contracts, especially the financial and WTI crude oil contracts.

[Insert Figure 3 here]

The Amihud metric and TBAS provide two measures of market quality and together yield a robust, representative analysis of market quality changes. These granular metrics are calculated over a long time period and capture a wide range of market environments.

3.4.3 Other variables

Our analysis includes a rich set of control variables detailed in Table A1 of the Appendix. The primary variable of interest is the HFT market share, while the other (control) variables are included in our specifications to better isolate the effect of HFTs on market quality. A subset of our robustness exercises uses several variations of the primary dependent variables, which are detailed in Table A2 of the Appendix. The summary statistics for our entire (pooled) dataset are presented in Table 2, and summary statistics for individual markets are available from the authors upon request.

[Insert Table 2 here]

3.5 Panel regressions: fixed effects vs. random effects

A quick note on our main econometric methods may be in order. We use panel estimation to simultaneously model both the cross-sectional and time series dimensions of our data. A variety of regression types are run to allow thorough comparisons. As is well known, pooled OLS estimation does not allow for group heterogeneity, and thus may lead to heterogeneity bias. If the reader's main interest is to compare the impact of HFTs across markets (i.e., the between effect), then the results of our random effects models would most informative.

If, instead, the reader's focus is on studying the impact of HFTs across time (i.e., the within effect), then the results of our fixed-effects models should receive primary attention.

Further, and also as is well-known, if there are no omitted variables, or if the omitted variables are uncorrelated with the regressors, random effects estimates are unbiased and efficient. A violation of this assumption would result in inconsistent estimation. In contrast, fixed effects estimates are always consistent, and fixed effects models control for the possible presence of the omitted variable bias. For these reasons—and since we are agnostic about which assumptions would be most realistic—we report the results for both fixed effects and random effects models (and variations thereof) and let the reader choose which specifications are most in line with her beliefs and whether the reader places more emphasis on the time-series or on the cross-sectional dimension. All of our main results are consistent across the variety of models used in the analysis.

4 Results

Our analysis identifies a significant, negative relationship between HFT market share and both market quality metrics used—namely, the Amihud measure of price impact and the traded bid-ask spread. Since lower values of these metrics correspond to higher market quality, the negative relationship implies that market quality improves with a greater presence of HFTs. This main finding holds with and without the rich set of control variables and in the variety of robustness exercises.

4.1 Preliminary results

Table 3 reports the results of a univariate analysis and shows a negative relationship between the market quality metrics—the Amihud measure (Panel A) and traded bid-ask spread (Panel B)—and HFT market share that is statistically significant at the 5% level or higher across a range of estimation methods, including pooled OLS with and without heteroskedasticity-robust standard errors, fixed effects with and without heteroskedasticity-robust standard errors, fixed effects with AR(1) disturbances, random effects with and without heteroskedasticity-robust standard errors, and random effects with AR(1) disturbances.

It is noteworthy that these results hold over such a long sample period and with a diverse selection of contracts. As noted by O'Hara (2015), much of the current literature shows that HFT involvement benefits market quality, but there is still evidence that HFTs can degrade market quality in certain circumstances (see Hoffmann, 2014; Brogaard et al., 2017).

[Insert Table 3 here]

4.2 Main results

4.2.1 Market quality and HFTs

Table 4 controls for a variety of factors that may also affect market quality (besides the HFT market share), including percent of aggressive volume, order-to-traded-volume ratio, volume per trade, directional ratio, total stop volume, number of actively trading accounts, market volatility, total and active volume and open interest, macroeconomic variables (including credit spread, TED spread, and VIX), as well as season and year effects.[15] As in the univariate regressions in Table 3, our estimation methods include pooled OLS with and without heteroskedasticity-robust standard errors, fixed effects with and without heteroskedasticity-robust standard errors, fixed effects with AR(1) disturbances, random effects with and without heteroskedasticity-robust standard errors, and random effects with AR(1) disturbances.

[Insert Table 4 here]

Table 4 shows a negative relationship between HFT market share and both market quality metrics—Amihud and bid-ask spread (shown in Panels A and B, respectively) that is statistically significant at the 1% level across the variety of estimation techniques. Thus, Table 4 confirms that the highly significant results from the univariate regressions in Table 3 continue to hold in the presence of a battery of control variables. Our results agree with many recent studies showing that HFTs generally provide market liquidity (e.g., Menkveld, 2013; O'Hara, 2015).

[15] See Table A1 of the Appendix for the definitions of these variables.

4.2.2 Market quality and aggressive trading

We calculate the percentage of aggressively executed volume traded by HFT accounts and find a statistically significant, positive relationship between this variable and both of our market quality metrics (Table 4). This result suggests that when HFTs trade more aggressively, they tend to reduce market quality. HFT market-making behavior is often characterized as "supplying" liquidity by placing limit orders in the market (e.g., O'Hara, 2015), often as two-sided quotes at the best bid and ask simultaneously. When these orders are matched for a trade, they are often executed passively since the "aggressor" to a trade is the one that crosses the bid-ask spread to execute and the HFTs' orders are often matched after waiting in the order book for a few seconds (or milliseconds). The HFT market share captures much of the HFT behavior in our sample that falls into the market-making category, partly by the design of the classification criteria (see Section 3), and partly because market-making activity naturally drives large traded volumes and higher market share. The percent of aggressively traded volume is important in that it captures instances when HFTs diverge from the market-making activity. One common example is when an HFT firm builds up an inventory within the trading day and attempts to reduce its position by trading aggressively in the market in one direction to return to a more neutral level of inventory.

Our analysis quantifies the dichotomy that the recent literature has struggled with, often because the data used heretofore are able to shed light only on one side of the question. On the one hand, HFT activity is often positively associated with market quality (e.g., Menkveld, 2013). On the other hand, when HFTs trade more aggressively, they can temporarily consume liquidity and diminish market quality (e.g., Brogaard et al., 2017). Our result is similar to what Conrad et al. (2015) found in their study of equities markets: HFT activity does not harm market quality overall, even when HFTs are associated with large liquidity extractions. The trade-related details in our regulatory data, combined with the long sample period, allow us to uncover this dynamic, nuanced relationship between HFTs and the markets they interact with. When HFTs behave in a market-making capacity (i.e., trading frequently on both sides of the market and keeping their inventory close to zero), they supply liquidity and augment market quality. When they trade more aggressively, they consume liquidity, thereby reducing market quality.

Our model also sheds light on the magnitude of each of these two opposing effects. The coefficient on the HFT market share variable is approximately three times the size of the coefficient on the percent of aggressive volume variable in Panel A of Table 4, and twice its size in Panel B. These results suggest that, on average, the effect of the HFTs' market-making has a much larger beneficial impact on market quality than the negative effect of their aggressive trading. Stated differently, HFTs enhance market quality overall, even after acknowledging and controlling for the temporary periods of aggressive, directional trading.

4.2.3 Other determinants of market quality

Volatility measures

To control for volatility, we include both the VIX and the within-contract volatility calculated as a 21-day annualized volatility observed in each contract, using changes in daily settlement prices for the 21 days preceding each reference date (to help avoid simultaneity concerns). VIX has a positive, significant effect on both measures of market quality, which is in line with expectations as higher market stress (or uncertainty) tends to diminish market quality even in non-financial contracts. The coefficient on within-contract volatility is also positive and significant and is much higher in terms of economic significance relative to the VIX coefficient. Our results suggest, therefore, that while each futures contract is influenced by overall financial market stress, it is much more sensitive to own price volatility.

Other control variables

Dummy variables for seasonality and year are included in all specifications. The year variable controls for large shifts that occur over time. The seasonality variable controls for the seasonal effects that are especially prevalent in commodity markets subject to crop cycles (corn, wheat, and soybeans) as well as in the demand for crude oil and natural gas during warm vis-á-vis cold weather. Our results point to significant seasonal and year effects.

Total volume and open interest for all contract expirations are included to control for size differences across markets and the difference in market size over time. Additionally, we include the percentage of total volume and open interest in the active contract. This is a measure of how much trading activity is spread out in the futures curve vs. being concentrated in the active contract. This controls for both the amount of spread trading and the roll period.

4.3 Robustness

This subsection evaluates the robustness of our results with respect to (i) clustering of standard errors, (ii) using alternative estimation methods (e.g., GLS consistent estimation), and (iii) modeling additional potential sources of heteroskedasticity, including groupwise heteroskedasiticy (i.e., differences in variance across markets), autocorrelation, and temporal correlation (i.e, correlation of errors across markets). We also re-run our main regressions under panel-corrected standard errors (PCSE), which are robust to general forms of cross-sectional and temporal dependence, including, for example, time-varying cross-correlation due to persistent common shocks.

Table 5 shows the results for pooled OLS with heteroskedasticity-robust standard errors and fixed effects regressions as a frame of reference, and checks the robustness of our main results to using panel-corrected and clustered standard errors. In particular, we consider panel-corrected standard errors with no autocorrelation within panels, panel-corrected standard errors with first-order autocorrelation within panels, panel-corrected standard errors with panel-specific first-order autocorrelation, first-order autocorrelation within panels with heteroskedastic panel-level disturbances, a fixed effects model with standard errors clustered at the market level, and a random effects model with standard errors clustered at the market level.

[Insert Table 5 here]

The key relationships between the market quality metrics, HFT market share, and the percent of aggressive trading remain the same under the panel-corrected and clustered standard errors. It should be noted that with the relatively small cross-sectional dimension (nine futures markets) and the relatively large time-series dimension (2399 days) of our sample, clustering at the market level leads to excessively large standard errors and, therefore, makes it much more difficult to find any significance. Notwithstanding the small number of clusters with a large number of observations in each cluster, our key variables of interest retain their significance. The clustered standard errors in our fixed-effects model correct for within-cluster correlation and heteroskedasticity, while our main fixed effects model is focused on the differences between groups. The fact that both models share similar results strengthens the key findings discussed in Subsection 4.2.

Table 6 uses feasible generalized least squares (FGLS) estimation to model various forms of cross-sectional and temporal dependence. FGLS methods estimate unknown parameters in the presence of some correlation among residuals. In particular, we allow for heteroskedasticity across panels (Model (4)), heteroskedasticity and cross-sectional correlation (Model (5)), and heteroskedasticity across panels along with autocorrelation within panels (Model (6)). The results reported in Table 6 are generally consistent with the results in Table 4, which lends further credence to our main findings.

[Insert Table 6 here]

Table 7 studies each commodity futures market separately and shows OLS results with heteroskedasticity-robust standard errors for each individual market. With few excepetions, the results and key relationships are consistent across markets. The nine contracts in our sample were selected because they are all large, liquid markets, and come from a diverse set of sectors and, therefore, reflect variation driven, in part, by differences in underlying market structures. The HFT market share is negative and highly significant for almost all markets for both market quality measures.[16] The percent of aggressively traded volume is positive and significant in almost all markets for both market quality measures. Further, the coefficients on the volatility variables (i.e., VIX and contract-specific volatility) are largely positive and significant. Thus, the individual markets results are consistent with the main results and show that no single contract or a small group of contracts drives our panel estimation findings.

[Insert Table 7 here]

[16]The only exceptions are natural gas and wheat in the Amihud regressions and gold (marginally significant at 10%) and natural gas in the bid-ask spread regressions.

Tables 8, 9, and 10 show results from OLS with robust standard errors and fixed-effects models under a variety of alternatively calculated market quality metrics.[17] First, in futures markets the daily settlement period is an important part of a trading day where a large portion of trading occurs and significant price movements are not uncommon. To check the sensitivity of our results to the settlement period as well as to outliers, Table 8 presents the results for pooled OLS with heteroskedasticity-robust standard errors as well as for the fixed effect model with the Amihud price impact computed either with or without outliers and for various trading periods: full day including the settlement period, full day excluding the settlement period, and the settlement period only. Our key main results hold for these alternative methods of computing the Amihud measure. This is an important finding because the trading environment during the settlement period can be vastly different from the rest of the day, especially in terms of elevated volume and volatility.

[Insert Tables 8, 9, and 10 here]

Next, all futures contracts have a minimum tick size set in the contract specifications and enforced by the matching engine. Many traders use the minimum tick size in each market to normalize their trade strategies and sizing when the market price can be quite different across markets. We take this into account and calculate the Amihud price impact in minimum ticks to contrast the percentage change in price used in the main results section. Table 9 presents OLS and fixed-effects regressions for the Amihud measure based on price changes measured in minimum ticks while also allowing for various trading periods (i.e., full trading day, excluding the settlement period, or settlement period only) as well as including or excluding outliers. Table 9 shows that our main results in Panel A of Table 4 are immune to computing the Amihud measure using minimum ticks instead of percent price changes.

Finally, the traded bid-ask spread can be calculated in a variety of ways, and we check the sensitivity of our main results to a representative set of these alternative methods: both simple and volume-weighted averages as well as averages that include minutes with negative spreads (which were excluded as outliers in the main results). Table 10 shows that our results in Panel B of Table 4 continue to hold under these alternative ways of computing the

[17] Specific alternative methods of computing the Amihud and the bid-ask spread are described in Table A2 of the Appendix.

bid-ask spread. Thus, this section confirms that (i) the HFT market share is positively and significantly associated with market quality, (ii) the percent of aggressively traded volume has a partially offsetting, adverse impact on market quality, (iii) both VIX and the contract-specific volatility are negatively and significantly linked to market quality, and (iv) all of these results are highly robust.

5 Addressing potential endogeneity

The extant research takes as given the decision of HFTs to change the extent of their presence in the market, and thus far we have followed the literature in this regard. Similar to the existing body of HFT research, HFTs were present in all contracts included in our study from the beginning of our sample. The HFT's varied market share may raise the endogeneity question and, by extension, the reverse causality question: was HFT participation changing in response to changes in market quality? Our use of the traded bid-ask spreads partially addresses endogeneity because HFTs are unlikely to be incentivized by relatively narrow spreads. To address potential endogeneity more formally, we (i) perform difference-in-difference estimation to analyze the effects of the change in CME's settlement methodology on the HFT market share and on market quality, and (ii) study a lagged impact of HFT market share on market quality.

5.1 Change in CME settlement methodology

As was mentioned in Subsection 3.3, in July 2012 the CME Group changed it daily settlement methodology for agricultural products to be a blend of pit-traded activity and electronically traded activity, which was not implemented in other sectors. While HFTs were already present in these markets, this change could have negatively affected the HFTs' involvement. Total volume traded in the pits was very low by the time our sample began. However, the use of pit trades in the calculation of the final daily settlement price in agricultural futures posed several issues for HFTs. For example, the HFT speed advantage was somewhat hindered during the settlement period when pit trades were included in the daily settlement calculations because HFTs did not receive pit trades until after the traders in the pit placed them and all traders in the pit would naturally be aware of those price-forming trades before or at least at the same time as the HFTs.

Additionally, during the settlement period a large percentage of daily volume is traded, on average, and large price movements are common. The settlement period also happens to be the end of the trading day when HFTs need to close out their positions for the trading session and calculate their daily profit and loss (P&L). Any uncertainty during this critical time would make it harder for HFTs to manage their price and inventory risk as they end the trading session. Further, trades that occurred in the pit could often be large enough to drive potentially significant price movements on their own. Finally, large pit trades might not be known to those outside the pit until after the market closed if they were placed in the last few seconds of the trading session, leaving almost no opportunity to react to them until the next trading session.

The blended settlement methodology was abandoned in 2015 when CME closed its agricultural futures pits. CME also closed most of its futures trading pits at this time but only the agricultural futures used the blended settlement methodology. The end of the blended settlement in 2015 could be part of the reason HFT involvement in the agricultural sector was the lowest in our sample's starting year of 2012 (see Table 1). All three agricultural contracts under consideration experienced increases in HFT activity of 100% or more during the remainder of our sample period. As a result, we use the close of the pits and the end of the blended settlement methodology as a natural experiment to identify any adverse effects this regime had on HFT involvement, as well as on the Amihud measure and bid-ask spreads.

5.2 Difference-in-difference estimation

We use a generalized, or fixed-effects difference-in-difference model to analyze the effect of the change in settlement methodology on the HFT market share in agricultural futures markets as compared to the other contracts in our sample. We then compare this differential effect with the impact of the change in settlement methodology on the two market quality metrics—Amihud and bid-ask spread. The change occurred on July 2, 2015, and we use daily data before and after the change to conduct the difference-in-difference estimation. We use the generalized form of the difference-in-difference model in order to capture the effects across multiple groups.

The generalized difference-in-differences model is run as a fixed-effects model using both market and time effects with clustered standard errors using the following equation:

$$log\left(MarketShare_{it}\right) = \alpha + \lambda_{it} BlendedSettlement_{it} + \sum_{k=1}^{K-1} \beta_k Contract_k + \sum_{t=1}^{T-1} \beta_t Day_t + \varepsilon_{st},$$

where K is the number of contracts ($K = 9$), T is the number of days in the pre- and post-change period (30, 60, or 90 days before and after the rule change), *Contract* and *Day* capture market and time fixed effects, respectively, and *BlendedSettlement* is a dummy variable equal to 1 if the contract is agricultural (i.e., corn, soybeans, or wheat) and the date is July 2, 2015 or earlier, and 0 for all other observations. In other words, *BlendedSettlement* takes on the value of 1 when the blended settlement was in effect and 0 after it ended, as well as for the unaffected markets, so that λ reflects the impact of the blended settlement and, therefore, represents the difference-in-difference estimator.

We use the natural log of HFT market share so that λ can be interpreted as a percentage change in market share. For robustness, we run the model over three different time windows: 30 days before and after the change, 60 days before and after, and 90 days before and after.

5.3 Difference-in-difference results

Table 11 summarizes the results of the difference-in-difference estimation for the HFT market share (Panel A), the Amihud price impact (Panel B), and the bid-ask spread (Panel C). The difference-in-difference coefficient on the HFT market share is negative and significant at the 5% level, and this result is consistent across the three event windows—namely, ± 30 days, ± 60 days, and ± 90 days. The magnitude and the negative sign of the difference-in-difference coefficient suggests that HFT involvement was depressed by over 35% during the blended settlement period compared to the period after its abandonment. Thus, the end of blended settlement had a profound impact on the HFTs' participation.

Panel B of Table 11 applies the same fixed-effects difference-in-difference methodology to the Amihud measure. The difference-in-difference coefficient (λ) is now not significant at the 5% level and only marginally significant (at the 10% level). The results are even more striking when we apply the same estimation method to the bid-ask spreads: λ is insignificant in all three regressions in Panel C of Table 11. Thus, the end of the blended settlement period did not affect market quality strongly (if measured by the Amihud price impact) or at all (if focusing on the traded bid-ask spread).

These results, coupled with the aforementioned significant impact on the HFT market share, suggest that while we may not be able to make strong causal statements, any impact of market quality on HFTs' market share is likely negligible in comparison with other factors that deter or attract HFTs to certain markets at various points in time, such as the change in CME's settlement methodology. Put differently, our results suggest that while we cannot rule out some feedback effect of market quality on HFT market share, this effect is small relative to the modeled impact of HFT share on market quality.

To summarize or difference-in-difference exercises, the change in settlement methodology rules on July 2, 2015, concerned three out of nine commodity futures markets in our sample. As a result, the market quality did not increase significantly (at the 5% level), yet the HFT presence did go up significantly (also at the 5% level). This implies that the significantly greater market share of HFTs was not precipitated by better market quality. This means that if anything, better market quality is not the driving force behind HFTs' incentives to modify the extent of their presence in various markets during various periods of time.

5.4 Lagged HFT market share results

Finally, as an additional way to address the potential issue of endogeneity, we re-run our main models in Table 4 using lagged values of HFT market share instead of contemporaneous values. We find that the positive and significant relationship between HFT market share and market quality is preserved when we allow for a lagged impact. These additional results are available from the authors upon request.

6 Conclusions and future research

This paper has studied the link between HFTs' presence and market quality using a unique regulatory dataset on intraday transactions and positions for nine diverse futures markets for the period 2012-2021. We find that HFTs' activity results in improved market quality, which is consistent with much of the existing literature (e.g., Carrion, 2013; Menkveld, 2013) and suggested mechanisms (e.g., Brogaard et al., 2019), and this effect is highly significant and robust. Additionally, we find a second-order adverse effect on market quality from HFTs' aggressive trading, and this finding is in line with some of the literature's evidence that HFTs may not continue to act in a market-making capacity in some segments of financial markets during stress periods (e.g., Brogaard et al., 2018). Our results suggest that the market quality improvements brought about by HFTs' market making more than offset the negative effects of HFTs' aggressive directional trading.

Subsequent research can apply a similar analysis to futures order book data if and when such data become available. Amihud and traded bid-ask spreads used in this paper provide comprehensive metrics of market quality, and futures order book data can complement this analysis with several related market quality measures, such as the quoted bid-ask spread and the order book depth.

To address potential endogeity, we took advantage of the 2015 change in CME Group's settlement rules that concerned a subset of commodity markets in our sample. While the results of our difference-in-difference estimation cannot rule out a possible feedback effect of market quality on HFTs' market share, the results suggest that market quality is not the primary determinant of the variation in HFTs' market presence either cross-sectionally or over time.

Our paper has followed the existing literature in taking as given the HFT firms' decisions to change the extent of their presence in various markets. It is unlikely that there is a substantial feedback loop from improved market quality to increased HFTs' participation (e.g., narrow bid-ask spreads per se may not be an attractive feature of a market from the HFTs' standpoint), especially in light of our difference-in-difference results. Nevertheless, it would be interesting to study which factors are behind the endogenous decisions of HFTs to enter—and expand their presence in—certain markets at certain points in time. Such an investigation would likely necessitate a look into how and to what extent various market structures allow HFTs to leverage their comparative advantage in trading technology. We leave these important questions to future research.

Appendix: Variables and descriptions

Table A1: Regressors and descriptions

Regressor	Description
HFT market share	The market share of HFTs as a percent of the total volume traded (main regressor).
Percent of aggressive volume	The percentage of volume traded by HFTs executed as aggressive order types. This could be thought of as a proxy for market-making behavior: less aggressive (more passive) order placement is sometimes associated with market-making behavior, whereas more aggressive trading is often use when HFTs attempt to lay off an accumulated position.
Ordered-to-traded volume ratio	The ratio of orders entered to volume actually traded. This could be construed as a proxy for market dislocation: if traders are having trouble getting their orders filled, they will have to accept partial fills and place new orders.
Volume per trade	Total traded volume divided by the number of trades placed. This is a rough measure of market liquidity since more liquid markets can handle larger trade sizes.
Directional ratio	The absolute ratio of buy volume to sell volume (this is agnostic for long and short, it just measures how far away from flat HFTs are where flat = 1).
Total stop volume	The total volume executed as stop orders for the entire market. Stop orders are often triggered in times of stress so higher stop volumes are used here as an indicator of market stress.
Total trading accounts	The total number of accounts actively trading in the market. More accounts tend to trade during times of stress, so this is used as another proxy for market stress.
Lagged volatility	Annualized 21-day volatility, lagged one period, using day-over-day price changes. This is arguably not endogenous to the model since it is the overall volatility environment of each market, not including any trades from the current trading date.
Total volume and open interest	Total volume and open interest (OI) in all traded contract expirations, not just the active contract. These control for the size differences across markets.
OI percent active and volume percent active	The percentage of total volume and open interest in the active contract. This is a measure of how much trading activity is spread out in the futures curve vs. concentrated in the active contract. This controls for both the amount of spread trading and the roll period.
Macro control variables	Credit spread (Baa corporate bond yield relative to 10-year treasury constant maturity), TED Spread, and VIX (all retrieved from the St. Louis FRED).
Season	Season (rather than quarter) to account for seasonality, especially in agricultural contracts. [Dec, Jan, Feb] = 1; [Mar, Apr, May] = 2; [Jun, Jul, Aug] = 3, [Sep, Oct, Nov] = 4.
Year	Calendar year to control for year fixed effects.

Table A2: Market Quality Metrics with Time Frame and Outlier Variations

Market quality metric	Description
Amihud_pct_all_ino	Amihud with settlement and non-settlement periods, including outliers (used in main regressions)
Amihud_pct_all_exo	Amihud with settlement and non-settlement periods, excluding outliers
Amihud_pct_ns_ino	Amihud with non-settlement period only, including outliers
Amihud_pct_ns_exo	Amihud with non-settlement period only, excluding outliers
Amihud_pct_s_ino	Amihud with settlement period only, including outliers
TBAS_simple_all	TBAS calculated with simple average of all volume-weighted minute averages, including outliers
TBAS_weighted_all	TBAS calculated with volume-weighted average of all volume-weighted minute averages, including outliers
TBAS_simple_adj	TBAS calculated with simple average of all volume-weighted minute averages, excluding outliers
TBAS_weighted_adj	TBAS calculated with volume-weighted average of all volume-weighted minute averages, excluding outliers (used in main regressions)

References

Amihud, Y. (2002). "Illiquidity and stock returns: Cross-section and time-series effects", *Journal of Financial Markets*, 5(1), 31–56.

Aquilina, M., Budish, E. and O'Neill, P. (2022). "Quantifying the high-frequency trading 'arms race'", *Quarterly Journal of Economics*, 137(1), 493–564.

Baker, L., McPhail, L. and Tuckman, B. (2020). "The liquidity hierarchy in the US Treasury cash and futures market", *Journal of Fixed Income*, 30(1), 90–99.

Biais, B. and Foucault, T. (2014). "HFT and market quality", *Bankers, Markets & Investors*, 128, 5–19.

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

Brogaard, J., Hendershott, T. and Riordan, R. (2014). "High frequency trading and price discovery", *Review of Financial Studies*, 27(8), 2267–2306.

Brogaard, J., Hendershott, T. and Riordan, R. (2017). "High frequency trading and the 2008 short-sale ban", *Journal of Financial Economics*, 124(1), 22–42.

Brogaard, J., Hendershott, T. and Riordan, R. (2019). "Price discovery without trading: Evidence from limit orders", *Journal of Finance*, 74(4), 1621–1658.

Budish, E., Cramton, P. and Shim, J. (2015). "The high-frequency trading arms race: Frequent batch auctions as a market design response", *Quarterly Journal of Economics*, 130(4), 1547–1621.

Carrion, A. (2013). "Very fast money: High-frequency trading on the NASDAQ", *Journal of Financial Markets*, 16(4), 680–711.

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E. and Vega, C. (2014). "Rise of the machines: Algorithmic trading in the foreign exchange market", *Journal of Finance*, 69(5), 2045–2084.

Coën, A. and de La Bruslerie, H. (2019). "The informational dimensions of the Amihud (2002) illiquidity measure: Evidence from the M&A market", *Finance Research Letters*, 29, 23–29.

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

Foley, S. and Putniņš, T.J. (2016). "Should we be afraid of the dark? Dark trading and market quality", *Journal of Financial Economics*, 122(3), 456–481.

Fotak, V., Raman, V. and Yadav, P.K. (2014). "Fails-to-deliver, short selling, and market quality", *Journal of Financial Economics*, 114(3), 493–516.

Goldstein, M.A., Kwan, A. and Philip, R. (2021). "High-frequency trading strategies", working paper, available at https://ssrn.com/abstract=2973019.

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

Han, J., Khapko, M. and Kyle, A.S. (2014). "Liquidity with high-frequency market making", Swedish House of Finance Research Paper No. 14-06, available at https://ssrn.com/abstract=2416396.

Hendershott, T. and Jones, C.M. (2005). "Island goes dark: Transparency, fragmentation, and regulation", *Review of Financial Studies*, 18(3), 743–793.

Hendershott, T., Jones, C.M. and Menkveld, A.J. (2011). "Does algorithmic trading improve liquidity?", *Journal of Finance*, 66(1), 1–33.

Hoffmann, P. (2014). "A dynamic limit order market with fast and slow traders", *Journal of Financial Economics*, 113(1), 156–169.

Kirilenko, A.A., Kyle, A.S., Samadi, M. and Tuzun, T. (2017). "The flash crash: High-frequency trading in an electronic market", *Journal of Finance*, 72(3), 967–998.

Korajczyk, R.A. and Murphy, D. (2019). "High-frequency market making to large institutional trades", *Review of Financial Studies*, 32(3), 1034–1067.

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

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

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

Menkveld, A.J. and Zoican, M.A. (2017). "Need for speed? exchange latency and liquidity", *Review of Financial Studies*, 30(4), 1188–1228.

Nie, J. (2019). "High-frequency price discovery and price efficiency on interest rate futures", *Journal of Futures Markets*, 39(11), 1394–1434.

O'Hara, M. (2015). "High frequency market microstructure", *Journal of Financial Economics*, 116(2), 257–270.

Raman, V., Robe, M.A. and Yadav, P.K. (2020). "The third dimension of financialization: Intraday institutional financial traders and commodity market quality", working paper, available at https://ssrn.com/abstract=3068248.

Stoll, H.R. (2000). "Friction", *Journal of Finance*, 55(4), 1479–1514.

Van Kervel, V. and Menkveld, A.J. (2019). "High-frequency trading around large institutional orders", *Journal of Finance*, 74(3), 1091–1137.

Figure 1: Market Share by Account Type, Quarterly Averages, January 2012 – August 2021

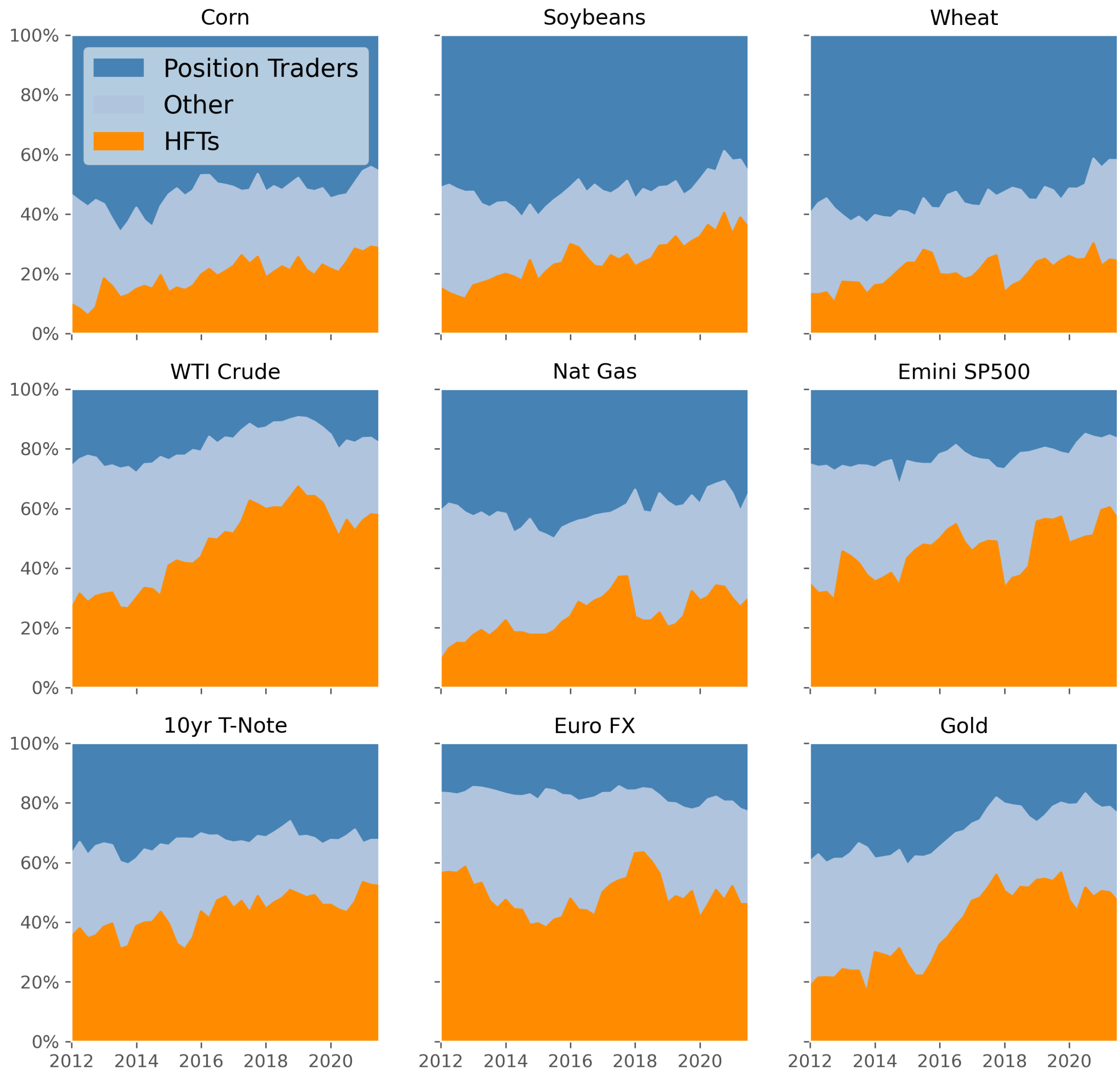

Figure 2: Daily Amihud Price Impact with a 21-Day Moving Average, January 2012 – August 2021

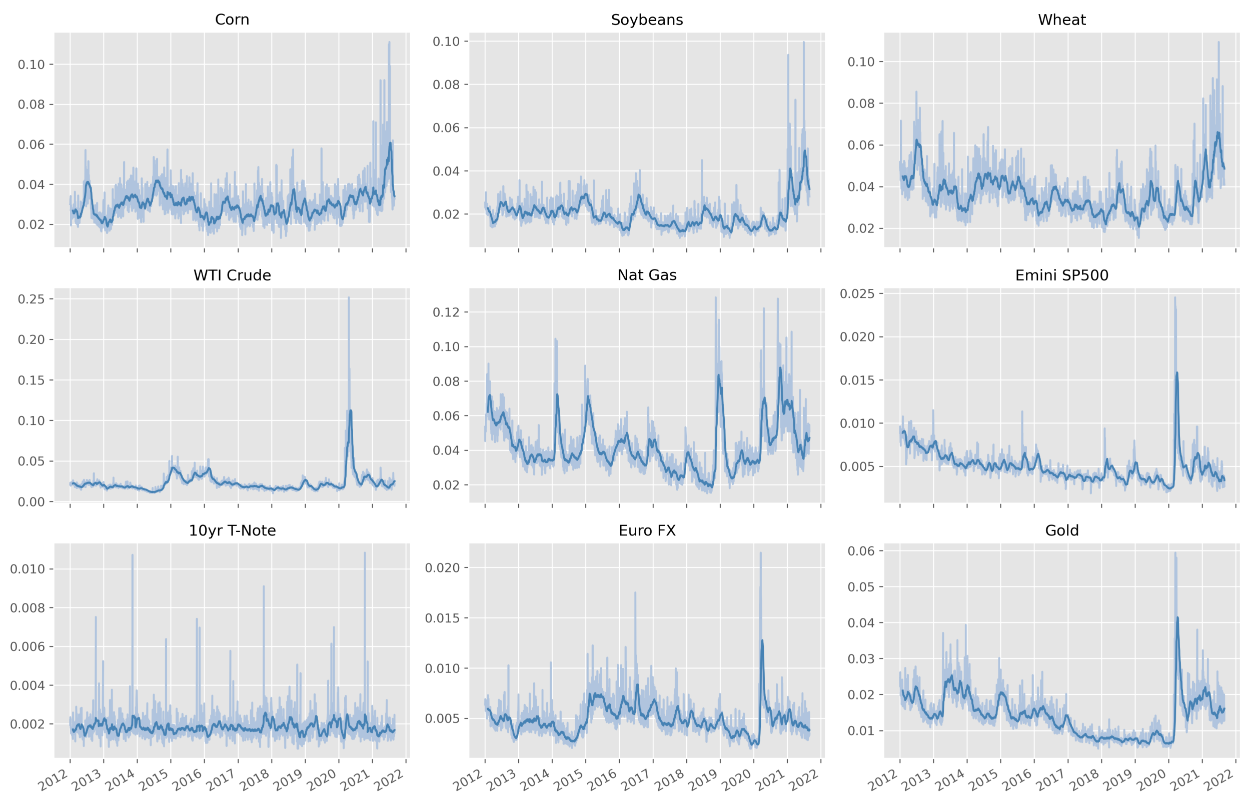

Figure 3: Daily Traded Bid-Ask Spread with a 21-Day Moving Average, January 2012 – August 2021

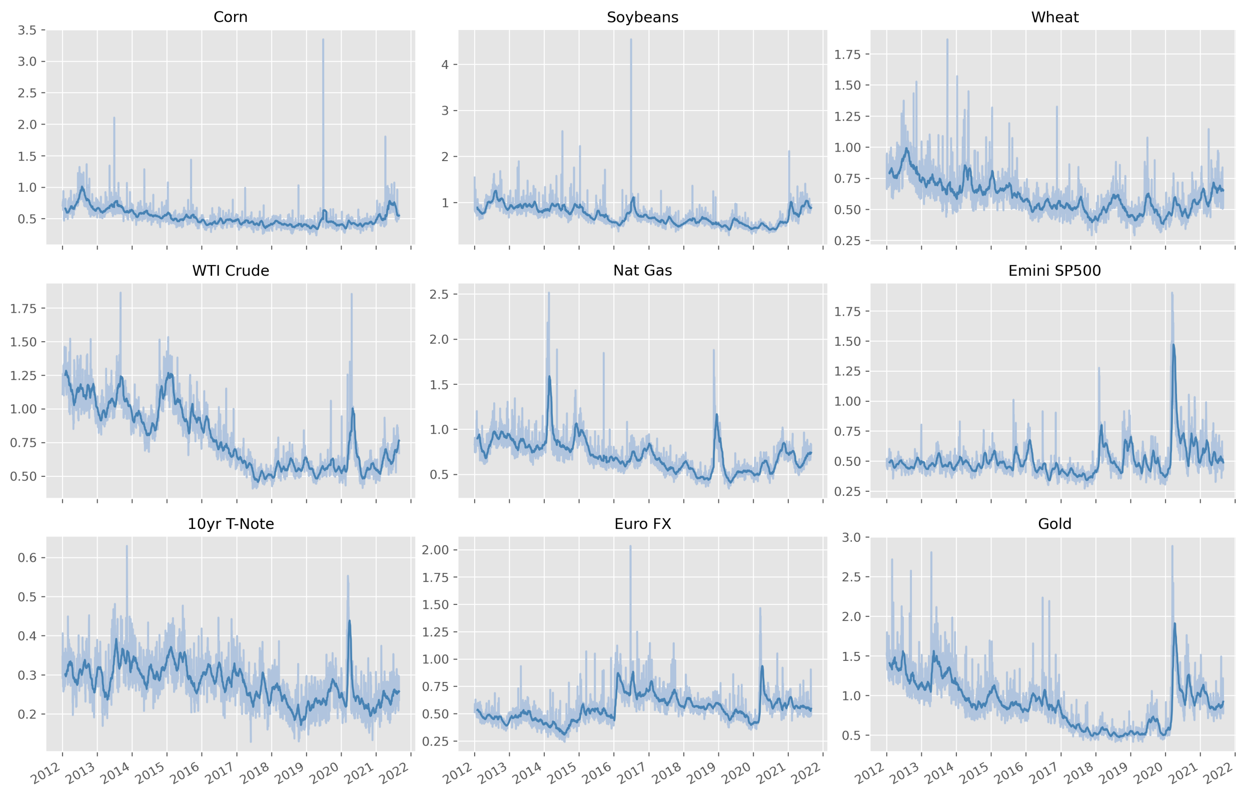

Table 1: HFT Market Share Summary, Select Time Periods

Sector	Commodity	HFT market share		Percent change	Average over entire sample
		2012	2020		
Agriculture	Corn	8.0%	23.9%	197.3%	19.4%
Agriculture	Soybeans	13.1%	36.0%	175.8%	24.8%
Agriculture	Wheat	12.5%	26.6%	112.0%	20.6%
Currencies	Euro FX	57.0%	46.2%	−18.9%	49.6%
Energy	Natural gas	12.9%	31.7%	145.8%	24.5%
Energy	WTI crude oil	29.1%	53.9%	85.2%	52.1%
Financials	10-year T-note	35.8%	45.2%	26.1%	43.5%
Financials	E-mini S&P500	31.8%	49.7%	56.3%	45.0%
Metals	Gold	20.6%	48.1%	133.4%	42.1%

Table 2: Summary Statistics

This table summarizes the data pooled across all contracts for the entire sample. Summary statistics for each market separately are available from the authors upon request (and will be made available in an Internet Appendix). The table summarizes all variables used in the analysis, including all alternative methods of calculating the market quality metrics. Please refer to the Appendix for the definitions of variables.

Variable	Obs.	Mean	St. dev.	Min	25th	Median	75th	Max	Skewness	Kurtosis
Amihud_pct_all_ino	21591	0.0196	0.0161	0.0007	0.0051	0.0175	0.0293	0.2517	1.3752	8.0787
Amihud_pct_all_exo	21591	0.0193	0.0158	0.0007	0.0050	0.0174	0.0292	0.2517	1.3400	8.1862
Amihud_pct_ns_ino	21591	0.0195	0.0161	0.0007	0.0050	0.0174	0.0289	0.2452	1.4217	8.1475
Amihud_pct_ns_exo	21591	0.0192	0.0158	0.0007	0.0050	0.0173	0.0288	0.2452	1.3873	8.2696
Amihud_pct_s_ino	21591	0.0208	0.0175	0	0.0063	0.0165	0.0306	0.4798	2.0510	26.9636
TBAS_simple_all	21591	0.5432	0.1377	0.0616	0.4450	0.5241	0.6205	2.0357	1.1256	6.5040
TBAS_weighted_all	21591	0.4348	0.1933	-2.3915	0.3152	0.4019	0.5166	3.5142	0.8492	15.6184
TBAS_simple_adj	21591	0.4919	0.1693	0.0317	0.3763	0.4589	0.5555	2.3265	1.8444	9.7100
TBAS_weighted_adj	21591	0.6164	0.2645	0.1284	0.4470	0.5604	0.7461	4.5427	1.4832	8.9762
HFT market share	21591	0.3457	0.1470	0.0152	0.2217	0.3255	0.4730	0.7482	0.2069	1.9591
avg volume per minute	21591	326	433	16	72	122	333	3997	2.0389	7.4022
pct aggressive volume	21591	0.4823	0.0766	0.2164	0.4393	0.4732	0.5135	0.9927	1.0821	6.3851
ordered-to-traded volume	21591	1.5106	0.6086	1.0214	1.1579	1.3109	1.6142	8.5493	3.2236	17.2911
volume per trade	21591	2.2561	1.8935	1.0646	1.2845	1.4683	1.9879	14.0061	2.5479	9.4247
pct directional	21591	-0.0003	0.0059	-0.0611	-0.0013	0.0000	0.0010	0.0879	0.0224	22.1947
total stop volume	21591	8059.54	11208.71	23	1741	3745	9469	175882	3.2365	18.7188
total active accounts	21591	3061	2528	555	1544	2267	3194	18510	2.1425	6.9757
total volume	21591	635046	703140	26157	190346	324238	924067	9025630	2.5482	13.7628
total open interest	21591	1404540	1017427	192845	498110	1194202	2079404	4704423	0.8495	2.6583
pct OI in most active	21591	0.5689	0.3176	0.0125	0.2811	0.5242	0.9608	1	0.0698	1.6251
pct volume in most active	21591	0.7016	0.2365	0.2173	0.4933	0.6303	0.9800	1	0.0861	1.4100
lagged volatility	21591	0.1974	0.1661	0.0166	0.0892	0.1651	0.2567	2.6555	4.0617	40.7113
TED spread	21591	0.0029	0.0015	0.0006	0.0020	0.0025	0.0037	0.0142	2.5732	15.7783
VIX	21591	17.0135	6.9011	9.14	12.93	15.2	18.84	82.69	3.4343	22.5053
credit spread	21591	2.5072	0.4733	1.56	2.16	2.43	2.85	4.31	0.4469	2.5657
season	21591	2.49	1.09	1	2	2	3	4	0.0085	1.7042
year	21591	2016.35	2.79	2012	2014	2016	2019	2021	0.0216	1.8044

Table 3: HFT Market Share and Market Quality: Univariate Analysis

The models are: (1) pooled OLS, (2) pooled OLS with heteroskedasticity-robust standard errors, (3) fixed effects model, (4) fixed effects model with heteroskedasticity-robust standard errors, (5) fixed effects model with AR(1) disturbances, (6) random effects model, (7) random effects model with heteroskedasticity-robust standard errors, (8) random effects model with AR(1) disturbances. The reported R-squared are adjusted R-squared for Models (1) through (5) (i.e., OLS and fixed effects) and overall R-squared for Models (6) through (8) (i.e., random effects). ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: Amihud Regressions

Variable	(1) OLS	(2) OLS robust	(3) FE	(4) FE robust	(5) FE AR(1)	(6) RE	(7) RE robust	(8) RE AR(1)
HFT market share	−0.062***	−0.062***	−0.011***	−0.011**	−0.004***	−0.011***	−0.011**	−0.013***
Constant	0.041***	0.041***	0.023***	0.023***	0.025***	0.023***	0.023***	0.024***
R-squared	0.318	0.318	0.012	0.012	0.001	0.318	0.318	0.318
Obs.	21591	21591	21591	21591	21582	21591	21591	21591

Panel B: Bid-Ask Spread Regressions

Variable	(1) OLS	(2) OLS robust	(3) FE	(4) FE robust	(5) FE AR(1)	(6) RE	(7) RE robust	(8) RE AR(1)
HFT market share	−0.537***	−0.537***	−1.137***	−1.137***	−0.249***	−1.136***	−1.136***	−1.048***
Constant	0.802***	0.802***	1.009***	1.009***	0.819***	1.009***	1.009***	0.978***
R-squared	0.089	0.089	0.266	0.266	0.006	0.089	0.089	0.089
Obs.	21591	21591	21591	21591	21582	21591	21591	21591

Table 4: HFT Market Share and Market Quality: Multivariate Analysis

The models are: (1) pooled OLS, (2) pooled OLS with heteroskedasticity-robust standard errors, (3) fixed effects model, (4) fixed effects model with heteroskedasticity-robust standard errors, (5) fixed effects model with AR(1) disturbances, (6) random effects model, (7) random effects model with heteroskedasticity-robust standard errors, (8) random effects model with AR(1) disturbances. The reported R-squared are adjusted R-squared for Models (1) through (5) (i.e., OLS and fixed effects) and overall R-squared for Models (6) through (8) (i.e., random effects). ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: Amihud Regressions

Variable	(1) OLS	(2) OLS robust	(3) FE	(4) FE robust	(5) FE AR(1)	(6) RE	(7) RE robust	(8) RE AR(1)
HFT market share	−0.0290***	−0.0290***	−0.0076***	−0.0076**	−0.0115***	−0.0290***	−0.0290***	−0.0089***
pct aggressive volume	0.0050***	0.0050***	0.0024***	0.0024	−0.0023***	0.0050***	0.005	−0.0068***
ordered-to-traded vol.	0.0001	0.0001	0.0013***	0.0013**	0.0022***	0.0001	0.0001	0.0018***
volume per trade	0.0008***	0.0008***	0.0012***	0.0012***	0.0003***	0.0008***	0.0008	0.0003***
pct directional	0.0063	0.0063	0.0171**	0.0171	−0.004	0.0063	0.0063	−0.0035
total stop volume	0.0000	0.0000	−0.0000***	0.0000	0.0000	0.0000	0.0000	0.0000
total active accounts	0.0000***	0.0000***	0.0000***	0.0000**	0.0000***	0.0000***	0.0000	0.0000***
total volume	−0.0000***	−0.0000***	0.0000	0.0000	0.0000	−0.0000***	−0.0000**	0.0000
total open interest	−0.0000***	−0.0000***	−0.0000***	0.0000	−0.0000***	−0.0000***	−0.0000*	0.0000
pct OI in most active	0.0005	0.0005	0.0040***	0.004	−0.0004	0.0005	0.0005	0.0009
pct vol. in most active	−0.0208***	−0.0208***	−0.0094***	−0.0094**	−0.0045***	−0.0208***	−0.0208***	−0.0061***
lagged volatility	0.0504***	0.0504***	0.0416***	0.0416***	0.0355***	0.0504***	0.0504***	0.0340***
TED spread	−0.0665	−0.0665	0.0939**	0.0939	−0.0001	−0.0665	−0.0665	0.1727**
VIX	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
credit spread	−0.0019***	−0.0019***	−0.0010***	−0.0010**	0.0043***	−0.0019***	−0.0019***	0.0001
season								
2	−0.0008***	−0.0008***	−0.0009***	−0.0009**	−0.0010***	−0.0008***	−0.0008***	−0.0009***
3	−0.0004***	−0.0004***	−0.0003**	−0.0003	−0.0002	−0.0004***	−0.0004	−0.0001
4	0.0003**	0.0003**	−0.0003**	−0.0003	−0.0001	0.0003**	0.0003	−0.0002
year								
2013	−0.0009***	−0.0009***	−0.0015***	−0.0015	0.0008*	−0.0009***	−0.0009	−0.0016***
2014	0.0002	0.0002	−0.0002	−0.0002	0.0044***	0.0002	0.0002	0.0004
2015	0.0003	0.0003	−0.0008***	−0.0008	0.0012***	0.0003	0.0003	−0.0008*
2016	0.0004*	0.0004*	−0.0028***	−0.0028	−0.0005	0.0004*	0.0004	−0.0030***
2017	0.0003	0.0003	−0.0041***	−0.0041	0.0017***	0.0003	0.0003	−0.0038***
2018	−0.0017***	−0.0017***	−0.0056***	−0.0056**	0.0015***	−0.0017***	−0.0017	−0.0050***
2019	−0.0003	−0.0003	−0.0048***	−0.0048*	0.0004	−0.0003	−0.0003	−0.0045***
2020	0.0009***	0.0009***	−0.0027***	−0.0027	0.0013**	0.0009***	0.0009	−0.0021***
2021	0.0065***	0.0065***	0.0030***	0.003	0.0096***	0.0065***	0.0065***	0.0039***
Constant	0.0358***	0.0358***	0.0147***	0.0147**	0.0036***	0.0358***	0.0358***	0.0179***
R-squared	0.7966	0.7966	0.4852	0.4854	0.2144	0.7968	0.7968	0.6162
Obs.	21591	21591	21591	21591	21582	21591	21591	21591

Table 4 continued

Panel B: Bid-Ask Spread Regressions

Variable	(1) OLS	(2) OLS robust	(3) FE	(4) FE robust	(5) FE AR(1)	(6) RE	(7) RE robust	(8) RE AR(1)
HFT market share	−0.3936***	−0.3936***	−0.9375***	−0.9375***	−0.7799***	−0.3936***	−0.3936*	−0.8148***
pct aggressive volume	0.1346***	0.1346***	0.4078***	0.4078**	0.2599***	0.1346***	0.1346	0.1463***
ordered-to-traded vol.	−0.0380***	−0.0380***	0.0014	0.0014	0.0174***	−0.0380***	−0.038	0.0051
volume per trade	−0.0207***	−0.0207***	−0.0163***	−0.0163	−0.0167***	−0.0207***	−0.0207	−0.0170***
pct directional	0.0237	0.0237	−0.0098	−0.0098	−0.2023*	0.0237	0.0237	−0.1925*
total stop volume	0.0000***	0.0000***	0.0000***	0.0000**	0.0000***	0.0000***	0.0000**	0.0000***
total active accounts	0.0000	0.0000	0.0000***	0.0000	0.0000***	0.0000	0.0000	0.0001***
total volume	−0.0000***	−0.0000***	−0.0000***	0.0000	0.0000	−0.0000***	0.0000	−0.0000**
total open interest	−0.0000***	−0.0000***	0.0000***	0.0000	0.0000	−0.0000***	0.0000	0.0000**
pct OI in most active	−0.6773***	−0.6773***	−0.1003***	−0.1003	−0.1580***	−0.6773***	−0.6773***	−0.1373***
pct vol. in most active	0.7868***	0.7868***	−0.0442**	−0.0442	0.2074***	0.7868***	0.7868**	0.1009***
lagged volatility	0.2479***	0.2479***	0.3022***	0.3022*	0.2643***	0.2479***	0.2479**	0.2636***
TED spread	6.0307***	6.0307***	1.7428*	1.7428	0.5183	6.0307***	6.0307*	4.7351**
VIX	0.0043***	0.0043***	0.0039***	0.0039***	0.0011***	0.0043***	0.0043***	0.0024***
credit spread	−0.0280***	−0.0280***	−0.0161***	−0.0161	0.1368***	−0.0280***	−0.028	0.0097
season								
2	0.0062*	0.0062*	−0.0032	−0.0032	−0.0098*	0.0062*	0.0062	−0.0061
3	0.0206***	0.0206***	0.0049*	0.0049	0.0045	0.0206***	0.0206	0.0062
4	−0.0025	−0.0025	0.001	0.001	0.0016	−0.0025	−0.0025	−0.0012
year								
2013	0.0158***	0.0158**	0.0161***	0.0161	0.0824***	0.0158***	0.0158	0.0158
2014	−0.0402***	−0.0402***	−0.0337***	−0.0337*	0.0990***	−0.0402***	−0.0402	−0.0176
2015	−0.0831***	−0.0831***	−0.0652***	−0.0652**	−0.0105	−0.0831***	−0.0831**	−0.0647***
2016	−0.0766***	−0.0766***	−0.0296***	−0.0296	0.0068	−0.0766***	−0.0766	−0.0525***
2017	−0.1121***	−0.1121***	−0.0468***	−0.0468	0.0824***	−0.1121***	−0.1121	−0.0635***
2018	−0.1513***	−0.1513***	−0.1145***	−0.1145	0.0746***	−0.1513***	−0.1513	−0.1081***
2019	−0.1540***	−0.1540***	−0.1061***	−0.1061	0.0202	−0.1540***	−0.1540*	−0.1081***
2020	−0.1178***	−0.1178***	−0.0872***	−0.0872	0.0147	−0.1178***	−0.1178	−0.0742***
2021	−0.0215**	−0.0215*	0.0154*	0.0154	0.1991***	−0.0215**	−0.0215	0.0435**
Constant	0.6724***	0.6724***	0.6579***	0.6579**	0.1041***	0.6724***	0.6724**	0.5895***
R-squared	0.5793	0.5793	0.479	0.4792	0.3185	0.5799	0.5799	0.2206
Obs.	21591	21591	21591	21591	21582	21591	21591	21591

Table 5: Robustness: Panel-Corrected and Clustered Standard Errors

The models are: (1) pooled OLS with heteroskedasticity-robust standard errors, (2) fixed effects model, (3) linear regression with panel-corrected standard errors and no autocorrelation within panels, (4) panel-corrected standard errors and first-order autocorrelation within panels, (5) panel-corrected standard errors and panel-specific first-order autocorrelation, (6) first-order autocorrelation within panels with heteroskedastic (but not contemporaneously correlated) panel-level disturbances, (7) fixed effects model with standard errors clustered at the market level, (8) random effects model with standard errors clustered at the market level. The reported R-squared are adjusted R-squared for Models (1), (2) and (7) (i.e., OLS and fixed effects) and overall R-squared for Model (8) (i.e., random effects). ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: Amihud Regressions

Variable	(1) OLS robust	(2) FE	(3) PCSE	(4) PCSE AR(1)	(5) PCSE PSAR(1)	(6) PCSE AR(1) het	(7) FE clustered	(8) RE clustered
HFT market share	−0.0290***	−0.0076***	−0.0290***	−0.0293***	−0.0309***	−0.0293***	−0.0076**	−0.0290***
pct aggressive volume	0.0050***	0.0024***	0.0050***	−0.0031***	0.0004	−0.0031***	0.0024	0.005
ordered-to-traded vol.	0.0001	0.0013***	0.0001	0.0008***	0.0006***	0.0008***	0.0013**	0.0001
volume per trade	0.0008***	0.0012***	0.0008***	0.0006***	0.0006***	0.0006***	0.0012***	0.0008
pct directional	0.0063	0.0171**	0.0063	−0.0065	−0.0039	−0.0065	0.0171	0.0063
total stop volume	0.0000	−0.0000***	0.0000	0.0000***	0.0000***	0.0000***	0.0000	0.0000
total active accounts	0.0000***	0.0000***	0.0000***	0.0000**	0.0000	0.0000**	0.0000**	0.0000
total volume	−0.0000***	0.0000	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000	−0.0000**
total open interest	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000	−0.0000*
pct OI in most active	0.0005	0.0040***	0.0005	−0.0023***	0.0004	−0.0023***	0.004	0.0005
pct vol. in most active	−0.0208***	−0.0094***	−0.0208***	−0.0163***	−0.0195***	−0.0163***	−0.0094**	−0.0208***
lagged volatility	0.0504***	0.0416***	0.0504***	0.0456***	0.0479***	0.0456***	0.0416***	0.0504***
TED spread	−0.0665	0.0939**	−0.0665	−0.0352	0.0117	−0.0352	0.0939	−0.0665
VIX	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
credit spread	−0.0019***	−0.0010***	−0.0019***	−0.0010*	−0.0014***	−0.0010**	−0.0010**	−0.0019***
season								
2	−0.0008***	−0.0009***	−0.0008***	−0.0009***	−0.0011***	−0.0009***	−0.0009**	−0.0008***
3	−0.0004***	−0.0003**	−0.0004***	−0.0003	−0.0004	−0.0003	−0.0003	−0.0004
4	0.0003**	−0.0003**	0.0003**	0.0002	0.0004*	0.0002	−0.0003	0.0003
year								
2013	−0.0009***	−0.0015***	−0.0009***	−0.0010*	−0.0007*	−0.0010**	−0.0015	−0.0009
2014	0.0002	−0.0002	0.0002	0.0009	0.0015***	0.0009*	−0.0002	0.0002
2015	0.0003	−0.0008***	0.0003	0.0003	0.0007*	0.0003	−0.0008	0.0003
2016	0.0004*	−0.0028***	0.0004*	0.0002	0.0002	0.0002	−0.0028	0.0004
2017	0.0003	−0.0041***	0.0003	0.0003	0.0005	0.0003	−0.0041	0.0003
2018	−0.0017***	−0.0056***	−0.0017***	−0.0015*	−0.0015**	−0.0015**	−0.0056**	−0.0017
2019	−0.0003	−0.0048***	−0.0003	−0.0003	−0.0002	−0.0003	−0.0048*	−0.0003
2020	0.0009***	−0.0027***	0.0009***	0.0012*	0.0019***	0.0012**	−0.0027	0.0009
2021	0.0065***	0.0030***	0.0065***	0.0070***	0.0076***	0.0070***	0.003	0.0065***
Constant	0.0358***	0.0147***	0.0358***	0.0366***	0.0374***	0.0366***	0.0147**	0.0358***
R-squared	0.7966	0.4852					0.4854	0.7968
Obs.	21591	21591	21591	21591	21591	21591	21591	21591

Table 5 continued

Panel B: Bid-Ask Spread Regressions

Variable	(1) OLS robust	(2) FE	(3) PCSE	(4) PCSE AR(1)	(5) PCSE PSAR(1)	(6) PCSE AR(1) het	(7) FE clustered	(8) RE clustered
HFT market share	−0.3936***	−0.9375***	−0.3936***	−0.4209***	−0.3361***	−0.4209***	−0.9375***	−0.3936*
pct aggressive volume	0.1346***	0.4078***	0.1346***	−0.0005	0.0571***	−0.0005	0.4078**	0.1346
ordered-to-traded vol.	−0.0380***	0.0014	−0.0380***	−0.0293***	−0.0258***	−0.0293***	0.0014	−0.038
volume per trade	−0.0207***	−0.0163***	−0.0207***	−0.0155***	−0.0159***	−0.0155***	−0.0163	−0.0207
pct directional	0.0237	−0.0098	0.0237	−0.1829	−0.1542	−0.1829	−0.0098	0.0237
total stop volume	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000**	0.0000**
total active accounts	0.0000	0.0000***	0.0000	0.0000***	0.0000***	0.0000***	0.0000	0.0000
total volume	−0.0000***	−0.0000***	−0.0000***	0.0000	0.0000	0.0000	0.0000	0.0000
total open interest	−0.0000***	0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000	0.0000
pct OI in most active	−0.6773***	−0.1003***	−0.6773***	−0.3745***	−0.4432***	−0.3745***	−0.1003	−0.6773***
pct vol. in most active	0.7868***	−0.0442**	0.7868***	0.3771***	0.4968***	0.3771***	−0.0442	0.7868**
lagged volatility	0.2479***	0.3022***	0.2479***	0.2639***	0.3102***	0.2639***	0.3022*	0.2479**
TED spread	6.0307***	1.7428*	6.0307***	6.6330***	5.7251***	6.6330***	1.7428	6.0307*
VIX	0.0043***	0.0039***	0.0043***	0.0030***	0.0033***	0.0030***	0.0039***	0.0043***
credit spread	−0.0280***	−0.0161***	−0.0280***	−0.0067	−0.0252***	−0.0067	−0.0161	−0.028
season								
2	0.0062*	−0.0032	0.0062**	0.0002	0.0078*	0.0002	−0.0032	0.0062
3	0.0206***	0.0049*	0.0206***	0.0155**	0.0254***	0.0155***	0.0049	0.0206
4	−0.0025	0.001	−0.0025	−0.0008	−0.0013	−0.0008	0.001	−0.0025
year								
2013	0.0158**	0.0161***	0.0158***	0.0112	0.0046	0.0112	0.0161	0.0158
2014	−0.0402***	−0.0337***	−0.0402***	−0.0287**	−0.0431***	−0.0287**	−0.0337*	−0.0402
2015	−0.0831***	−0.0652***	−0.0831***	−0.0812***	−0.1043***	−0.0812***	−0.0652**	−0.0831**
2016	−0.0766***	−0.0296***	−0.0766***	−0.0833***	−0.1095***	−0.0833***	−0.0296	−0.0766
2017	−0.1121***	−0.0468***	−0.1121***	−0.1044***	−0.1455***	−0.1044***	−0.0468	−0.1121
2018	−0.1513***	−0.1145***	−0.1513***	−0.1275***	−0.1700***	−0.1275***	−0.1145	−0.1513
2019	−0.1540***	−0.1061***	−0.1540***	−0.1376***	−0.1774***	−0.1376***	−0.1061	−0.1540*
2020	−0.1178***	−0.0872***	−0.1178***	−0.0979***	−0.1443***	−0.0979***	−0.0872	−0.1178
2021	−0.0215*	0.0154*	−0.0215**	0.0127	−0.0259*	0.0127	0.0154	−0.0215
Constant	0.6724***	0.6579***	0.6724***	0.7902***	0.7478***	0.7902***	0.6579**	0.6724**
R-squared	0.5793	0.479					0.4792	0.5799
Obs.	21591	21591	21591	21591	21591	21591	21591	21591

Table 6: Robustness: Panel Data Models Estimated by FGLS

The models are: (1) pooled OLS with heteroskedasticity-robust standard errors, (2) fixed effects model, (3) homoskedastic panel data model, (4) panel data model with heteroskedasticity across panels, (5) panel data model with heteroskedasticity and cross-sectional correlation, (6) panel data model with heteroskedasticity across panels and autocorrelation within panels. ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: Amihud Regressions

Variable	(1) OLS robust	(2) FE	(3) Panel	(4) Panel het	(5) Panel corr het	(6) Panel het auto
HFT market share	−0.0290***	−0.0076***	−0.0290***	−0.0245***	−0.0198***	−0.0167***
pct aggressive volume	0.0050***	0.0024***	0.0050***	0.0072***	0.0050***	−0.0021***
ordered-to-traded vol.	0.0001	0.0013***	0.0001	0.0010***	0.0010***	0.0011***
volume per trade	0.0008***	0.0012***	0.0008***	0.0007***	0.0005***	0.0002***
pct directional	0.0063	0.0171**	0.0063	−0.0148*	−0.0123	−0.0074**
total stop volume	0.0000	−0.0000***	0.0000	−0.0000***	−0.0000**	0.0000***
total active accounts	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
total volume	−0.0000***	0.0000	−0.0000***	−0.0000***	−0.0000***	−0.0000***
total open interest	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***
pct OI in most active	0.0005	0.0040***	0.0005	−0.0034***	−0.0029***	−0.0028***
pct vol. in most active	−0.0208***	−0.0094***	−0.0208***	−0.0147***	−0.0143***	−0.0069***
lagged volatility	0.0504***	0.0416***	0.0504***	0.0489***	0.0473***	0.0387***
TED spread	−0.0665	0.0939**	−0.0665	−0.2260***	−0.1782***	−0.0665
VIX	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
credit spread	−0.0019***	−0.0010***	−0.0019***	−0.0018***	−0.0013***	−0.0009***
season						
2	−0.0008***	−0.0009***	−0.0008***	−0.0005***	−0.0004***	−0.0002
3	−0.0004***	−0.0003**	−0.0004***	−0.0002	−0.0002	−0.0001
4	0.0003**	−0.0003**	0.0003**	0.0000	0.0000	−0.0001
year						
2013	−0.0009***	−0.0015***	−0.0009***	−0.0006***	−0.0009***	−0.0003
2014	0.0002	−0.0002	0.0002	−0.0003	−0.0004*	−0.0006*
2015	0.0003	−0.0008***	0.0003	−0.0003*	−0.0011***	−0.0004
2016	0.0004*	−0.0028***	0.0004*	0.0014***	0.0002	0.0007**
2017	0.0003	−0.0041***	0.0003	0.0013***	−0.0004	0.0008**
2018	−0.0017***	−0.0056***	−0.0017***	−0.0003	−0.0018***	−0.0005
2019	−0.0003	−0.0048***	−0.0003	0.0008***	−0.0008***	0.0001
2020	0.0009***	−0.0027***	0.0009***	0.0005**	−0.0007***	−0.0003
2021	0.0065***	0.0030***	0.0065***	0.0038***	0.0033***	0.0023***
Constant	0.0358***	0.0147***	0.0358***	0.0284***	0.0277***	0.0224***
Adjusted R-squared	0.7966	0.4852				
Obs.	21591	21591	21591	21591	21591	21591

Table 6 continued

Panel B: Bid-Ask Spread Regressions

Variable	(1) OLS robust	(2) FE	(3) Panel	(4) Panel het	(5) Panel corr het	(6) Panel het auto
HFT market share	−0.3936***	−0.9375***	−0.3936***	−0.2576***	−0.2091***	−0.4155***
pct aggressive volume	0.1346***	0.4078***	0.1346***	0.1743***	0.0514***	−0.0206
ordered-to-traded vol.	−0.0380***	0.0014	−0.0380***	−0.0350***	−0.0188***	−0.0245***
volume per trade	−0.0207***	−0.0163***	−0.0207***	−0.0266***	−0.0224***	−0.0179***
pct directional	0.0237	−0.0098	0.0237	0.0849	−0.0065	−0.1098
total stop volume	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
total active accounts	0.0000	0.0000***	0.0000	−0.0000***	0.0000	0.0000***
total volume	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000	0.0000
total open interest	−0.0000***	0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***
pct OI in most active	−0.6773***	−0.1003***	−0.6773***	−0.5400***	−0.2956***	−0.1677***
pct vol. in most active	0.7868***	−0.0442**	0.7868***	0.5903***	0.3287***	0.0848***
lagged volatility	0.2479***	0.3022***	0.2479***	0.2793***	0.3068***	0.2790***
TED spread	6.0307***	1.7428*	6.0307***	4.5151***	4.4546***	6.2656***
VIX	0.0043***	0.0039***	0.0043***	0.0041***	0.0040***	0.0025***
credit spread	−0.0280***	−0.0161***	−0.0280***	−0.0356***	−0.0374***	−0.0111
season						
2	0.0062*	−0.0032	0.0062*	0.0041	0.0016	−0.0026
3	0.0206***	0.0049*	0.0206***	0.0100***	0.0038	−0.0002
4	−0.0025	0.001	−0.0025	−0.0051	−0.0056**	−0.0037
year						
2013	0.0158**	0.0161***	0.0158***	0.0120**	0.0003	0.0181*
2014	−0.0402***	−0.0337***	−0.0402***	−0.0318***	−0.0534***	−0.0003
2015	−0.0831***	−0.0652***	−0.0831***	−0.0679***	−0.0854***	−0.0257**
2016	−0.0766***	−0.0296***	−0.0766***	−0.0670***	−0.1157***	−0.0198*
2017	−0.1121***	−0.0468***	−0.1121***	−0.0896***	−0.1590***	−0.0259*
2018	−0.1513***	−0.1145***	−0.1513***	−0.1264***	−0.1904***	−0.0346**
2019	−0.1540***	−0.1061***	−0.1540***	−0.1407***	−0.2039***	−0.0561***
2020	−0.1178***	−0.0872***	−0.1178***	−0.1232***	−0.1812***	−0.0447***
2021	−0.0215*	0.0154*	−0.0215**	−0.0523***	−0.1081***	0.0324**
Constant	0.6724***	0.6579***	0.6724***	0.6606***	0.7755***	0.8230***
Adjusted R-squared	0.5793	0.479				
Obs.	21591	21591	21591	21591	21591	21591

Table 7: HFT Market Share and Market Quality with Controls: Individual Markets Results

Coefficients are estimated using OLS with heteroskedasticity-robust standard errors. ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: Amihud Regressions

Variable	(1) 10-yr T-note	(2) Corn	(3) Emini SP500	(4) Euro FX	(5) Gold	(6) Nat. Gas	(7) Soybeans	(8) WTI Crude	(9) Wheat
HFT market share	−0.0044***	−0.0370***	−0.0057***	−0.0078***	−0.0068***	0.0092	−0.0095**	−0.0409***	0.0046
pct aggressive volume	−0.0043***	−0.0113***	−0.0026***	0.0087***	0.0141***	0.0102**	0.0036**	0.0022	0.001
ordered-to-traded vol.	0.0000	0.0002	0.0011***	0.0009***	0.0004	0.0062***	0.0003	0.0033***	0.0032***
volume per trade	0.0000	−0.001	0.0006***	−0.0013***	−0.0056***	−0.0279***	−0.0078***	−0.0023	−0.0012
pct directional	−0.0127*	0.0046	−0.0087	−0.0039	−0.001	0.0846***	0.0047	0.1545***	−0.0396*
total stop volume	0.0000	0.0000	0.0000	0.0000***	0.0000***	0.0000***	0.0000***	0.0000	0.0000
total active accounts	−0.0000***	0.0000**	0.0000	0.0000	−0.0000**	−0.0000**	0.0000***	−0.0000***	0.0000*
total volume	−0.0000***	−0.0000***	0.0000	0.0000***	0.0000	0.0000	0.0000	0.0000**	0.0000
total open interest	0.0000	−0.0000***	0.0000	0.0000	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000
pct OI in most active	0.0006***	0.0136***	0.0025***	−0.0015***	0.0005	0.0088***	−0.0015	0.0051**	0.0006
pct vol. in most active	−0.0007*	−0.0289***	−0.0004	0.0033***	−0.0033*	−0.0277***	−0.0093***	0.0188*	−0.0156***
lagged volatility	−0.0001	0.0166***	0.0015***	0.0173***	0.0315***	0.0455***	0.0232***	0.0289***	0.0390***
TED spread	0.0284**	0.1091	−0.0251	0.0316	0.5788***	0.7468***	0.1132	0.6312**	0.5700***
VIX	0.0000**	0.0001	0.0002***	0.0001***	0.0003***	0.0001***	0.0001***	0.0002***	0.0000
credit spread	−0.0002***	−0.0036***	0.0002	0.0001	−0.0013***	0.0009	−0.0013***	0.0008	−0.0008
season									
2	0.0001**	−0.0003	0.0002***	0.0001	−0.0007***	−0.0003	−0.0006**	−0.0004	0.0006
3	0.0000	0.0028***	−0.0002***	0.0001	−0.0001	−0.0024***	0.0004	−0.0002	0.0038***
4	0.0001***	0.0029***	−0.0003***	−0.0001**	−0.0004**	−0.0012*	0.0003	0.0003	0.0010**
year									
2013	0.0002**	−0.0011	−0.0002	0.0002	0.0030***	−0.0097***	−0.0021***	0.0027***	−0.0042***
2014	−0.0001	0.0068***	−0.0008***	−0.0002	0.0013***	−0.0090***	0.0005	0.0015**	−0.0005
2015	−0.0003***	0.0028***	−0.0006***	0.0009***	0.0004	−0.0054***	0.0002	0.0095***	−0.0060***
2016	0.0002***	0.0027**	−0.0009***	0.0017***	−0.0008	−0.0159***	0.0007	0.0031***	−0.0099***
2017	0.0002	0.0021	−0.0012***	0.0013***	−0.0016**	−0.0138***	−0.0024**	0.0073***	−0.0143***
2018	0.0000	0.0036**	−0.0016***	0.0011***	−0.0047***	−0.0100***	0.0001	0.0053***	−0.0152***
2019	0.0004**	0.0040**	−0.0011***	−0.0003	−0.0031***	−0.0139***	0.0001	0.0061***	−0.0131***
2020	−0.0001	0.0088***	−0.0023***	−0.0006**	0.0008	−0.0080***	0.0008	0.0084***	−0.0118***
2021	0.0003	0.0162***	−0.0010***	0.0000	0.0027***	−0.0021	0.0160***	0.0084***	0.0073***
Constant	0.0066***	0.0635***	0.0006	−0.001	0.0178***	0.0699***	0.0329***	0.0131**	0.0327***
R-squared	0.2935	0.4176	0.8328	0.682	0.8304	0.7146	0.6383	0.7801	0.562
Obs.	2399	2399	2399	2399	2399	2399	2399	2399	2399

Table 7 continued

Panel B: Bid-Ask Spread Regressions

Variable	(1) 10-yr T-note	(2) Corn	(3) Emini SP500	(4) Euro FX	(5) Gold	(6) Nat. Gas	(7) Soybeans	(8) WTI Crude	(9) Wheat
HFT market share	−0.3921***	−0.4414***	−0.5854***	−0.5226***	−0.1839*	−0.0185	−0.4800***	−0.9036***	−0.2401***
pct aggressive volume	−0.0486	0.0306	0.1120**	0.5017***	0.6724***	0.4765***	0.032	0.4802***	0.1920***
ordered-to-traded vol.	−0.0201***	0.0015	0.0680***	0.0004	−0.0087	0.0352***	−0.0075	0.0369***	0.0225
volume per trade	−0.0079***	−0.0679***	−0.0073**	−0.1153***	−0.2480***	−0.3066***	−0.2075***	−0.1602***	−0.0493
pct directional	−0.1826	0.1967	0.5005	−0.9534	−1.1329	0.4764*	−0.1444	0.9561**	−0.5153**
total stop volume	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0001***	0.0000***	0.0000***	0.0000***
total active accounts	0.0000***	0.0001***	−0.0000***	0.0000	0.0000**	0.0000	0.0001***	0.0000	0.0001***
total volume	0.0000	−0.0000**	0.0000***	0.0000***	−0.0000***	0.0000	0.0000	0.0000	0.0000***
total open interest	−0.0000***	−0.0000***	0.0000***	0.0000	−0.0000***	0.0000	0.0000	0.0000	−0.0000***
pct OI in most active	0.0365***	−0.0556**	0.1192***	−0.0416	0.0108	−0.0582	−0.0834**	−0.0545	−0.0878***
pct vol. in most active	−0.0458**	−0.1168**	0.2457***	0.2818***	−0.2709***	−0.2825***	−0.0549	−0.0995	0.1066**
lagged volatility	0.4860***	0.1016***	0.2610***	0.4558***	0.9390***	0.4179***	0.4850***	0.1465***	0.4285***
TED spread	3.0439***	2.6513**	−0.5819	−3.0985	28.2284***	4.6104***	0.0914	−2.3409	1.9381
VIX	0.0020***	0.0008*	0.0130***	0.0036***	0.0082***	0.0016***	0.0019***	0.0016***	0.0001
credit spread	−0.0378***	−0.0339***	−0.0374***	0.0662***	−0.0132	−0.0731***	−0.0450***	−0.0347***	0.0135
season									
2	−0.002	0.0196***	0.0213***	0.0017	−0.0131*	−0.0016	0.0367***	−0.0042	−0.0015
3	−0.0050**	0.0185***	0.0003	0.0007	0.0243***	0.0002	0.0301***	−0.0015	0.0092
4	−0.0092***	0.0022	−0.0054	−0.0049	−0.0085	0.0254***	−0.0121	0.0032	−0.0066
year									
2013	0.0080*	0.0026	0.0878***	0.0427***	−0.0087	−0.0135	−0.0862***	−0.0475***	−0.0228
2014	0.0164***	−0.0508***	0.0470***	0.0465***	−0.1394***	−0.0232	−0.1055***	−0.1728***	−0.0458***
2015	0.0055	−0.0802***	0.0839***	0.0575***	−0.1785***	−0.0898***	−0.1551***	−0.0911***	−0.1279***
2016	0.0548***	−0.0933***	0.0913***	0.3360***	−0.2156***	−0.1788***	−0.1458***	−0.2290***	−0.2024***
2017	0.0282***	−0.0691***	0.0688***	0.3266***	−0.3322***	−0.2192***	−0.2128***	−0.2766***	−0.1964***
2018	0.0047	−0.0600**	0.0229	0.3082***	−0.4402***	−0.2598***	−0.2180***	−0.2814***	−0.2002***
2019	0.0186*	−0.0494	0.1125***	0.1629***	−0.3685***	−0.3213***	−0.1906***	−0.2496***	−0.2194***
2020	−0.0135	−0.0585**	0.0461***	0.1434***	−0.1584***	−0.3600***	−0.2166***	−0.2913***	−0.1993***
2021	0.0125	0.0354	0.1180***	0.2315***	−0.1087***	−0.2454***	0.0333	−0.2222***	−0.1084***
Constant	0.5257***	0.8108***	−0.0712	−0.0223	0.9864***	1.0129***	1.0815***	1.3206***	0.4281***
R-squared	0.6719	0.6687	0.8625	0.7269	0.8521	0.7467	0.6574	0.9094	0.6378
Obs.	2399	2399	2399	2399	2399	2399	2399	2399	2399

Table 8: Robustness with respect to Alternative Methods of Computing the Amihud Measure: Outliers and Settlement Period

In all previous tables the Amihud measure was based on all data throughout a trading day and including any and all outliers ("all ino"). This table conducts additional robustness checks by isolating the settlement and non-settlement periods as well as data outliers (please see Subsection 3.4 and Table A2 of the Appendix for details). The models are: (1) pooled OLS with heteroskedasticity-robust standard errors using the main Amihud measure (i.e., settlement and non-settlement periods and including outliers), (2) robust OLS using Amihud based on settlement and non-settlement periods and excluding outliers ("all exo"), (3) robust OLS using Amihud based on non-settlement period only and including outliers ("ns ino"), (4) robust OLS using Amihud based on non-settlement period only and excluding outliers ("ns exo"), (5) robust OLS using Amihud based on settlement period only ("s ino"), (6) fixed effects model using the main Amihud measure, (7) fixed effects model using Amihud based on settlement and non-settlement periods and excluding outliers ("all exo"), (8) fixed effects model using Amihud based on non-settlement period only and including outliers ("ns ino"), (9) fixed effects model using Amihud based on non-settlement period only and excluding outliers ("ns exo"), (10) fixed effects model using Amihud based on settlement period only ("s ino"). ***, **, and * denote significance at 1%, 5%, and 10% level.

Variable	(1) OLS robust	(2) OLS robust all exo	(3) OLS robust ns ino	(4) OLS robust ns exo	(5) OLS robust s ino	(6) FE	(7) FE all exo	(8) FE ns ino	(9) FE ns exo	(10) FE s ino
HFT market share	−0.0290***	−0.0282***	−0.0282***	−0.0273***	−0.0453***	−0.0076***	−0.0068***	−0.0082***	−0.0074***	−0.0058***
pct aggressive volume	0.0050***	0.0049***	0.0040***	0.0037***	0.0200***	0.0024***	0.0021***	0.0025***	0.0022***	0.0006
ordered-to-traded vol.	0.0001	0.0004***	0.0000	0.0003**	0.0018***	0.0013***	0.0014***	0.0014***	0.0015***	0.0006***
volume per trade	0.0008***	0.0007***	0.0009***	0.0007***	0.0011***	0.0012***	0.0011***	0.0012***	0.0012***	0.0011***
pct directional	0.0063	0.0055	0.0028	0.0022	0.0495***	0.0171**	0.0155**	0.0144**	0.0130*	0.0484***
total stop volume	0.0000	0.0000	0.0000	0.0000	−0.0000*	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000
total active accounts	0.0000***	0.0000***	0.0000***	0.0000***	0.0000*	0.0000***	0.0000***	0.0000***	0.0000***	−0.0000***
total volume	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000	0.0000	0.0000	0.0000	0.0000
total open interest	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000***
pct OI in most active	0.0005	0.0006	−0.0004	−0.0004	0.0116***	0.0040***	0.0044***	0.0033***	0.0038***	0.0090***
pct vol. in most active	−0.0208***	−0.0200***	−0.0200***	−0.0191***	−0.0277***	−0.0094***	−0.0087***	−0.0091***	−0.0084***	−0.0092***
lagged volatility	0.0504***	0.0495***	0.0512***	0.0503***	0.0353***	0.0416***	0.0409***	0.0421***	0.0415***	0.0323***
TED spread	−0.0665	−0.0656	−0.0662	−0.0652	−0.1046	0.0939**	0.0948**	0.0899**	0.0906**	0.0826
VIX	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
credit spread	−0.0019***	−0.0018***	−0.0019***	−0.0017***	−0.0015***	−0.0010***	−0.0009***	−0.0009***	−0.0008***	−0.0013***
season										
2	−0.0008***	−0.0009***	−0.0009***	−0.0010***	−0.0003	−0.0009***	−0.0010***	−0.0010***	−0.0010***	−0.0003*
3	−0.0004***	−0.0005***	−0.0004***	−0.0005***	−0.0002	−0.0003**	−0.0004***	−0.0003**	−0.0004***	0.0002
4	0.0003**	0.0002*	0.0003**	0.0002*	0.0004**	−0.0003**	−0.0003***	−0.0003**	−0.0003***	−0.0002
year										
2013	−0.0009***	−0.0009***	−0.0011***	−0.0010***	0.0012***	−0.0015***	−0.0014***	−0.0016***	−0.0015***	−0.0006*
2014	0.0002	0.0002	0.0000	0.0000	0.0025***	−0.0002	−0.0002	−0.0003	−0.0003	0.0004
2015	0.0003	0.0002	0.0001	0.0000	0.0021***	−0.0008***	−0.0010***	−0.0008***	−0.0010***	−0.0014***
2016	0.0004*	0.0004*	0.0003	0.0002	0.0030***	−0.0028***	−0.0029***	−0.0027***	−0.0028***	−0.0040***
2017	0.0003	0.0003	0.0002	0.0002	0.0027***	−0.0041***	−0.0041***	−0.0038***	−0.0038***	−0.0070***
2018	−0.0017***	−0.0017***	−0.0018***	−0.0018***	0.0013**	−0.0056***	−0.0056***	−0.0053***	−0.0053***	−0.0078***
2019	−0.0003	−0.0005*	−0.0004	−0.0005*	0.0013***	−0.0048***	−0.0049***	−0.0045***	−0.0046***	−0.0081***
2020	0.0009***	0.0007**	0.0009**	0.0007*	0.0022***	−0.0027***	−0.0029***	−0.0024***	−0.0026***	−0.0064***
2021	0.0065***	0.0053***	0.0066***	0.0054***	0.0048***	0.0030***	0.0019***	0.0036***	0.0024***	−0.0044***
Constant	0.0358***	0.0345***	0.0356***	0.0342***	0.0327***	0.0147***	0.0139***	0.0141***	0.0133***	0.0178***
R-squared	0.7966	0.8019	0.796	0.8023	0.6209	0.4852	0.4949	0.4863	0.4969	0.2438
Obs.	21591	21591	21591	21591	21591	21591	21591	21591	21591	21591

Table 9: Robustness: Ticks-Based Amihud Measure, Outliers and Settlement Period

This table uses ticks-based Amihud measures while also isolating the settlement and non-settlement periods as well as data outliers (Subsection 3.4 and Table A2 of the Appendix). The models are: (1) pooled OLS with heteroskedasticity-robust standard errors using ticks Amihud based on settlement and non-settlement periods and including outliers ("all ino"), (2) robust OLS using ticks Amihud based on settlement and non-settlement periods and excluding outliers ("all exo"), (3) robust OLS using ticks Amihud based on non-settlement period only and including outliers ("ns ino"), (4) robust OLS using ticks Amihud based on non-settlement period only and excluding outliers ("ns exo"), (5) robust OLS using ticks Amihud based on settlement period only ("s ino"), (6) fixed effects model using ticks Amihud measure based on settlement and non-settlement periods and including outliers ("all ino"), (7)) fixed effects model using ticks Amihud based on settlement and non-settlement periods and excluding outliers ("all exo"), (8) fixed effects model using ticks Amihud based on non-settlement period only and including outliers ("ns ino"), (9) fixed effects model using ticks Amihud based on non-settlement period only and excluding outliers ("ns exo"), (10) fixed effects model using ticks Amihud based on settlement period only ("s ino"). ***, **, and * denote significance at 1%, 5%, and 10% level.

Variable	(1) OLS robust	(2) OLS robust all exo	(3) OLS robust ns ino	(4) OLS robust ns exo	(5) OLS robust s ino	(6) FE	(7) FE all exo	(8) FE ns ino	(9) FE ns exo	(10) FE s ino
HFT market share	−0.0258***	−0.0223***	−0.0234***	−0.0198***	−0.0859***	−0.1617***	−0.1565***	−0.1629***	−0.1577***	−0.1808***
pct aggressive volume	0.0126**	0.0142***	0.0096*	0.0108**	0.0645***	0.1085***	0.1086***	0.1083***	0.1082***	0.1139***
ordered-to-traded vol.	−0.0111***	−0.0107***	−0.0113***	−0.0108***	−0.0087***	0.0032***	0.0031***	0.0035***	0.0034***	0.0005
volume per trade	−0.0017***	−0.0021***	−0.0018***	−0.0022***	0.0031***	0.0022***	0.0020***	0.0022***	0.0020***	0.0044***
pct directional	0.0208	0.0257	0.0106	0.0160	0.1555***	0.0186	0.0215	0.0105	0.0139	0.1186**
total stop volume	0.0000***	0.0000***	0.0000***	0.0000***	0.0000	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
total active accounts	0.0000***	0.0000***	0.0000***	0.0000***	0.0000*	0.0000***	0.0000***	0.0000***	0.0000***	0.0000
total volume	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***
total open interest	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
pct OI in most active	−0.2376***	−0.2340***	−0.2399***	−0.2364***	−0.2056***	−0.0280***	−0.0253***	−0.0297***	−0.0269***	−0.0221***
pct vol. in most active	0.3012***	0.2977***	0.3034***	0.3001***	0.2965***	−0.0266***	−0.0270***	−0.0255***	−0.0258***	−0.0291***
lagged volatility	0.0905***	0.0888***	0.0929***	0.0913***	0.0482***	0.0929***	0.0917***	0.0944***	0.0931***	0.0738***
TED spread	1.5315**	1.4775**	1.5253**	1.4701**	1.7339**	0.3439	0.3037	0.3321	0.2903	0.4689
VIX	0.0013***	0.0013***	0.0013***	0.0013***	0.0012***	0.0012***	0.0013***	0.0012***	0.0013***	0.0011***
credit spread	−0.0099***	−0.0094***	−0.0099***	−0.0093***	−0.0056**	−0.0048***	−0.0043***	−0.0046***	−0.0041***	−0.0030*
season										
2	0.0000	−0.0003	−0.0001	−0.0004	0.0018	−0.0036***	−0.0038***	−0.0037***	−0.0039***	−0.0019**
3	0.0045***	0.0041***	0.0045***	0.0041***	0.0054***	−0.0014*	−0.0017**	−0.0015*	−0.0018**	−0.0004
4	−0.0013	−0.0016*	−0.0013	−0.0016*	−0.0023**	−0.0001	−0.0004	−0.0001	−0.0004	−0.0008
year										
2013	0.0021	0.0018	0.0016	0.0013	0.0081***	0.0014	0.0012	0.0012	0.0009	0.0037**
2014	−0.0093***	−0.0098***	−0.0097***	−0.0103***	−0.004	−0.0085***	−0.0091***	−0.0086***	−0.0092***	−0.0083***
2015	−0.0258***	−0.0260***	−0.0261***	−0.0264***	−0.0229***	−0.0221***	−0.0225***	−0.0219***	−0.0224***	−0.0258***
2016	−0.0241***	−0.0243***	−0.0244***	−0.0246***	−0.0193***	−0.0127***	−0.0132***	−0.0122***	−0.0127***	−0.0170***
2017	−0.0266***	−0.0266***	−0.0268***	−0.0268***	−0.0217***	−0.0107***	−0.0110***	−0.0099***	−0.0102***	−0.0200***
2018	−0.0287***	−0.0286***	−0.0290***	−0.0288***	−0.0224***	−0.0213***	−0.0212***	−0.0205***	−0.0204***	−0.0308***
2019	−0.0283***	−0.0286***	−0.0286***	−0.0288***	−0.0225***	−0.0192***	−0.0196***	−0.0184***	−0.0187***	−0.0276***
2020	−0.0062*	−0.0069**	−0.0065**	−0.0072**	−0.002	−0.0031	−0.0039**	−0.0025	−0.0033	−0.0138***
2021	0.0338***	0.0298***	0.0341***	0.0299***	0.0239***	0.0395***	0.0353***	0.0409***	0.0367***	0.0117***
Constant	0.0620***	0.0579***	0.0620***	0.0577***	0.0337***	0.0906***	0.0880***	0.0898***	0.0872***	0.0891***
R-squared	0.5353	0.5351	0.5358	0.5357	0.3976	0.3521	0.3512	0.3524	0.3514	0.275
Obs.	21591	21591	21591	21591	21591	21591	21591	21591	21591	21591

Table 10: Robustness with respect to Alternative Methods of Computing the Bid-Ask Spreads

This table conducts additional robustness checks by using alternative ways to compute the traded bid-ask spread. Main bid-ask spread variable: tbas_weighted_adj. Alternatives: tbas_simple_all, tbas_weighted_all, tbas_simple_adj. Please see Subsection 3.4 and Table A2 of the Appendix for details. ***, **, and * denote significance at 1%, 5%, and 10% level.

Variable	(1) OLS robust	(2) OLS robust simple all	(3) OLS robust simple adj	(4) OLS robust weighted all	(5) FE	(6) FE simple all	(7) FE simple adj	(8) FE weighted all
HFT market share	−0.3936***	−0.3455***	−0.0323***	−0.1780***	−0.9375***	−0.5540***	−0.5194***	−0.5587***
pct aggressive volume	0.1346***	0.0124	−0.0697***	0.0049	0.4078***	0.1740***	0.2277***	0.1977***
ordered-to-traded vol.	−0.0380***	0.0186***	−0.0151***	−0.0226***	0.0014	0.0127***	0.0081***	−0.0078***
volume per trade	−0.0207***	0.0120***	−0.0029**	−0.0177***	−0.0163***	0.0251***	0.0110***	−0.0105***
pct directional	0.0237	0.0766	−0.0646	0.1791	−0.0098	0.0106	−0.0972	0.1157
total stop volume	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
total active accounts	0.0000	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
total volume	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***	−0.0000***
total open interest	−0.0000***	0.0000***	−0.0000***	−0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
pct OI in most active	−0.6773***	−0.4058***	−0.5893***	−0.4604***	−0.1003***	0.0193**	−0.0443***	0.0179
pct vol. in most active	0.7868***	0.4546***	0.7995***	0.4708***	−0.0442**	0.0201*	0.0218	−0.0931***
lagged volatility	0.2479***	0.1411***	0.2378***	0.0951***	0.3022***	0.1527***	0.2568***	0.1379***
TED spread	6.0307***	3.9233***	4.3633***	5.0888***	1.7428*	2.6415***	1.1289*	2.2665***
VIX	0.0043***	0.0015***	0.0035***	0.0020***	0.0039***	0.0012***	0.0031***	0.0017***
credit spread	−0.0280***	0.0020	−0.0145***	−0.0013	−0.0161***	0.0102***	−0.0026	0.0060
season								
2	0.0062*	−0.0080***	−0.0070***	−0.0045*	−0.0032	−0.0126***	−0.0153***	−0.0093***
3	0.0206***	−0.0131***	0.0025	−0.003	0.0049*	−0.0202***	−0.0109***	−0.0132***
4	−0.0025	−0.0098***	−0.0128***	−0.002	0.0010	−0.0110***	−0.0094***	−0.0005
year								
2013	0.0158**	0.0012	−0.0011	0.0013	0.0161***	0.0015	0.0045	0.0038
2014	−0.0402***	−0.0197***	−0.0367***	−0.0290***	−0.0337***	−0.0173***	−0.0254***	−0.0239***
2015	−0.0831***	−0.0635***	−0.0693***	−0.0458***	−0.0652***	−0.0580***	−0.0456***	−0.0340***
2016	−0.0766***	−0.0507***	−0.0493***	−0.0568***	−0.0296***	−0.0366***	0.0023	−0.0229***
2017	−0.1121***	−0.0250***	−0.0611***	−0.0769***	−0.0468***	0.0019	0.0138***	−0.0243***
2018	−0.1513***	−0.0930***	−0.0979***	−0.1066***	−0.1145***	−0.0681***	−0.0402***	−0.0646***
2019	−0.1540***	−0.1031***	−0.1190***	−0.1113***	−0.1061***	−0.0746***	−0.0532***	−0.0653***
2020	−0.1178***	−0.0850***	−0.0845***	−0.0783***	−0.0872***	−0.0616***	−0.0337***	−0.0436***
2021	−0.0215*	−0.0642***	−0.0350***	−0.0420***	0.0154*	−0.0403***	0.0233***	−0.0024
Constant	0.6724***	0.4717***	0.3008***	0.4756***	0.6579***	0.4300***	0.3369***	0.5032***
R-squared	0.5793	0.4537	0.4699	0.4303	0.479	0.429	0.3931	0.2926
Obs.	21591	21591	21591	21591	21591	21591	21591	21591

Table 11: Difference-in-Difference Estimation of the Effects of the Change in Settlement Methodology on HFT Market Share and on Market Quality

This table uses difference-in-differences estimation to analyze how the change in settlement methodology affected HFT involvement in agricultural futures markets (Panel A), Amihud price impact (Panel B), and traded bid-ask spread (Panel C) vis-à-vis the other contracts in our sample. The change occurred on July 2, 2015 and we use daily data before and after the change. We use the generalized form of the difference-in-differences model in order to capture the effects across multiple groups. ***, **, and * denote significance at 1%, 5%, and 10% level.

Panel A: HFT Market Share

Variable	30 days pre- and post-change		60 days pre- and post-change		90 days pre- and post-change	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
Blended settlement	−0.3647**	0.047	−0.3747**	0.035	−0.3966**	0.031
Constant	−1.2529***	0.000	−1.2379***	0.000	−1.2348***	0.000
Market fixed effects	Yes		Yes		Yes	
Time (day) fixed effects	Yes		Yes		Yes	
R-squared	0.114		0.123		0.139	

Panel B: Amihud

Variable	30 days pre- and post-change		60 days pre- and post-change		90 days pre- and post-change	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
Blended settlement	0.9138*	0.060	0.8656*	0.073	0.8609*	0.075
Constant	−11.3556***	0.000	−11.3414***	0.000	−11.336***	0.000
Market fixed effects	Yes		Yes		Yes	
Time (day) fixed effects	Yes		Yes		Yes	
R-squared	0.115		0.098		0.098	

Panel C: Bid-Ask Spread

Variable	30 days pre- and post-change		60 days pre- and post-change		90 days pre- and post-change	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
Blended settlement	0.0366	0.813	0.0119	0.938	−0.0017	0.991
Constant	−0.5115***	0.004	−0.5187***	0.004	−0.5126***	0.005
Market fixed effects	Yes		Yes		Yes	
Time (day) fixed effects	Yes		Yes		Yes	
R-squared	0.002		0.000		0.000	