THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Economic and Social Research Council

High-frequency Trading in the Stock Market and the Costs of Option Market Making

Mahendrarajah Nimalendran
Khaladdin Rzayev
Satchit Sagade

SRC Discussion Paper No 113
December 2021

Systemic Risk Centre

Discussion Paper Series

ISSN 2054-538X

Abstract
Using a comprehensive panel of 2,969,829 stock-day data provided by the Securities and Exchange Commission (MIDAS), we find that HFT activity in the stock market increases market-making costs in the options markets. We consider two potential channels – the hedging channel and the arbitrage channel – and find that HFTs' liquidity-demanding orders increase the hedging costs due to a higher stock bid-ask spread and a higher price impact for larger hedging demand. The arbitrage channel subjects the options market-maker to the risk of trading at stale prices. We show that the hedging (arbitrage) channel is dominant for ATM (ITM) options. Given the significant growth in options trading, we believe that our study highlights the need to better understand the costs/risks due to HFT activities in equity markets on derivative markets.

Keywords: market microstructure; high-frequency trading; options market-making; hedging; liquidity

This paper is published as part of the Systemic Risk Centre's Discussion Paper Series. The support of the Economic and Social Research Council (ESRC) in funding the SRC is gratefully acknowledged [grant number ES/R009724/1].

Mahendrarajah Nimalendran, University of Florida

Khaladdin Rzayev, University of Edinburgh, Koç University and Systemic Risk Centre, London School of Economics

Satchit Sagade, Goethe University, Frankfurt

Published by
Systemic Risk Centre
The London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© Mahendrarajah Nimalendran, Khaladdin Rzayev and Satchit Sagade, submitted 2021

High-frequency Trading in the Stock Market and the Costs of Option Market Making

Mahendrarajah Nimalendran [a,*], Khaladdin Rzayev [b,c,d*], Satchit Sagade [e,*]
[a] *University of Florida*
[b] *University of Edinburgh*
[c] *Koç University*
[d] *Systemic Risk Centre, London School of Economics*
[e] *Goethe University, Frankfurt*

Abstract

Using a comprehensive panel of 2,969,829 stock-day data provided by the Securities and Exchange Commission (MIDAS), we find that HFT activity in the stock market increases market-making costs in the options markets. We consider two potential channels – the *hedging channel* and the *arbitrage channel* – and find that HFTs' liquidity-demanding orders increase the hedging costs due to a higher stock bid-ask spread and a higher price impact for larger hedging demand. The arbitrage channel subjects the options market-maker to the risk of trading at stale prices. We show that the hedging (arbitrage) channel is dominant for ATM (ITM) options. Given the significant growth in options trading, we believe that our study highlights the need to better understand the costs/risks due to HFT activities in equity markets on derivative markets.

Keywords: market microstructure; high-frequency trading; options market-making; hedging; liquidity

*Emails: mahen.nimalendran@warrington.ufl.edu, khaladdin.rzayev@ed.ac.uk, sagade@safe.uni-frankfurt.de.
Corresponding author's contact information: University of Edinburgh Business School, University of Edinburgh; e-mail: khaladdin.rzayev@ed.ac.uk
We thank Phil Mackintosh and Jeffrey W. Smith at NASDAQ for data provision. The data is made available at no cost to academics following the provision of a project description and the signing of a nondisclosure agreement. Khaladdin gratefully acknowledges the support of the Trans-Atlantic Platform by the Economic and Social Research Council (ESRC) (grant number ES/R004021/1). This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 329107530.

1. Introduction

High-frequency trading (HFT) has materially impacted the dynamics of electronic markets. The extensive and growing literature examining the implications of HFT has predominantly focused on the within-market, mainly the stock market, quality effects of HFT.[1] Investigations into the cross-asset impact of HFT are limited. In this paper, we attempt to fill this gap by examining the impact of HFT in the stock market on the options market microstructure.[2] Figure 1 shows the time-series evolution of trading volume in the US equity and options markets. We see that from 1996 to 2020 the options market volume grew at an annualized compound rate of 15% compared to 11% for the stock trading volume.

INSERT FIGURE 1 ABOUT HERE

An increase in options trading volume has been observed alongside some notable trends and records in the last two years. For instance, for the first time in history, in 2020 the number of shares traded with options contracts was higher than the underlying stock market trading volume.[3] Trading volume in the US equity options markets also hit a record high in two consecutive years in 2020[4] and 2021.[5] Thus, we believe that studying the potential externalities imposed by HFTs in the stock market on the options market is timely and important.

---

[1] Hendershott *et al.* (2011), Brogaard *et al.* (2015), Van Kervel and Menkveld (2019) and Hagströmer and Nordén (2013) examine the effects on stock market liquidity; Kirilenko *et al.* (2017) and Lee (2015) investigate the impact in the futures market; Chaboud *et al.* (2014) and Jiang *et al.* (2014) focus on FX and fixed-income securities markets respectively.

[2] Throughout this paper, we use the phrase "high-frequency trading" or "HFT" to refer to HFT activity in the stock market, unless we explicitly indicate otherwise. The acronym HFT is used interchangeably to refer to high-frequency traders and high-frequency trading.

[3] https://finance.yahoo.com/news/option-trading-volume-higher-underlying-211006236.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAEDVmOUhGSccv7vJXLzChzsOdqv_dwRYGuoAr4To9lPwq1ho_ANZqf8yViK5YWjwDoNZAawTz64F1XrmDdCFkag0FKL5OBmTJ1K0OvgXGljjm_wmfjPiDhIEsOjo3HMIO9sghsOBOjYIpvj9KrYsEGRvPPimhzoNXO1gxEtP0ZKS

[4] https://www.thetradenews.com/occ-clears-record-volumes-for-us-exchange-listed-options-in-2020/

[5] https://www.reuters.com/article/us-usa-stocks-options-idUSKBN29K2OI

In this study, we specifically address the following questions: (i) How does HFT in stocks impact the liquidity of options written on these stocks? (ii) Does the effect vary by option moneyness? (iii) Do different HFT strategies affect options market liquidity differently? (iv) Is the effect exclusively via the stock liquidity channel, or is there a direct effect after controlling for stock liquidity?

To answer these questions, we construct HFT measures by using the Securities and Exchange Commission's (SEC) Market Information Data Analytics System (MIDAS) data and show that HFT activity in the equity markets is associated with a significant deterioration in liquidity (increase in quoted bid-ask spreads) in the options markets. On average, a one standard deviation increase in HFT activity is associated with a 4.27% (8.89%) higher proportional (dollar) bid-ask spread. To address potential endogeneity concerns between the options market spread and stock market HFT, we employ a two-stage least squares (2SLS) instrumental variable (IV) approach by using two different sets of instruments. First, we follow Lee and Watts (2021) and use the randomized experiment of tick size changes launched by the SEC as an instrument for the level of HFT (Hagströmer & Nordén 2013 also employ tick size change as an exogenous shock on HFT). Second, following Hasbrouck and Saar (2013), we instrument the level of HFT in a stock-day with the average level of HFT on that day in all other stocks in the same market size quintile. Our results remain robust to both approaches. The consistency between the 2SLS IV approach and the fixed effects OLS regression confirms that the relationship between HFT and options market liquidity can be interpreted causally.

We then propose two channels to explain the relationship between HFT and options liquidity: (i) the *hedging* channel; and (ii) the *arbitrage* channel. The hedging channel is based on HFT activity in the stock market, affecting the option market bid-ask spread through its effect on option market-makers' hedging costs. Black and Scholes (1973) show that option market-makers can perfectly hedge their exposure by acquiring an offsetting position in the underlying asset and continuously rebalancing their portfolio to ensure it remains delta-neutral. However, due to friction and market imperfections, they can only imperfectly hedge their positions. Consequently, they require compensation for the transaction costs and the risks associated with imperfect hedging of their exposure (see Cho & Engle 1999; Kaul *et al.* 2004; Wu *et al.* 2014).[6]

Boyle and Vorst (1992) and Cho and Engle (1999) demonstrate that option market-makers' hedging costs are proportional to the stock market bid-ask spread (see also Kaul *et al.* 2004). This implies that the hedging channel leads to two opposing predictions regarding the impact of HFT on the options spread. On the one hand, options market-makers can more precisely hedge their exposure at lower cost if liquidity in the underlying stock market improves. This is likely when liquidity-supplying HFT activity is high, as argued by Hendershott *et al.* (2011) and Brogaard *et al.* (2015). On the other hand, if liquidity is lower in the underlying stock market, options market-makers will likely quote wider spreads due to increased costs of hedging their exposure and keeping their positions partially unhedged by reducing the rebalancing frequency. The arbitrage and momentum/directional strategies employed by HFT firms rely on aggressive trades and are harmful to the overall market liquidity (see Budish *et al.* 2015; Foucault *et al.* 2016; Foucault *et al.* 2017).

Kaul *et al.* (2004) and Engle and Neri (2010) show that the transaction costs of hedging are due to setting up and unwinding the initial hedge and rebalancing costs. At-the-money (ATM) options have high rebalancing costs due to their high gamma. In-the-money (ITM) options have the highest absolute delta and high costs of initially setting up and unwinding the hedge. By contrast, out-of-the-money (OTM) options have the lowest delta and gamma, and, therefore, the hedging cost is relatively lower in these options. Above discussion suggests that,

[6] Leland (1985) and Boyle and Vorst (1992) develop alternative discrete-time option replication strategies in the presence of transaction costs.

if the hedging channel indeed explains the association between HFT and the cost of options market-making, then the impact of HFT activity in the underlying market on options spreads should be higher for ATM and ITM options.

The arbitrage channel relates to violations of the put-call parity relationship from the asynchronous adjustment in stock and option prices. These "toxic" arbitrage opportunities, to the extent that they are driven by a delay in incorporating information into stock and option prices, will induce market-makers to post wider quotes to protect against adverse selection losses (see Budish *et al.* 2015; Foucault *et al.* 2017). The short-lived nature of such arbitrage opportunities requires arbitrageurs to rely on aggressive orders in the stock (see Kozhan & Tham 2012). Options market-makers are particularly exposed to such toxic arbitrage losses due to the exchange-imposed caps on the number of quote updates and fines on traders with higher message-to-transaction ratios (see Muravyev & Pearson 2020). These restrictions severely limit the options market-makers' ability to update their quotes in response to new information. In addition, liquidity-consuming HFTs engaging in cross-market arbitrage strategies may exploit violations of the put-call parity relation by sniping stale quotes in the options market. Hence, aggressive HFTs can expose options market-makers to the options market risk of trading at stale quotes. Notably, Halpern and Turnbull (1985) and Galai (1978) observe that violations of the put-call parity relationship are more frequent for ITM options. Thus, if the arbitrage channel is the dominant channel to explain the relationship between HFT and the cost of options market-making, then the impact of HFT on the options spread should be higher for ITM options and should be weakened on days without information.

To test the hedging and arbitrage channels, we first examine whether the effects of HFT on options spreads vary by three moneyness groups: (i) ATM, (ii) ITM, and (iii) OTM options. While the results are statistically significant in all moneyness groups, consistent with the arbitrage and hedging channels, the economic magnitudes of the impact are relatively higher for ATM and ITM options. However, it is important to note that our main data (the SEC's MIDAS data) does not allow us to directly test the hedging and arbitrage channels. The direct testing of the hedging and arbitrage channels requires us to use more granular HFT data; it is necessary to disaggregate HFT activity into aggressive and passive trades to test these channels. The MIDAS data only allows us to construct a general measure of HFT. Therefore, we employ proprietary HFT data provided by NASDAQ to further test the hedging and arbitrage channels. Unfortunately, this data is only available for 2009, but, due to its granularity, it complements the MIDAS data that covers a much longer period: 2012-2019.

We conduct two additional tests using NASDAQ's proprietary data. In the first analysis, we find that after controlling for the stock market bid-ask spread, the impact of liquidity-consuming HFT activity on options spread remains positive and significant for ATM and ITM options. Specifically, a one standard deviation increase in aggressive HFT activity increases ATM options proportional (dollar) spreads by about 5.71% (13.89%).[7] For ITM options, a one standard deviation increase in aggressive HFT activity is associated with a +6.22% (+9.66%) change in the proportional (dollar) options market spread. Importantly, the association between aggressive HFT activity and ITM options spreads is weakly statistically significant, while the economic magnitude is quite high. Finally, the results are insignificant for OTM options. These findings are consistent with the predictions of the hedging and arbitrage channels.

In the second analysis, we further isolate the arbitrage channel by identifying and excluding days involving the release of firm-specific news from our sample. We observe that the impact of aggressive HFT on ITM options spread becomes insignificant. On the other hand, for ATM options, while the magnitude of the effect decreases by about 19% for proportional

[7] The estimates based on the NASDAQ proprietary data are similar to the estimates we obtain from the longer panel based on the SEC-MIDAS data (4.27% (8.89%) for the proportional and (dollar) spreads).

spread, the relationship remains significant. These two tests indicate that the hedging (arbitrage) channel predominantly explains the relationship between aggressive HFTs, and options market spread for ATM (ITM) options. However, we caution against over-interpretation of this particular result because the NASDAQ HFT data has two important limitations. First, it only covers 103 randomly-selected stocks for the year 2009. Second, the data includes HFT activity in the NASDAQ only. We also find that the economic magnitude of the impact of HFT on options spreads is substantial in ITM contracts when we use the MIDAS HFT data, which covers most of the US-listed common stocks and the period of 2012 to 2019. Therefore, we believe that the association between HFT and ITM options spreads is likely due to both hedging and arbitrage channels.

Interestingly, the effect of liquidity-providing HFT activity on options spread is not statistically significant. We show that this is because 100% of passive HFT trades' effect on options market spread is captured by the bid-ask spread in the stock market. Conversely, 75% of the impact of aggressive HFT activity on spreads in the options market is direct and unrelated to the stock market bid-ask spread. This is likely due to several factors: the bid-ask spread-capturing transaction costs for small(er) orders due to the limited depth available at the best quotes, the predictability of options market makers' hedging demand, options market-makers employing order-splitting algorithms to execute large hedging trades, and aggressive HFTs engaging in predatory trading (see Brunnermeier & Pedersen 2005).

Our findings offer insights into the role of HFTs in explaining the cross-asset market microstructure dynamics between the stock market and the options market. To the best of our knowledge, this study is the first to provide evidence on the direct impact of various strategies used by HFTs on the options spread. In a recent study, Kapadia and Linn (2019) use the Glosten and Milgrom (1985) framework to develop a model of trading in the primary market (stock) and a derivative market (option). The model is used to relate the bid-ask spread in the options market to the volatility of the bid-ask spread in the stock market via a synthetic stock based on put-call parity. The authors use a glitch in Knight Capital's trading platform that erroneously executed a large number of small orders for a selected number of stocks. While the glitch increased uninformed order flow, it also resulted in persistent liquidity-related uncertainty. The study finds that options spreads widened among impacted stocks and remained wide for a quarter of an hour after the broker-dealer fixed the glitch in their computer system.

Our study also examines how low-latency traders (HFTs) in the stock market impact on the options spreads. While Kapadia and Linn (2019) analyze the effects of stock market quote uncertainty on the options market spread, we focus directly on the impact of HFTs on options spreads by using comprehensive HFT data. Due to the rich and granular MIDAS and NASDAQ HFT datasets, our study provides direct evidence of the negative impact of HFT on options market-making. While HFTs may cause uncertainty in the underlying stock liquidity, this is not solely driven by HFTs. Thus, the uncertainty in liquidity is not necessarily a proxy for HFT activity and is not commonly used in the literature. Nevertheless, for robustness, we control for the volatility of the stock bid-ask spread by including it as an explanatory variable, and we obtain the same results (see Table A.1).

The focus of our study is also related to the work of Mishra *et al.* (2012), one of the first studies to examine the impact of automation on options markets. Mishra *et al.* (2012) use high-frequency data sourced from the Options Price Reporting Authority (OPRA), and show that automation reduces bid-ask spreads and increases liquidity in options markets. Our study differs from theirs in at least two ways. First, they focus on the impact of automated trading on option liquidity. While HFT strategies require automated trading, HFTs are only a subset of all the traders on the automated trading platform, and they have different impacts on market dynamics. Second, and most importantly, Mishra *et al.* (2012) investigate the relationship between option liquidity and option automated trading. However, as discussed, liquidity in options markets is also determined by equity market dynamics (see Cho & Engle 1999). Since HFTs are one of the most important drivers of equity market liquidity, it is expected that the options spread will also be impacted by HFT activity in the underlying market.

Although HFTs use both market-making and speculative trading strategies in the stock market, it is widely accepted that the majority of them (about 80%) follow the market-making strategy, implying that the adverse-selection-avoidance channel dominates the picking-off channel in this market (see Hagströmer & Nordén 2013; Menkveld 2013). In this context, if HFTs predominantly engage in market-making in the options market as well, then our results suggest that liquidity-demanding HFTs in the stock market will increase the costs of liquidity-supplying HFTs in the options market. As such, this study provides a complementary perspective to that of Menkveld and Zoican (2017), who show that speculative HFTs impose adverse-selection costs on market-making HFTs. Focusing on the stock and options markets, we find that this is also the case in the cross-market setting.

In other related literature, studies examine the effects of various market microstructure determinants of the options spread. Easley *et al.* (1998) propose a model where informed traders choose between stock and options markets based on the relative transaction costs in the markets and the "bang-for-buck" in the form of leverage afforded by the options market. The authors conclude that, depending on the relative transaction costs in the markets, there can be a separating equilibrium where informed traders only trade in the stock market, or a pooling equilibrium where informed traders trade in both markets. Subsequent empirical work that focused on informed trading in these two venues has largely supported the model's theoretical predictions. For example, Cao *et al.* (2005) find informed options trading before takeovers. Hu

(2014) provides evidence of an information channel by documenting that the options market-makers' initial delta hedging strategy is reflected in stock prices.[8]

Roadmap: The remainder of this paper is organized as follows. In Section 2, a description of the data is provided. Section 3 presents the estimation approaches, and Sections 4 and 5 provide the estimation results. Section 6 outlines the conclusion.

2. Data and Variable Construction

2.1. Data Sources

We compile data from several sources. The Securities and Exchange Commission's Market Information Data Analytics System (MIDAS) data is used to construct HFT proxies. The SEC offers MIDAS data to promote investigations on the US equity markets structure. The MIDAS collects data across all major US exchanges since 2012, and it includes the following variables: lit trade and order volume, hidden trade and order volume, odd lot trade and order volume, and counts of cancellations (full or partial) for each day. The dataset covers over 5,570 stocks and 2,730 exchange-traded funds.

End-of-the-day option bid and ask prices, trading volumes, Greeks, and implied volatilities are obtained from OptionMetrics. Greeks and implied volatilities are computed by using a binomial tree where an interest rate is constant. We obtain options data for all US-listed common stocks from 2012 to 2019 as MIDAS data is available from 2012 only. Our sample does not include 2020 due to the Covid-19-induced volatility. We follow the existing literature and exclude long-term options, i.e., those with maturities greater than 180 days. This allows us to restrict our analysis to the most actively traded options contracts, i.e., options contracts with higher trading volume (see Brenner *et al.* 2001; Christoffersen *et al.* 2017). There are over 1

---

[8] Pan and Poteshman (2006), Ni *et al.* (2008), Cremers and Weinbaum (2010), Ge *et al.* (2016) and Collin-Dufresne *et al.* (2020) document the role of options markets in the price discovery process, which is consistent with the information channel.

billion option transactions in the final sample, and the total nominal and USD volumes of these transactions are approximately 28.7 billion and 73.9 billion USD respectively.

Daily ask, bid, and trading prices are obtained from the Center for Research in Securities Prices (CRSP) dataset. The main analysis includes all CRSP common stocks matched in the MIDAS and OptionMetrics databases. The resulting sample has 2,746 unique securities and 2,969,829 security-day observations.

2.2. Variable Construction

We construct five HFT measures from the MIDAS data. Our first proxy is the ratio of the quote-to-trade volume (see Hendershott *et al.* 2011), $QT_{i,d}$, which is computed as the sum of order volume for all order messages divided by the sum of trade volume for all trades that are not against hidden orders. The second HFT measure is the cancel-to-trade ratio, $CT_{i,d}$, or the number of all cancel messages (full or partial) divided by the number of trades (see Weller 2018). The third (fourth) HFT proxy is the odd lot rate (odd-lot volume), $OR_{i,d}$ ($OV_{i,d}$), which is calculated as the number of odd lot trade messages (odd lot trade volume) divided by the number of all trade messages (trade volume) (see O'Hara *et al.* 2014).[9] The final HFT measure is the inverse of the average trade size, $ITS_{i,d}$, or the number of trades divided by the number of shares traded (see Conrad *et al.* 2015).

Consistent with the literature, we measure option liquidity using bid-ask spreads (see as an example, Muravyev & Pearson 2020). Our liquidity proxies are the proportional quoted spread, computed as the difference between the ask and bid prices divided by the midpoint and the dollar quoted spread, which equals the difference between the ask and bid prices for each options transaction. Options contracts with various characteristics (different maturities and strike prices) are traded for each stock and day. As a result, for each stock, we observe multiple

[9] Odd lots are trades that have a volume of less than 100 shares.

bid-ask spreads during the day. Given that our main analysis is based on panel regressions at a daily frequency, we follow Muravyev and Pearson (2020) and compute the daily proportional spread ($OPspread_{i,d}$) and dollar spread ($ODspread_{i,d}$) as the dollar-volume-weighted average of all spreads for stock *i* and day *d*.

Apart from the variables mentioned above, we employ several control variables to capture stock and options market dynamics. Our options market variables are the options volume ($Ovolume_{i,d}$), implied volatility ($Oimplied_{i,d}$), absolute option delta ($|Odelta_{i,d}|$), option vega ($Ovega_{i,d}$), and option gamma ($Ogamma_{i,d}$). The $Ovolume_{i,d}$ is the natural logarithm of the daily trading volume (contracts) for each stock *i* and day *d*. OptionMetrics provide the implied volatility and option Greeks. Similar to options market spread measures, we compute the daily $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ovega_{i,d}$, and $Ogamma_{i,d}$ as the dollar-volume-weighted averages of all implied volatilities, absolute deltas, vegas, and gammas for stock *i* and day *d*. We use the absolute value of delta as call and put options and have different signs for deltas, i.e., a call option delta is positive, while delta is negative for put options.

We employ the following variables to control for stock market activity: stock spreads ($SPspread_{i,d}$ and $SDspread_{i,d}$), and stock price volatility ($SVolatility_{i,d}$). $SPspread_{i,d}$ (proportional) and $SDspread_{i,d}$ (dollar) are computed using the best-ask and bid prices for each stock and day. Specifically, $SPspread_{i,d}$ is computed as the difference between the best-ask and bid prices for stock *i* and day *d*, divided by the midpoint of the two prices on the same day. $SDspread_{i,d}$ is the difference between the best-ask and bid prices for stock *i* and day *d*. $SVolatility_{i,d}$ is the absolute difference between the last transaction prices for stock *i* on days *d* and *d-1*.

INSERT TABLE 1 ABOUT HERE

Table 1 provides an overview of all the variables and their computation methods.

2.3. Descriptive Statistics

Table 2 provides descriptive statistics for the 2,746 stocks in the full sample and 1,235 stocks in the tick size pilot sample and their listed options (see Section 3.2 for detailed information on the tick size pilot sample). We winsorized all variables at the 1st and 99th percentile values.

INSERT TABLE 2 ABOUT HERE

Panel A reports the summary statistics for the underlying stock market variables. For the full sample, the average $SPspread_{i,d}$ ($SDspread_{i,d}$) is 0.11% (0.02 USD), which is lower than the average $OPspread_{i,d}$ ($ODspread_{i,d}$) by a factor of approximately 310 (22.5). $SPspread_{i,d}$ and $SDspread_{i,d}$ are higher for the tick size pilot sample. This is expected as relatively small and illiquid stocks are included in the Tick Size Pilot Program (see Chung *et al.* 2020).

Panel B reports the summary statistics for the options market variables for the full sample and the three moneyness groups (ATM, OTM, and ITM) based on the classification provided by Bollen and Whaley (2004). We define OTM options as those with absolute option delta $|\Delta| \leq 0.375$, ATM options as those with $0.375 < |\Delta| \leq 0.625$, and ITM options as those with $|\Delta| > 0.625$. For ATM options, the average $OPspread_{i,d}$ and $ODspread_{i,d}$ are 21.64% and 0.38 USD respectively. $ODspread_{i,d}$ increases as we move from OTM (0.27 USD) to ITM (0.62 USD) contracts. This is not surprising as ITM (OTM) contracts have higher (lower) prices. Wei and Zheng (2010) also show that the dollar spread is higher for options with higher prices. By contrast, $OPspread_{i,d}$ increases as we move from ITM (15.73%) to OTM (52.10%) contracts. Wei and Zheng (2010) link this to leverage; in particular, they argue that option contracts with higher leverage (OTM contracts) attract more informed traders and are associated with a higher spread. ATM options have the highest $Ovolume_{i,d}$, $Ovega_{i,d}$ and

$Ogamme_{i,d}$. This is expected as these contracts are the most active. Consistent with the equity market spreads, the options spreads are higher for the tick size pilot sample.

2.4. Correlation Between Control Variables and Dependent Variables

We use several variables as controls for the stock and options spread regression models. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ovega_{i,d}$ and $Ogamma_{i,d}$, and the stock market variables are $Sspread_{i,d}$ and $SVolatility_{i,d}$. To show that the selected control variables are indeed valid and capture the important variation in options spreads, we first estimate the association between these control variables and options spreads.

INSERT TABLE 3 ABOUT HERE

Table 3 reports the correlation between the control variables and dependent variables used in the study. The estimates suggest that the associations between control variables and options spreads are significant and in line with the relevant literature. For instance, $Ovolume_{i,d}$ is negatively correlated with options spreads, suggesting that a higher trading volume means higher liquidity. Consistent with Cho and Engle (1999) and Engle and Neri (2010), we also find that the stock spread is positively and significantly correlated to the options spread, implying that the stock spread is a significant determinant of the options market liquidity.

3. Estimation Approaches

3.1. Fixed Effect Estimation

We begin testing the impact of HFT on the options spreads by estimating the following fixed-effect models:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 HFT_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (1)$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), and $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$,

$CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$). The α_i and β_d are stock and time (day) fixed effects. Standard errors are double clustered on stock and day.[10] The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$, and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. All these variables are defined in Table 1.

3.2. Two-stage Least Squares (2SLS) Instrumental Variable (IV) Approach

The presence of HFT in the stock markets and option bid-ask spreads could be determined by unobserved common factors, implying they are jointly endogenous. To address these potential endogeneity concerns, we estimate the following 2SLS IV approach:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT_{i,d}} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (2)$$

$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (3)$$

where $\widehat{HFT_{i,d}}$ is the fitted values of $HFT_{i,d}$ obtained by regressing $HFT_{i,d}$ on $IV_{i,d}$. We use two different sets of instruments for robustness. Our first instrument is based on tick size changes (see Hagströmer & Nordén 2013). We follow Lee and Watts (2021) and employ the introduction of the Tick Size Pilot Program as an exogenous shock on HFT. In October 2016, the SEC launched a two-year pilot program to investigate the effect of increased tick size on market quality. The pilot program consists of treatment and control groups of 1,200 randomly selected stocks each. The treatment securities are split into three treatment groups, each group subject to different changes. The stocks that are included in the first treatment group must be quoted in 5 cent increments. In addition to being subject to the increments in tick size applying

[10] Double-clustered standard errors are used in all subsequent models.

to the first group, the stocks in the second treatment group must be traded in 5 cent increments. The stocks in the third treatment group have the same quoting and trading rules as the second treatment group. However, they are also subject to the "trade-at" rule: orders must not be executed in dark venues for these stocks unless there is a meaningfully better price in dark markets. The tick size sample includes all stocks in the Tick Size Pilot Experiment matched in the MIDAS and OptionMetrics databases. The resulting sample has 1,235 unique securities (617 control and 618 treated stocks) and 640,306 security-day observations.

We follow Lee and Watts (2021) and Chung *et al.* (2020) and use these three groups as our treatment group. The Tick Size Pilot Program commenced gradually from October 3, 2016, with all treatment firms included by the end of October; the program was implemented over the next two years. Thus, for the treatment stocks, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018) and zero before (from October 1, 2014 to October 2, 2016). For the control stocks, $IV_{i,d}$ equals zero during the sample period (from October 1, 2014 to September 28, 2018). Given that we include matched samples of stocks into the first-stage model, the estimation results of the second-stage model effectively give us a difference-in-differences estimate (see Malceniece *et al.* 2019).

While the change in tick size is used as an exogenous shock on HFT in the relevant market microstructure literature (see Hagströmer & Nordén 2013; Lee & Watts 2021), it is not a perfect setting. The main concern with this identification strategy is that the tick size changes do not only influence the amount of HFT. For instance, Albuquerque *et al.* (2020) show that an increase in tick size led to increased transaction costs (see also Chung *et al.* 2020). Thus, the changes in tick size in equity markets can affect options spreads through a direct impact on the transaction costs in the equity markets. Although we control the stock spread in our first and second stage models to address this issue, we employ an additional instrument as our second identification strategy for robustness. We follow Hasbrouck and Saar (2013) and instrument the level of HFT in a stock-day with the average level of HFT on that day in all other stocks in the corresponding size quintile (see also Comerton-Forde & Putniņš 2015). This variable meets the requirements for an instrument for two reasons. First, the level of HFT in other stocks is correlated with the level of HFT in a particular stock. Second, HFT in other stocks is unlikely to be driven by the nature of liquidity in options on a particular stock. This instrument alone is sufficient for identification; however, we follow Foley and Putniņš (2016) and include the first lag values of the dependent variable (HFT measures) as an additional control variable in the first-stage regression.

4. Estimation Results

4.1. The Impact of HFT on Options Spreads

The estimation results of Equations 1 and 2 are presented in Table 4.

INSERT TABLE 4 ABOUT HERE

The first-stage results of the 2SLS IV approaches are reported in the Appendix (see Table A.2). Overall, our selected instruments are significantly correlated with all HFT measures, and the signs of the associations are as expected. We scale all independent variables by their standard deviation as this allows us to directly estimate the economic significances of the effect (see Foucault & Fresard 2014). Column 1 presents the results of the standard OLS approach with stock and day fixed effects, while Columns 2 and 3 report the results for the 2SLS IV approach. Overall, we see that HFT in the underlying markets is positively related to the options spreads. This suggests that HFTs play a negative role in options market liquidity. The results are remarkably consistent across various HFT proxies and specifications. Consistency between the OLS and 2SLS IV methods allows us to establish a causal link between HFT and the options spread while ruling out endogeneity concerns.

The magnitude of the impact is economically significant as well. For example, estimates using the standard OLS approach with stock and day fixed effects show that, on average, a one standard deviation increase in HFT increases $ODspread_{i,t}$ and $OPspread_{i,t}$ by 8.89% and 4.27% respectively. This means that, for the average option, $ODspread_{i,t}$ ($OPspread_{i,t}$) will increase from 0.45 USD (34.97%) to 0.49 USD (36.46%). The economic magnitudes of the 2SLS IV estimates depend on which set of instruments is used. When we use the average HFT of the other stocks in the same size quintile as our instrument in the 2SLS-IV, the economic effects are even greater in magnitude. A one standard deviation shock on HFT raises $ODspread_{i,t}$ and $OPspread_{i,t}$ by 14.22% and 6.46% respectively. However, when we use the changes in tick size as an instrument, a one standard deviation increase in HFT is associated with a +3.30% and +2.98% change in $ODspread_{i,t}$ and $OPspread_{i,t}$ respectively. As seen, tick size results are lower than those of other specifications. This is because only small stocks are included in the Tick Size Pilot Program. This is a plausible explanation as Brogaard *et al.* (2014) show that HFTs are less active in small stocks.

In this paper, we employ five HFT measures and three different model specifications to investigate the relationship between HFT and options spread. Hence, we believe that these estimations are sufficient to isolate the effects of HFT on options spread. Nevertheless, the results can still be driven by some common factors that drive both HFT and options spread. To rule out this concern, we include several control variables in Equation 2.

We also want to specifically discuss one of these common factors: stock market volatility. The stock market volatility is one of the determinants of HFT trading volume (see Brogaard *et al.* 2014). Furthermore, as reported in Table 3, volatility is an important determinant of options liquidity. The implication is that our model may capture the impact of stock market volatility on both HFT and options spreads rather than the association between HFT and options spreads. While we are controlling for absolute price changes – one of the common volatility measures in the literature (see Karpoff 1987) – in Equation 2, we conduct some moderation tests to further strengthen the interpretation of our results. Specifically, we estimate two additional models. We first test the association between HFT and stock realized volatility (where the realized spread is measured as the variance estimates based on the returns calculated using the midpoint of the quoted bid and ask prices at every second during the trading hours). Second, we examine the association between HFT and options spread after controlling for both the absolute value of price changes and the realized volatility. The estimation results are reported in the Appendix section (Tables A.3 and A.4). Two points stand out. First, although the association between HFT and the realized spread is statistically significant, the signs of the association are not consistent across different HFT measures and specifications. Second, and importantly, the effects of HFT on options spread are statistically significant even after controlling for the realized volatility. These results suggest that volatility does not drive our results.

The results provided in Table 4 show the average effects (across firms and years) of HFT on options spread, suggesting that the estimations can mask time variation in the effects. Hence, to further understand the time variation in the association between HFT and options spread, we estimate Equation 2 year-by-year.

INSERT FIGURE 2 ABOUT HERE

Figure 2 displays the yearly estimates of γ_1. For $ODspread_{i,t}$ ($OPspread_{i,t}$), the associations between HFT (measured by $QT_{i,t}$) and options spread are significant in seven (six) of eight years. The results suggest that, overall, the effect of HFT on options spreads is persistent across the years.

4.2. The Impact of HFT on Options Spreads by Moneyness

To further study the impact of HFT on options spreads and understand the channels through which HFT impacts options spreads, we split our sample into moneyness groups and estimate Equations 2 and 3 for each group.

INSERT TABLE 5 ABOUT HERE

Table 5 presents the estimation results for ATM, ITM, and OTM options respectively. Similar to the results reported in Table 4, there is a positive and significant relationship between HFT and options spreads across the three moneyness groups in most specifications. There are a few exceptions for the OTM options. Specifically, for two HFT proxies ($OV_{i,t}$ and $TS_{i,t}$), the impact of HFT on $OPspread_{i,t}$ is negative when we use the standard OLS approach. Nevertheless, the results show that the impact of HFT on the options market-making is robust and persistent. In addition to the statistical significance of the results, we also study the economic significance among different moneyness groups.

INSERT TABLE 6 ABOUT HERE

Table 6 presents the economic magnitudes of the effect of HFT on options spreads for the full sample and three moneyness groups. Two observations stand out. First, as discussed above, the magnitude of the increase in options spreads is economically meaningful. Second, and more importantly, the magnitudes of the impact of HFT on options spreads differ based on the moneyness of the options contract. In all specifications, the economic magnitude is higher for ITM (17.10% for $ODspread_{i,t}$ and 6.80% for $OPspread_{i,t}$) and ATM (6.32% for $ODspread_{i,t}$ and 7.55% for $OPspread_{i,t}$) options in comparison to OTM options (3.93% for $ODspread_{i,t}$ and 0.91% for $OPspread_{i,t}$). We conjecture that this finding is important as it may give us insights into the channels that drive the HFT option market-making relationship.

As suppliers of immediacy, market-makers in equity markets face adverse selection, inventory holding, and order-processing costs. Cho and Engle (1999) and Kaul *et al.* (2004)

argue that, in addition to adverse selection, inventory holding, and order-processing costs, hedging costs play an important role in determining options spreads. Battalio and Schultz (2011) show that options market-makers' exposure to adverse selection and inventory risks is typically larger. First, the options market-makers' inventory positions can be highly volatile due to the implicit leverage in options contracts and uncertainty relating to stock return volatility (see also Jameson & Wilhelm 1992). Second, the options market-makers have limited control over their inventory positions due to options market dynamics. For example, traders are more likely to write call options than buy them, while they use buy and sell orders relatively evenly in equity markets (see Lakonishok *et al.* 2007). For these reasons, options market-makers hedge their inventories by taking an offsetting position in the underlying cash market.

In a discrete time setting, options market-makers' hedging costs consist of two components: the cost of setting up and liquidating the initial delta-neutral position and the cost of continuously rebalancing the portfolio and maintaining a delta-neutral position (see Jameson & Wilhelm 1992; Cho & Engle 1999; Kaul *et al.* 2004; Engle & Neri 2010). Kaul *et al.* (2004) show that to the extent that options market-makers employ market orders to hedge their inventories, their hedging costs are proportional to the stock market bid-ask spread. In addition to the bid-ask spread, the rebalancing component of the hedging costs is positively related to the volatility of the underlying asset and the sensitivity of the option to changes in underlying volatility ($option\ vega, v$) is inversely related to the revision interval.

The magnitude of the options market-makers' hedging costs differs by the moneyness of the options contract. For example, ITM options have the highest absolute delta and hence the highest cost associated with the setup and unwinding of the initial hedge. On the other hand, ATM options have the highest gamma and vega, and hence the market-makers' positions in these contracts need to be rebalanced/hedged much more frequently (see Kaul *et al.* 2004). Consequently, ATM options have higher rebalancing costs than ITM and OTM options, and

ITM options have higher setup and unwinding costs for the delta-neutral position than ATM and OTM options (see Wu *et al.* 2014). This suggests that OTM options have the lowest initial and rebalancing hedging costs. Notably, as reported in Table 6, we also find that the economic magnitudes of the impact of HFT on options spreads are the lowest for the OTM contracts, suggesting that the *hedging* channel is a plausible explanation for our findings on the impact of HFTs on options spreads.

The hedging channel discussed above may not be the only channel through which HFT activity in the stock market affects options market spreads. The payoffs of a stock and its listed option contracts are correlated through put-call parity. This parity relationship states that a portfolio that consists of a short put option and a long call option – both with the same strike price and maturity date – will have the same return as holding a forward contract with the same strike price and maturity. If this relationship does not hold, there will be a violation of the law of one price, implying the existence of an arbitrage opportunity between the stock and options markets (see Galai 1978; Halpern & Turnbull 1985; Ofek *et al.* 2004).[11]

Options market-makers are particularly exposed to such toxic arbitrage losses as the exchange-imposed caps on the number of quote updates and fines on traders with higher message-to-transaction ratios (see Muravyev & Pearson 2020) severely limit the options market-makers' ability to update their quotes in response to new information. Therefore, liquidity-consuming HFTs engaging in cross-market arbitrage strategies may exploit the put-call parity relation violations by sniping stale quotes in the options market. Halpern and Turnbull (1985) and Galai (1978) observe that the frequency of profitable put-call parity violations in ITM contracts is due to the prices of such options following stock prices very closely and the resulting difficulty in keeping the option and stock prices arbitrage-free in these

[11] In the case of American options, the above-mentioned put-call parity relation is typically expressed as an inequality due to the early-exercise premium of American calls and puts.

contracts. Consistent with these findings, our results also suggest that the impact of HFT on options spread (especially for the dollar spread) is relatively higher for ITM options. Thus, the *arbitrage* channel might also be a potential channel to explain the association between HFT and options spreads.

The main implication of the hedging and arbitrage channels is that they predict heterogeneity in the impact of various HFT strategies (liquidity-supplying and -demanding) on options spreads. In the hedging channel, HFT activity in the stock market likely affects options market-makers' hedging costs by affecting the stock bid-ask spread. On the one hand, liquidity-supplying HFTs in the stock market rely on their speed advantage to better manage the adverse-selection and inventory-holding risks (see, as an example, Brogaard *et al.* 2015). On the other hand, liquidity-consuming HFTs may pick off slower traders and impose adverse-selection costs on liquidity providers (see, as an example, Shkilko & Sokolov 2020). The resulting impact on the bid-ask spread may affect the options market-makers' hedging costs positively or negatively depending on the underlying strategy employed by the HFT firms.

Further, the bid-ask spread may not fully capture options market-makers' hedging costs as it only captures transaction costs for small orders. Lee (2008) estimates that more than half of the orders in the options markets originate from institutional investors. This, combined with the fact that options market-makers typically hedge their entire inventory of options, may induce them to employ more complex execution strategies involving the use of limit and market orders and order-splitting algorithms to minimize their transaction costs. Market-making HFTs may be more willing to supply liquidity to options market-makers due to their uninformed nature and the ability of HFTs to continuously reprice their orders. On the other hand, the predictability of options market-makers' hedging demand can allow liquidity-consuming HFTs to exploit the intraday (temporary) price impact generated by options market-makers' hedging demand (see Van Kervel & Menkveld 2019; Yang & Zhu 2020).

In the arbitrage channel, liquidity-demanding HFTs engaging in "toxic" arbitrage strategies may expose dealers to the risk of trading at stale quotes and force them to charge a larger bid-ask spread in options markets. Given that HFTs' liquidity-demanding and -supplying strategies are expected to have different effects on options spreads via the hedging and arbitrage channels, testing this hypothesis requires us to use more granular HFT data. More explicitly, we need to deconstruct HFT activity into liquidity-demanding and -supplying components. Unfortunately, the SEC's MIDAS data is not granular at this level. The MIDAS supplied proxies are general measures of HFTs and include the effects of both liquidity-demanding and liquidity-supplying HFT activities. Therefore, in the next section, we employ a more granular dataset – the NASDAQ HFT dataset – to test these two channels.

5. Heterogenous HFT Strategies and their Impact on Options Spread: Hedging and Arbitrage Channels

5.1. Description of the NASDAQ HFT Data

The NASDAQ HFT dataset is used to analyze the two proposed channels (hedging and arbitrage) to explain the role of HFTs in options market-making. This data contains transactions for 120 randomly selected NASDAQ and NYSE-listed stocks trading in 2009. The dataset stamps transactions into those initiated by HFTs and non-HFTs.[12] The following variables are included: date, time (in milliseconds), trading volume, price, buy-sell indicator, and the liquidity nature of the two sides to each trade (HH, HN, NH, and NN). HH refers to a trade in which both liquidity-providers and -takers are HFTs. HN (NH) implies that an HFT (a non-HFT) demands liquidity, and a non-HFT (HFT) supplies liquidity. Finally, NN indicates a transaction between two non-HFTs demanding and supplying liquidity. Consistent with Brogaard *et al.* (2014), we define the sum of HH, HN, and NH as the total HFT volume. The

[12] The disaggregation process is done by the NASDAQ. Brogaard *et al.* (2014) provide full details of the disaggregation.

total trading volume is about 44,800 million shares, for which 31,968 million or 71.30% have HFTs as counterparties. The total value of all trades is 1,381 billion USD.

We construct HFTs' liquidity-demanding ($SHFT^{D}_{i,d}$) and liquidity-supplying ($SHFT^{S}_{i,d}$) trades from the NASDAQ dataset. $SHFT^{D}_{i,d}$ ($SHFT^{S}_{i,d}$) is computed as the sum of HH and HN (HH and NH) divided by the total trading volume. $SHFT^{All}_{i,d}$ is the ratio of the total HFT trading volume (HH, HN, and NH) to the total trading volume. The summary statistics of the NASDAQ HFT data is provided in Table A.5.

While the NASDAQ dataset contains transactions for 120 stocks, we only use 103 stocks as we cannot match all options contracts in OptionMetrics with the corresponding stock in the NASDAQ dataset because of inconsistencies in ticker symbols across the two datasets. Thus, we obtain the option data for these 103 stocks in 2009.

5.2. Estimation Approaches

Similar to the baseline model, we first estimate the OLS approach with stock and time fixed effects:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 SHFT^{All}_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (4)$$

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 SHFT^{D}_{i,d} + \gamma_1 SHFT^{S}_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (5)$$

where $SHFT^{All}_{i,d}$, $SHFT^{D}_{i,d}$ and $SHFT^{S}_{i,d}$ are the measures of HFTs' total, liquidity-demanding and -supplying activities respectively and α_i and β_d are stock and time (day) fixed effects. All these variables are defined in Table 1.

As noted, in order to address endogeneity, we employ the 2SLS IV approach in the baseline model; the same approach is used for the following analysis:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{SHFT}^{All}_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (6)$$

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{SHFT}^{D}_{i,d} + \gamma_2 \widehat{SHFT}^{S}_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (7)$$

$$SHFT_{i,d}^{All} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (8)$$

$$SHFT_{i,d}^{D} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (9)$$

$$SHFT_{i,d}^{S} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d} \quad (10)$$

where $\widehat{SHFT}_{i,d}^{All}$, $\widehat{SHFT}_{i,d}^{D}$ and $\widehat{SHFT}_{i,d}^{S}$ are the fitted values of $SHFT_{i,d}^{All}$, $SHFT_{i,d}^{D}$ and $SHFT_{i,d}^{S}$ obtained by regressing the respective variables on $IV_{i,d}$. We again employ two sets of instruments for robustness. First, we use the NASDAQ HFT dataset from 2009, and any exogenous shock to the volume of HFT on the NASDAQ stock exchange during this period is a candidate for the instrument. A potential instrument satisfying these criteria is proposed by Skjeltorp *et al.* (2016). On June 5, 2009, the NASDAQ stock exchange introduced NASDAQ Only Flash Orders.[13]

The implementation details for these orders and some numerical examples are also provided by Skjeltorp *et al.* (2016). After an unsuccessful execution attempt in the NASDAQ limit order book, the NASDAQ gave an additional 500 milliseconds to its market participants and vendors to expose the orders before reaching the general marketplace. It is clear from the time constraint that only qualified low-latency traders, i.e., HFTs, are expected to use flash orders (see Harris & Namvar 2016).[14] This expectation is also consistent with the flash order implementation of Direct Edge – the first company to introduce flash orders on January 27, 2006 – which states that such orders allow brokers and HFTs to see and execute flash orders (see Skjeltorp *et al.* 2016). Thus, HFTs benefit from flash orders, such that the latter are expected to increase HFTs' participation. In this specification, $IV_{i,d}$ is a dummy equal to *one* from June 5, 2009 to August 31, 2009, and *zero* for the other periods (from January 1, 2009 to June 4, 2009, and from September 1, 2009 to December 31, 2009) in our sample. It is important

[13] https://www.nasdaqtrader.com/TraderNews.aspx?id=ETA2009-35

[14] Some regulators and investors have argued that flash orders give an unfair advantage to market participants who are able to use them. For example, Mary Schapiro, SEC Chairman, says that "flash orders have the potential to discourage publicly displayed trading interest and harm quote competition among markets".

to note that, in this specification, we only include stock fixed effect because our instrument does not have a time variation.

Second, similar to the main analysis, and inspired by Hasbrouck and Saar (2013), the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of three HFT proxies ($SHFT_{i,d}^{All}$, $SHFT_{i,d}^{D}$ and $SHFT_{i,d}^{S}$) in all other stocks in the corresponding size quintile.

5.3. Estimation Results

The estimation results are presented in Table 7.

INSERT TABLE 7 ABOUT HERE

The first-stage results of the 2SLS IV approach are reported in the Appendix section (Table A.6). The selected instruments are significantly related to $SHFT_{i,d}^{All}$, $SHFT_{i,d}^{D}$ and $SHFT_{i,d}^{S}$ and the signs of association are as expected. There are three important points in Table 7. First, consistent with the results based on the SEC's MIDAS data, HFT is positively and significantly related to options spreads. The magnitude of the increase in options spreads is also economically meaningful. For example, estimates using the OLS regression with fixed effects show that a one standard deviation increase in $SHFT_{i,d}^{All}$ raises $ODspread_{i,d}$ and $OPspread_{i,d}$ by about 14.7% and 3.0% respectively.

Second, HFTs' liquidity-demanding orders ($SHFT_{i,d}^{D}$) are positively and significantly (both statistically and economically) related to both options spread proxies ($OPspread_{i,d}$ and $ODspread_{i,d}$) for ATM and ITM contracts; for ITM contracts, the association between $SHFT_{i,d}^{D}$ and $OPspread_{i,d}$ is weakly significant (10% level). The economic magnitudes of the impact are again substantial. Specifically, a one standard deviation shock to $SHFT_{i,d}^{D}$ increases $OPspread_{i,d}$ by 5.71% and $ODspread_{i,d}$ by 13.89% for ATM options. For ITM options, the economic magnitudes are 6.22% for $OPspread_{i,d}$ and 9.66% for $ODspread_{i,d}$. The association between $SHFT_{i,d}^{D}$ and $Ospread_{i,d}$ is not significant for OTM options. These results are consistent with the *hedging* channel.

Second, liquidity-supplying orders initiated by HFTs ($SHFT_{i,d}^{S}$) are weakly significantly (negatively) related to both options spread proxies for ITM contracts only; this is not significant for ATM and OTM options. This result seems puzzling as it is expected that the hedging costs of the options market-maker will decrease with the HFTs' liquidity-supplying trading. We argue that this may be linked to the fact that we control for the stock spread in our regression setting. A significant portion of the effect of $SHFT_{i,d}^{S}$ on the options market spread can be captured by the bid-ask spread in the stock market. To test this argument, we estimate models 4 to 10 without controlling for the stock spread.

INSERT TABLE 8 ABOUT HERE

Panels A ($OPspread_{i,d}$) and B ($ODspread_{i,d}$) of Table 8 show the estimation results for models 4 to 10 without controlling for the stock spread. The results show that both $SHFT_{i,d}^{D}$ and $SHFT_{i,d}^{S}$ are statistically significantly related to $OPspread_{i,d}$ and $ODspread_{i,d}$ in this specification. Importantly, the associations are statistically significant across all moneyness groups, albeit weakly significant in some specifications. The results are economically meaningful as well. For instance, for ATM options, a one standard deviation increase in $SHFT_{i,d}^{S}$ ($SHFT_{i,d}^{D}$) is associated with a −5.00% (+6.43%) and −13.89% (+20.83%) change in $OPspread_{i,d}$ and $ODspread_{i,d}$ respectively.

Overall, the results imply that HFTs' effect on the options market-making is restricted to their liquidity-demanding orders in the underlying market after controlling for the stock spread. This can be explained by the fact that it is unlikely that the options market-maker observes who (HFT or non-HFT) is supplying liquidity and strategically chooses one over the other, as these are anonymous markets. However, this explanation raises an interesting question about why $SHFT_{i,d}^{D}$ is still significant. We argue that this is because the relationship between HFT in the stock market and the options market bid-ask spread is unlikely to be fully captured by including the stock market bid-ask spread as a control variable. Bid-ask spreads only capture the transaction costs associated with small orders and do not capture the price impact of large parent orders being split into smaller child orders.[15]

This is indeed a plausible explanation as Lee (2008) shows that half of the orders in the options markets originate from institutional investors who are generally splitting their orders (see Menkveld 2008; Chemmanur *et al.* 2010). It is also known that options market-makers typically hedge their entire inventory of options, and thus they may need to split their large orders to reduce their transaction costs. Liquidity-demanding HFTs employ different aggressive strategies (e.g., arbitrage, back-running, etc.) (see Brogaard *et al.* 2015), which allow them to profit from the price impact of options market-makers' large orders being split into child orders. This finding further suggests that the stock spread mainly includes the impact of HFTs' liquidity supply trades on the stock market liquidity. Therefore, the studies that investigate the role of HFTs in market quality using only the spread as a proxy for liquidity should be interpreted with caution.

As noted, in addition to the hedging channel discussed and tested above, the arbitrage channel, which suggests that the options market-makers can charge a higher bid-ask spread due to liquidity-consuming HFTs engaging in cross-market arbitrage strategies between stock and options markets, may also explain the impact of HFT activity in the stock market affects options market spreads. In the next analysis, we test this channel.

5.3.1. The Arrival of New Information and HFT

[15] The literature examining the effect of HFT in the stock market similarly finds that HFT simultaneously leads to lower bid-ask spreads (see Hendershott *et al.* 2011; Brogaard *et al.* 2015) and higher execution costs for large orders (see Korajczyk & Murphy 2019; Van Kervel & Menkveld 2019).

Foucault *et al.* (2017) argue that arbitrage opportunities arising due to the asynchronous adjustments in the price of correlated assets are "toxic" and lead to increased bid-ask spreads if they result from the arrival of new information.[16] Inspired by this, to test the arbitrage channel and investigate the relative importance of the hedging and arbitrage channels, we study the differences in the impact of HFT on the options market spread during news and no-news days.

We follow Hirschey (2020) to identify the days with firm news. First, we use *Factiva*, which contains news from over 35,000 sources, and identify the news days for each firm. As suggested by Hirschey (2020), traders may respond to signals not covered by these sources. Therefore, we compute absolute market-adjusted returns in a second step and exclude days on which they were greater than specific thresholds (1%, 0.5%, and 0.25%).[17]

INSERT TABLE 9 ABOUT HERE

Panels A and B in Table 9 present the model's estimation results for days with no firm news. The results show that $SHFT_{i,d}^{D}$ has a positive and statistically significant influence on the ATM options spread, even on these days. Thus, the hedging channel is more dominant for ATM options as options market-makers need to rebalance their entire inventory positions. In contrast, the relationship between $SHFT_{i,d}^{D}$ and the ITM options spread loses its significance once we exclude days with firm news. Halpern and Turnbull (1985) and Galai (1978) document that the profitable arbitrage opportunities arising from violations' put-call parity are more frequently observed in ITM options. Therefore, the association between HFT and options spreads disappears on days with fewer arbitrage opportunities (i.e., without firm news), suggesting that the *arbitrage* channel is more dominant for ITM contracts.

---

[16] This is also consistent with Rzayev and Ibikunle (2019) and Brolley and Zoican (2020), who show that HFTs can make profits at slow traders' expense due to their ability to react faster to public news (i.e., *latency arbitrage*).
[17] Our sample period covers 2009. Due to the 2008 financial crisis, there may be some abnormal returns during our sample period not related to any specific information. Therefore, for robustness, in the second specification, we exclude days with market-adjusted returns higher than the mean of the market-adjusted returns for the sample period. The results obtained are qualitatively similar to those we report.

Importantly, we do not claim that the arbitrage channel is the only channel that can explain the association between HFT and ITM options spreads. This is because the NASDAQ HFT data contains a limited number of stocks, and it covers one year – 2009. The data also includes HFT activity in NASDAQ only. In the main analysis, where we use the comprehensive panel data from the SEC's MIDAS database, we find that the economic magnitude of the effect of HFT on options spreads is quite substantial for ITM options. Thus, it is plausible to expect that both hedging and arbitrage channels contribute to the association between HFT and the costs of market-making in ITM options. We nevertheless believe that the fact that the relationship between $SHFT_{i,d}^{D}$ and the ITM options spread loses its significance on no-news days gives us some insights into the association between HFT activity and ITM options spreads. Thus, this should be of general interest.

6. Conclusion

This study uses a comprehensive sample of HFT data provided by the SEC MIDAS to investigate the relationship between HFT activity in the stock market and its impact on options market-making. We find that HFT in the stock market leads to increased bid-ask spread and deterioration of liquidity in options markets. We propose two channels to explain this finding: (i) hedging channel; and (ii) arbitrage channel. The hedging channel suggests that the options market-makers' hedging costs increase (decrease) due to higher (lower) bid-ask spread and the price impact-sourced HFTs' liquidity-demanding trades. On the other hand, the arbitrage channel implies that HFTs' aggressive (liquidity-demanding) trades expose the options market-maker to the risk of trading at stale prices upon the arrival of new information.

We test these channels by using proprietary HFT data from NASDAQ. Our findings suggest that the hedging channel dominates for ATM options. On the other hand, the significant relationship between HFT and the options spread might be mainly driven by the arbitrage channel for ITM options. For OTM contracts, the economic magnitude of the impact of HFT on options spreads is lower in the main analysis and not statistically significant based on the NASDAQ HFT data.

Our findings also suggest that while HFTs' liquidity-supplying trades increase the options market-makers' hedging abilities by improving liquidity in the underlying markets, the significant relationship between HFTs' liquidity-supply trades and options market liquidity disappears after controlling for the stock spread.

The results of this paper highlight the need for a better understanding of the costs/risks due to HFTs in today's highly fragmented and complicated market structures. Our findings in particular suggest that practitioners, academics, and policy makers should carefully consider the cross-asset effects of HFTs' activities in equity markets on derivative market quality.

References

Albuquerque, R., Song, S., Yao, C., 2020. The price effects of liquidity shocks: A study of the SEC's tick size experiment. Journal of Financial Economics 138, 700-724

Battalio, R., Schultz, P., 2011. Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets. The Journal of Finance 66, 2013-2053

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of political economy 81, 637-654

Bollen, N.P., Whaley, R.E., 2004. Does net buying pressure affect the shape of implied volatility functions? The Journal of Finance 59, 711-753

Boyle, P.P., Vorst, T., 1992. Option replication in discrete time with transaction costs. The Journal of Finance 47, 271-293

Brenner, M., Eldor, R., Hauser, S., 2001. The price of options illiquidity. The Journal of Finance 56, 789-805

Brogaard, J., Hagströmer, B., Nordén, L., Riordan, R., 2015. Trading fast and slow: Colocation and liquidity. The Review of Financial Studies 28, 3407-3443

Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. The Review of Financial Studies 27, 2267-2306

Brolley, M., Zoican, M., 2020. Liquid speed: A congestion fee for low-latency exchanges. Rotman School of Management Working Paper

Brunnermeier, M.K., Pedersen, L.H., 2005. Predatory trading. The Journal of Finance 60, 1825-1863

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

Cao, C., Chen, Z., Griffin, J.M., 2005. Informational content of option volume prior to takeovers. The Journal of Business 78, 1073-1109

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., Vega, C., 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. The Journal of Finance 69, 2045-2084

Chemmanur, T.J., Hu, G., Huang, J., 2010. The role of institutional investors in initial public offerings. The Review of Financial Studies 23, 4496-4540

Cho, Y.-H., Engle, R.F., 1999. Modeling the impacts of market activity on bid-ask spreads in the option market. National Bureau of Economic Research

Christoffersen, P., Goyenko, R., Jacobs, K., Karoui, M., 2017. Illiquidity Premia in the Equity Options Market. The Review of Financial Studies 31, 811-851

Chung, K.H., Lee, A.J., Rösch, D., 2020. Tick size, liquidity for small and large orders, and price informativeness: Evidence from the Tick Size Pilot Program. Journal of Financial Economics 136, 879-899

Collin-Dufresne, P., Fos, V., Muravyev, D., 2020. Informed trading and option prices: Theory and evidence from activist trading. Journal of Financial and Quantitative Analysis Forthcoming

Comerton-Forde, C., Putniņš, T.J., 2015. Dark trading and price discovery. Journal of Financial Economics 118, 70-92

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

Cremers, M., Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. Journal of Financial and Quantitative Analysis, 335-367

Easley, D., O'Hara, M., Srinivas, P.S., 1998. Option volume and stock prices: Evidence on where informed traders trade. The Journal of Finance 53, 431-465

Engle, R., Neri, B., 2010. The impact of hedging costs on the bid and ask spread in the options market. Unpublished Working Paper, New York University

Foley, S., Putniņš, T.J., 2016. Should we be afraid of the dark? Dark trading and market quality. Journal of Financial Economics 122, 456-481

Foucault, T., Fresard, L., 2014. Learning from peers' stock prices and corporate investment. Journal of Financial Economics 111, 554-577

Foucault, T., Hombert, J., Roşu, I., 2016. News trading and speed. The Journal of Finance 71, 335-382

Foucault, T., Kozhan, R., Tham, W.W., 2017. Toxic arbitrage. The Review of Financial Studies 30, 1053-1094

Galai, D., 1978. Empirical tests of boundary conditions for CBOE options. Journal of Financial Economics 6, 187-211

Ge, L., Lin, T.-C., Pearson, N.D., 2016. Why does the option to stock volume ratio predict stock returns? Journal of Financial Economics 120, 601-622

Glosten, L.R., Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71-100

Hagströmer, B., Nordén, L., 2013. The diversity of high-frequency traders. Journal of Financial Markets 16, 741-770

Halpern, P.J., Turnbull, S.M., 1985. Empirical tests of boundary conditions for Toronto Stock Exchange options. The Journal of Finance 40, 481-500

Harris, L., Namvar, E., 2016. The economics of flash orders and trading. Journal of Investment Management, Fourth Quarter

Hasbrouck, J., Saar, G., 2013. Low-latency trading. Journal of Financial Markets 16, 646-679

Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? The Journal of Finance 66, 1-33

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

Hu, J., 2014. Does option trading convey stock price information? Journal of Financial Economics 111, 625-645

Jameson, M., Wilhelm, W., 1992. Market making in the options markets and the costs of discrete hedge rebalancing. The Journal of Finance 47, 765-779

Jiang, G.J., Lo, I., Valente, G., 2014. High-frequency trading around macroeconomic news announcements: Evidence from the US treasury market. Bank of Canada Working Paper

Kapadia, N., Linn, M., 2019. Option Spreads and the Uncertain Cost of Equity Liquidity: Evidence from the Knight Capital Trading Glitch. Available at SSRN 3517429

Karpoff, J.M., 1987. The relation between price changes and trading volume: A survey. Journal of Financial and Quantitative Analysis 22, 109-126

Kaul, G., Nimalendran, M., Zhang, D., 2004. Informed trading and option spreads. Available at SSRN 547462

Kirilenko, A., Kyle, A.S., Samadi, M., Tuzun, T., 2017. The Flash Crash: High-frequency Trading in an Electronic Market. The Journal of Finance 72, 967-998

Korajczyk, R.A., Murphy, D., 2019. High-Frequency Market Making to Large Institutional Trades. The Review of Financial Studies 32, 1034-1067

Kozhan, R., Tham, W.W., 2012. Execution risk in high-frequency arbitrage. Management Science 58, 2131-2149

Lakonishok, J., Lee, I., Pearson, N.D., Poteshman, A.M., 2007. Option market activity. The Review of Financial Studies 20, 813-857

Lee, C.M., Watts, E.M.J.T.A.R., 2021. Tick size tolls: Can a trading slowdown improve earnings news discovery? 96, 373-401

Lee, E.J., 2015. High frequency trading in the Korean index futures market. Journal of Futures Markets 35, 31-51

Lee, S., 2008. U.S. equity options market: Changing competitive landscape (November). Aite Group, Boston

Leland, H.E., 1985. Option pricing and replication with transactions costs. The journal of finance 40, 1283-1301

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

Menkveld, A.J., 2008. Splitting orders in overlapping markets: A study of cross-listed stocks. Journal of Financial Intermediation 17, 145-174

Menkveld, A.J., 2013. High frequency trading and the new market makers. Journal of Financial Markets 16, 712-740

Menkveld, A.J., Zoican, M.A., 2017. Need for speed? Exchange latency and liquidity. The Review of Financial Studies 30, 1188-1228

Mishra, S., Daigler, R.T., Holowczak, R., 2012. The effect of algorithmic trading on liquidity in the options market. Citeseer

Muravyev, D., Pearson, N.D., 2020. Options Trading Costs Are Lower than You Think. The Review of Financial Studies Forthcoming

Ni, S.X., Pan, J., Poteshman, A.M., 2008. Volatility information trading in the option market. The Journal of Finance 63, 1059-1091

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

Ofek, E., Richardson, M., Whitelaw, R.F., 2004. Limited arbitrage and short sales restrictions: Evidence from the options markets. Journal of Financial Economics 74, 305-342

Pan, J., Poteshman, A.M., 2006. The information in option volume for future stock prices. The Review of Financial Studies 19, 871-908

Rzayev, K., Ibikunle, G., 2019. A state-space modeling of the information content of trading volume. Journal of Financial Markets

Shkilko, A., Sokolov, K., 2020. Every cloud has a silver lining: Fast trading, microwave connectivity and trading costs. Journal of Finance Forthcoming

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

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

Wei, J., Zheng, J., 2010. Trading activity and bid–ask spreads of individual equity options. Journal of Banking & Finance 34, 2897-2916

Weller, B.M., 2018. Does algorithmic trading reduce information acquisition? The Review of Financial Studies 31, 2184-2226

Wu, W.-S., Liu, Y.-J., Lee, Y.-T., Fok, R.C., 2014. Hedging costs, liquidity, and inventory management: The evidence from option market makers. Journal of Financial Markets 18, 25-48

Yang, L., Zhu, H., 2020. Back-running: Seeking and hiding fundamental information in order flows. The Review of Financial Studies 33, 1484-1533

Table 1. Definitions of variables
This table reports the notation, description, and source of variables. The units of the variables are in parentheses following the variable names. Panel A reports the equity market variables and Panel B reports options market variables.

Panel A. Equity market variables		
Variable	Description	Data source
$QT_{i,d}$	Quote-to-trade ratio for firm i on day d is computed as the sum of order volume for all order messages divided by the sum of trade volume for all trades that are not against hidden orders.	MIDAS
$CT_{i,d}$ (%)	Cancel-to-trade for firm i on day d is computed as the number of all cancel messages (full or partial) divided by the number of all trade messages.	MIDAS
$OR_{i,d}$ (%)	Odd lot rate for firm i on day d is computed as the number of odd lot trade messages for all exchanges divided by the number of trades from exchange that report individual trades.	MIDAS
$OV_{i,d}$ (%)	Odd lot rate for firm i on day d is computed as the sum of odd lot trade volume for all exchanges divided by the sum of trade volume from exchange that report individual trades.	MIDAS
$ITS_{i,d}$	Inverse trade size for firm i on day d is computed as the number of trades divided by the number of shares traded.	MIDAS
$SHFT_{i,d}^{All}$ (%)	HFTs' total trades percentage for stock *i* and day *d* computed as the ratio of the HFTs' total trading volume (the sum of HH, HN, and NH) to the total trading volume.	NASDAQ
$SHFT_{i,d}^{D}$ (%)	HFTs' liquidity-demanding trades percentage for stock *i* and day *d* computed as the ratio of the HFTs' liquidity-demanding trading volume (the sum of HH and HN) to the total trading volume.	NASDAQ
$SHFT_{i,d}^{S}$ (%)	HFTs' liquidity-supplying trades percentage for stock *i* and day *d* computed as the ratio of the HFTs' liquidity-supplying trading volume (the sum of HH and NH) to the total trading volume.	NASDAQ
$SDspread_{i,d}$	Stock dollar spread for firm i on day d is computed as the difference between best-ask and bid prices for day *d*.	CRSP
$SPspread_{i,d}$ (%)	Stock proportional spread for firm i on day d is computed as bid-ask spread divided by bid-ask midpoint for stock i each day.	CRSP
$SVolatility_{i,d}$	Stock price volatility for firm i on day t is computed as the absolute value of the difference between mid-prices for days *d* and *d-1*.	CRSP
Panel B. Option market variables		
$ODspread_{i,d}$	Option market dollar spread for stock *i* and day *d* computed as the dollar-volume-weighted average of the dollar spread (the difference between the best-ask and bid prices).	OptionMetrics
$OPspread_{i,d}$ (%)	Option market proportional spread for stock *i* and day *d* computed as the dollar-volume-weighted average of the proportional spread (the difference between the best-ask and bid prices divided by the midpoint of the ask and bid prices).	OptionMetrics
$Ovolume_{i,d}$	Option volume for stock *i* and day *d* computed as the natural logarithm of the daily trading volume (contracts)	OptionMetrics
$Oimplied_{i,d}$	Option-implied volatility for stock *i* and day *d* computed as the dollar-volume-weighted average of the implied volatility provided by OptionMetrics.	OptionMetrics
$\lvert Odelta_{i,d} \rvert$	Absolute option delta for stock *i* and day *d* computed as the dollar-volume-weighted average of the absolute delta provided by OptionMetrics.	OptionMetrics
$Ogamma_{i,d}$	Option gamma for stock *i* and day *d* computed as the dollar-volume-weighted average of the gamma provided by OptionMetrics.	OptionMetrics
$Ovega_{i,d}$	Option vega for stock *i* and day *d* computed as the dollar-volume-weighted average of the vega provided by OptionMetrics.	OptionMetrics

Table 2. Summary statistics for MIDAS sample

This table reports the descriptive statistics for the variables used in our analysis. Panel A shows the descriptive statistics for all variables from the underlying stock market. Panel B provides the descriptive statistics for all options-related variables separately for the full sample and three groups based on moneyness. For the definitions and computation methods of the variables, see Table 1. We follow Bollen and Whaley (2004) and define OTM options as those with absolute option delta $|\Delta| \leq 0.375$, ATM options as those with $0.375 < |\Delta| \leq 0.625$, and ITM options as those with $|\Delta| > 0.625$. In both panels, we have two samples: (i) the full sample; and (ii) the tick size pilot sample. The full sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. The tick size pilot sample contains 1,235 stocks (617 control stocks and 618 treated stocks), including the Tick Size Pilot Program launched by the SEC. The Tick Size Pilot Program commenced on October 3, 2016, and was implemented over the following two years. Therefore, the tick size pilot sample covers October 1, 2014 to September 28, 2018.

Panel A. Equity market variables				
	Variable	Mean	Median	St dev
Full sample	$QT_{i,d}$	42.07	33.73	29.43
	$CT_{i,t}$ (%)	23.15	19.69	13.51
	$OR_{i,t}$ (%)	37.21	35.07	17.44
	$OV_{i,t}$ (%)	15.87	14.09	9.81
	$ITS_{i,t}$	0.01	0.01	0.01
	$SDspread_{i,t}$	0.02	0.01	0.03
	$SPspread_{i,t}$ (%)	0.11	0.05	0.18
	$SVolatility_{i,t}$	0.69	0.35	1.02
Tick size pilot sample	$QT_{i,d}$	44.04	32.24	36.81
	$CT_{i,t}$ (%)	24.18	19.68	16.15
	$OR_{i,t}$ (%)	39.15	38.33	14.59
	$OV_{i,t}$ (%)	16.83	15.96	8.35
	$ITS_{i,t}$	0.01	0.01	0.01
	$SDspread_{i,t}$	0.03	0.02	0.03
	$SPspread_{i,t}$ (%)	0.18	0.09	0.26
	$SVolatility_{i,t}$	0.60	0.32	0.81
Panel B. Option market variables				
Full sample	$ODspread_{i,t}$	0.45	0.27	0.55
	$OPspread_{i,t}$ (%)	34.07	19.63	40.75
	$Ovolume_{i,t}$	5.06	4.96	2.59
	$Oimplied_{i,t}$	0.45	0.36	0.28
	$\|Odelta_{i,t}\|$	0.51	0.50	0.17
	$Ogamma_{i,t}$	0.12	0.08	0.12
	$Ovega_{i,t}$	5.98	4.10	6.25
Tick size pilot sample	$ODspread_{i,t}$	0.60	0.40	0.62
	$OPspread_{i,t}$ (%)	50.44	32.43	47.79
	$Ovolume_{i,t}$	3.62	3.58	1.95
	$Oimplied_{i,t}$	0.52	0.43	0.28
	$\|Odelta_{i,t}\|$	0.50	0.49	0.19
	$Ogamma_{i,t}$	0.12	0.10	0.10
	$Ovega_{i,t}$	4.06	2.86	3.79
	$ODspread_{i,t}$	0.38	0.23	0.47
	$OPspread_{i,t}$ (%)	21.64	13.57	25.30

ATM	$Ovolume_{i,t}$	4.53	4.54	2.42
	$Oimplied_{i,t}$	0.40	0.33	0.23
	$\lvert Odelta_{i,t} \rvert$	0.49	0.49	0.05
	$Ogamma_{i,t}$	0.13	0.09	0.12
	$Ovega_{i,t}$	8.23	5.79	8.29
ITM	$ODspread_{i,t}$	0.62	0.39	0.71
	$OPspread_{i,t}$ (%)	15.73	10.57	16.38
	$Ovolume_{i,t}$	3.81	3.64	2.31
	$Oimplied_{i,t}$	0.48	0.38	0.31
	$\lvert Odelta_{i,t} \rvert$	0.78	0.78	0.08
	$Ogamma_{i,t}$	0.11	0.07	0.11
	$Ovega_{i,t}$	4.86	3.11	0.11
OTM	$ODspread_{i,t}$	0.27	0.17	0.32
	$OPspread_{i,t}$ (%)	52.10	32.50	52.08
	$Ovolume_{i,t}$	4.34	4.17	2.50
	$Oimplied_{i,t}$	0.43	0.35	0.26
	$\lvert Odelta_{i,t} \rvert$	0.25	0.26	0.06
	$Ogamma_{i,t}$	0.10	0.06	0.10
	$Ovega_{i,t}$	5.93	4.13	6.05

Table 3. The relationship between control variables and option spread – MIDAS sample

This table presents the results for the estimation of the association between control variables and the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the option and underlying markets. The option market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. The sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

	$ODspread_{i,t}$	$OPspread_{i,t}$
$Ovolume_{i,d}$	-0.24***	-25.55***
	(-256.37)	(-351.10)
$Oimplied_{i,d}$	0.23***	13.44***
	(211.08)	(187.84)
$\|Odelta_{i,d}\|$	0.42***	21.18***
	(455.23)	(316.27)
$Ogamme_{i,d}$	-0.09***	9.56***
	(-224.78)	(170.34)
$Ovega_{i,d}$	0.16***	-13.72***
	(191.56)	(301.59)
$SPspread_{i,d}$		2.38***
		(51.55)
$SDspread_{i,d}$	0.05***	
	(78.00)	
$SVolatility_{i,d}$	0.06***	3.89***
	(132.43)	(170.27)
Time and stock FEs	Yes	Yes
N	2,969,829	2,969,829

Table 4. The impact of HFT on options spread – MIDAS sample
This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT_{i,d}} + \sum\nolimits_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$
$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum\nolimits_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), and $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$). α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the option and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Column 1, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 3, we use the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks, and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $ODspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.02*** (28.42)	0.02*** (10.95)	0.03*** (25.16)
	$CT_{i,d}$	0.01*** (15.52)	0.02*** (11.12)	0.01*** (10.10)
	$OR_{i,d}$	0.07*** (60.39)	0.02*** (6.85)	0.11*** (65.74)
	$OV_{i,d}$	0.05*** (54.45)	0.02*** (8.52)	0.09*** (62.46)
	$ITS_{i,d}$	0.05*** (41.60)	0.02*** (6.30)	0.08*** (44.01)
	Controls	Yes	Yes	Yes
	Time and Stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Panel B: $OPspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.89*** (19.51)	1.46*** (11.19)	1.31*** (17.98)
	$CT_{i,d}$	0.60*** (11.99)	1.35*** (11.59)	0.96*** (11.65)
	$OR_{i,d}$	2.76*** (32.88)	1.67*** (5.81)	3.59*** (28.06)
	$OV_{i,d}$	1.89*** (27.72)	1.50*** (7.17)	2.61*** (25.39)
	$ITS_{i,d}$	1.13*** (12.45)	1.54*** (5.18)	2.53*** (19.84)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Table 5. The impact of HFT on options spread by moneyness – MIDAS sample

This table presents the results for the estimation of the impact of HFT on options spread by moneyness:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT}_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$ and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Column 1, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 3, we use the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. We follow Bollen and Whaley (2004) and define OTM options as those with absolute option delta $|\Delta| \leq 0.375$, ATM options as those with $0.375 < |\Delta| \leq 0.625$, and ITM options as those with $|\Delta| > 0.625$. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $ODspread_{i,t}$ is the dependent variable.

	ATM			ITM			OTM		
$HFT_{i,t}$	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)	OLS (4)	IV (Tick size pilot) (5)	IV (Average HFT) (6)	OLS (7)	IV (Tick size pilot) (8)	IV (Average HFT) (9)
$QT_{i,t}$	0.02***	0.02***	0.02***	0.04***	0.05***	0.06***	0.01***	0.01***	0.02***
	(25.32)	(10.37)	(22.86)	(39.46)	(10.49)	(33.88)	(30.26)	(12.32)	(27.87)
$CT_{i,t}$	0.01***	0.02***	0.02***	0.01***	0.02***	0.01***	0.01***	0.01***	0.01***
	(19.53)	(9.79)	(14.33)	(14.41)	(7.64)	(6.07)	(24.73)	(12.60)	(20.67)
$OR_{i,t}$	0.04***	0.03***	0.07***	0.18***	0.08***	0.27***	0.02***	0.01***	0.03***
	(34.90)	(5.13)	(40.63)	(107.91)	(4.34)	(110.07)	(28.02)	(4.51)	(33.89)
$OV_{i,t}$	0.03***	0.03***	0.05***	0.15***	0.09***	0.23***	0.01***	0.01***	0.02***
	(27.83)	(7.08)	(35.95)	(104.94)	(6.75)	(109.69)	(19.02)	(7.55)	(27.01)
$ITS_{i,t}$	0.02***	0.03***	0.04***	0.15***	0.15***	0.22***	0.003***	0.01***	0.01***
	(17.65)	(5.32)	(21.40)	(81.49)	(8.84)	(82.73)	(4.71)	(5.61)	(9.04)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Time and stock FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,266,764	402,656	2,264,040	2,080,605	343,460	2,077,881	2,543,347	480,324	2,540,617

Panel B: $OPspread_{i,t}$ is the dependent variable.

	ATM			ITM			OTM		
$HFT_{i,t}$	OLS (1)	IV (Tick Size Pilot) (2)	IV (Average HFT) (3)	OLS (4)	IV (Tick Size Pilot) (5)	IV (Average HFT) (6)	OLS (7)	IV (Tick Size Pilot) (8)	IV (Average HFT) (9)
$QT_{i,t}$	1.10***	1.24***	1.82***	0.75***	1.30***	1.19***	1.28***	2.35***	2.28***
	(30.47)	(9.38)	(32.13)	(30.74)	(10.27)	(31.19)	(22.80)	(11.19)	(25.60)
$CT_{i,t}$	1.21***	0.97***	2.04***	0.64***	0.62**	1.03***	1.63***	2.24***	3.12***
	(30.53)	(9.06)	(31.64)	(24.00)	(7.63)	(23.90)	(26.55)	(11.90)	(31.02)
$OR_{i,t}$	2.61***	1.28***	4.00***	1.78***	1.85***	3.02***	0.92***	1.66***	2.05***
	(38.34)	(3.29)	(39.78)	(40.27)	(4.79)	(46.83)	(8.80)	(5.42)	(13.31)
$OV_{i,t}$	1.70***	1.46***	2.80***	1.21***	2.35***	2.12***	-0.09	1.57***	0.66***
	(30.84)	(5.34)	(34.67)	(33.83)	(8.06)	(41.30)	(-1.02)	(5.51)	(5.35)
$ITS_{i,t}$	1.55***	1.09***	3.10***	0.97***	4.44***	1.99***	-1.36***	1.72***	-0.19
	(20.71)	(2.67)	(30.46)	(19.45)	(7.78)	(29.69)	(-12.06)	(4.17)	(-1.25)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time and stock FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,266,764	402,656	2,264,040	2,080,605	343,460	2,077,881	2,543,347	480,324	2,540,617

Table 6. Economic effect of HFT on options spreads – MIDAS sample

This table presents the economic effect of a one standard deviation shock to HFT proxies on the options spreads:

$$Economic\ effect = \gamma_1/\mu(Ospread_{i,s}),$$

where γ_1 is the coefficient of various HFT proxies obtained from estimation of the regression, and $\mu(Ospread_{i,t})$ is the average value of various $ODspread_{i,t}$ and $OPspread_{i,t}$. Panel A (B) shows the economic magnitude of the impact of HFT on the $ODspread_{i,t}$ ($OPspread_{i,t}$). The first column shows the results for a standard OLS approach with stock and day fixed effects, whilst the second and third columns present the results for the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. We follow Bollen and Whaley (2004) and define OTM options as those with $|Odelta_{i,t}| \leq 0.375$, ATM options as those with $0.375 < |Odelta_{i,t}| \leq 0.625$, and ITM options as those with $|Odelta_{i,t}| > 0.625$. For the definitions and computation methods of all the variables, see Table 1.

Panel A: $\boldsymbol{ODspread_{i,t}}$ is the dependent variable.

	OLS (1)				IV (Tick size pilot) (2)				IV (Average HFT) (3)			
$HFT_{i,t}$	Full	ATM	ITM	OTM	Full	ATM	ITM	OTM	Full	ATM	ITM	OTM
$QT_{i,t}$	4.44%	5.26%	6.45%	3.70%	3.33%	3.64%	5.81%	2.63%	6.67%	5.26%	9.68%	7.41%
$CT_{i,t}$	2.22%	2.63%	1.61%	3.70%	3.33%	3.64%	2.33%	2.63%	2.22%	5.26%	1.61%	3.70%
$OR_{i,t}$	15.56%	10.53%	29.03%	7.41%	3.33%	5.45%	9.30%	2.63%	24.44%	18.42%	43.55%	11.11%
$OV_{i,t}$	11.11%	7.89%	24.19%	3.70%	3.33%	5.45%	10.47%	2.63%	20.00%	13.16%	37.10%	7.41%
$ITS_{i,t}$	11.11%	5.26%	24.19%	1.11%	3.33%	5.45%	17.44%	2.63%	17.78%	10.53%	35.48%	3.70%
Average	8.89%	6.32%	17.10%	3.93%	3.33%	4.73%	9.07%	2.63%	14.22%	10.53%	25.48%	6.67%

Panel B: $OPspread_{i,t}$ is the dependent variable.

	OLS (1)				IV (Tick size pilot) (2)				IV (Average HFT) (3)			
$HFT_{i,t}$	Full	ATM	ITM	OTM	Full	ATM	ITM	OTM	Full	ATM	ITM	OTM
$QT_{i,t}$	2.61%	5.08%	4.77%	2.46%	2.89%	3.47%	5.16%	3.04%	3.85%	8.41%	7.57%	4.38%
$CT_{i,t}$	1.76%	5.59%	4.07%	3.13%	2.68%	2.71%	2.46%	1.60%	2.82%	9.43%	6.55%	5.99%
$OR_{i,t}$	8.10%	12.06%	11.32%	1.77%	3.31%	3.58%	7.35%	2.15%	10.54%	18.48%	19.20%	3.93%
$OV_{i,t}$	5.55%	7.86%	7.69%	-0.17%	2.97%	4.08%	9.33%	2.03%	7.66%	12.94%	13.48%	1.27%
$ITS_{i,t}$	3.32%	7.16%	6.17%	-2.61%	3.05%	3.05%	17.63%	2.22%	7.43%	14.33%	12.65%	-0.36%
Average	4.27%	7.55%	6.80%	0.91%	2.98%	3.38%	8.39%	2.21%	6.46%	12.72%	11.89%	3.04%

Table 7. The impact of HFT on options spread – NASDAQ sample

This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{SHFT}^D_{i,d} + \gamma_2 \widehat{SHFT}^S_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^D_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^S_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $SHFT^D_{i,d}$ and $SHFT^S_{i,d}$ are the measures of HFTs' liquidity-demanding and -supplying activities respectively, and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Columns 1 and 4, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 5, we use the 2SLS IV approach. $IV_{i,d}$ is a dummy variable that takes the value 1 during the flash-orders period (from June 5, 2009 to August 31, 2009) initiated by the NASDAQ. In Columns 3 and 6, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of two HFT proxies ($SHFT^D_{i,d}$ and $SHFT^S_{i,d}$) in all other stocks in the corresponding size quintile. The sample contains 103 stocks traded between January 1, 2009 and December 31, 2009 on the NASDAQ. We follow Bollen and Whaley (2004) and define OTM options as those with $|Odelta_{i,t}| \leq 0.375$, ATM options as those with $0.375 < |Odelta_{i,t}| \leq 0.625$, and ITM options as those with $|Odelta_{i,t}| > 0.625$. Standard errors are double clustered on stock and day, and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $\boldsymbol{ODspread_{i,t}}$ is the dependent variable.

	$SHFT^{All}_{i,d}$			$SHFT^D_{i,d}$			$SHFT^S_{i,d}$		
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.02***	0.02***	0.02***	0.03***	0.03***	0.03***	-0.01	-0.01	-0.01
	(3.93)	(3.85)	(3.41)	(3.11)	(3.27)	(3.57)	(-1.26)	(-1.33)	(-1.28)
ATM	0.02**	0.02**	0.02**	0.02**	0.03***	0.03***	-0.01	-0.01	-0.01
	(1.98)	(2.17)	(2.22)	(2.28)	(2.80)	(2.89)	(-1.05)	(-1.17)	(-1.06)
ITM	0.01*	0.01*	0.01*	0.02***	0.02***	0.02***	-0.01**	-0.01*	-0.01*
	(1.80)	(1.83)	(1.82)	(2.58)	(2.63)	(2.91)	(-2.38)	(-1.77)	(-1.85)
OTM	0.01	0.00	0.01	0.01	0.01	0.01	-0.01	-0.01	-0.01
	(0.33)	(0.04)	(0.16)	(0.94)	(0.55)	(1.12)	(-0.77)	(-0.90)	(-0.83)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

N	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600
Panel B: $OPspread_{i,t}$ is the dependent variable.									
		$SHFT_{i,d}^{All}$			$SHFT_{i,d}^{D}$			$SHFT_{i,d}^{S}$	
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.17***	0.16***	0.19***	0.32***	0.35***	0.33***	-0.12	-0.13	-0.10
	(2.88)	(3.47)	(3.95)	(3.18)	(3.49)	(3.32)	(-1.34)	(-1.41)	(-1.27)
ATM	0.14**	0.10**	0.11	0.32***	0.30***	0.34***	-0.19	-0.19	-0.22
	(2.03)	(1.99)	(2.26)	(2.72)	(2.93)	(2.73)	(-0.30)	(-0.44)	(-0.31)
ITM	-0.00	0.01	0.00	0.23*	0.22*	0.22*	-0.21*	-0.20*	-0.21*
	(-0.17)	(0.97)	(0.83)	(1.85)	(1.93)	(1.89)	(-1.91)	(-1.94)	(-1.90)
OTM	0.00	0.00	-0.00	0.17	0.18	0.19	-0.20	-0.21	-0.23
	(0.01)	(0.05)	(-0.27)	(1.02)	(1.23)	(1.04)	(-1.45)	(-1.57)	(-1.58)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
N	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600

Table 8. The impact of HFT on options spread without controlling for the stock spread – NASDAQ sample
This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{SHFT}^D_{i,d} + \gamma_2 \widehat{SHFT}^S_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^D_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^S_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $SHFT^D_{i,d}$ and $SHFT^S_{i,d}$ are the measures of HFTs' liquidity-demanding and -supplying activities respectively, and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the option and underlying markets. The option market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variable is $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Columns 1 and 4, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 5, we use the 2SLS IV approach. $IV_{i,d}$ is a dummy variable that takes the value 1 during the flash-orders period (from June 5, 2009 to August 31, 2009) initiated by the NASDAQ. In Columns 3 and 6, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of two HFT proxies ($SHFT^D_{i,d}$ and $SHFT^S_{i,d}$) in all other stocks in the corresponding size quintile. The sample contains 103 stocks traded between January 1, 2009 and December 31, 2009 on the NASDAQ. We follow Bollen and Whaley (2004) and define OTM options as those with $|Odelta_{i,t}| \leq 0.375$, ATM options as those with $0.375 < |Odelta_{i,t}| \leq 0.625$, and ITM options as those with $|Odelta_{i,t}| > 0.625$. Standard errors are double clustered on stock and day, and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $\boldsymbol{ODspread_{i,t}}$ is the dependent variable.									
	$SHFT^{All}_{i,d}$			$SHFT^D_{i,d}$			$SHFT^S_{i,d}$		
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.03**	0.02*	0.02**	0.04***	0.04***	0.04***	-0.02**	-0.02**	-0.02**
	(2.40)	(1.91)	(2.22)	(3.54)	(3.74)	(3.72)	(-2.26)	(-2.07)	(-2.16)
ATM	0.02*	0.02*	0.02*	0.03***	0.03***	0.03***	-0.02**	-0.02**	-0.02**
	(1.87)	(1.95)	(1.88)	(2.71)	(3.05)	(2.91)	(-2.29)	(-2.44)	(-2.21)
ITM	0.02	0.01	0.01*	0.03***	0.03***	0.03***	-0.02***	-0.02**	-0.02***
	(1.55)	(1.12)	(1.74)	(3.01)	(2.99)	(3.13)	(-2.93)	(-2.26)	(-2.99)
OTM	0.00	0.00	0.01	0.01*	0.01*	0.01*	-0.01*	-0.01*	-0.01*
	(1.09)	(0.77)	(1.10)	(1.88)	(1.89)	(1.90)	(-1.82)	(-1.76)	(-1.87)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
N	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600

Panel B: $OPspread_{i,t}$ is the dependent variable.

	$SHFT_{i,d}^{All}$			$SHFT_{i,d}^{D}$			$SHFT_{i,d}^{S}$		
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.21**	0.18**	0.23***	0.43***	0.40***	0.45***	-0.21**	-0.29**	-0.20**
	(2.42)	(2.05)	(2.89)	(3.63)	(4.10)	(3.68)	(-2.25)	(-2.32)	(-2.25)
ATM	0.19**	0.17**	0.12**	0.36***	0.37***	0.36***	-0.28*	-0.26*	-0.29*
	(2.16)	(1.99)	(2.07)	(3.18)	(3.86)	(3.19)	(-1.73)	(-1.90)	(-1.75)
ITM	0.11*	0.10*	0.14**	0.38**	0.36**	0.39**	-0.24**	-0.24**	-0.21**
	(1.82)	(1.72)	(2.35)	(2.02)	(2.19)	(2.06)	(-2.64)	(-2.21)	(-2.39)
OTM	-0.09	-0.11*	-0.06	0.11*	0.10*	0.13*	-0.29**	-0.24**	-0.27**
	(-1.47)	(-1.67)	(-0.93)	(1.70)	(1.88)	(1.78)	(-2.02)	(-2.08)	(-1.99)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
N	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600	24,600

Table 9. The impact of HFT on options spread during "no-news" days – NASDAQ sample
This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{SHFT}^D_{i,d} + \gamma_2 \widehat{SHFT}^S_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^D_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^S_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{6} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $SHFT^D_{i,d}$ and $SHFT^S_{i,d}$ are the measures of HFTs' liquidity-demanding and -supplying activities respectively and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. We follow Hirschey (2020) to identify days with firm news. Three specifications of the model are estimated. In Columns 1 and 4, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 5, we use the 2SLS IV approach. $IV_{i,d}$ is a dummy variable that takes the value 1 during the flash-orders period (from June 5, 2009 to August 31, 2009) initiated by the NASDAQ. In Columns 3 and 6, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of two HFT proxies ($SHFT^D_{i,d}$ and $SHFT^S_{i,d}$) in all other stocks in the corresponding size quintile. The sample contains 103 stocks traded between January 1, 2009 and December 31, 2009 on the NASDAQ. We follow Bollen and Whaley (2004) and define OTM options as those with $|Odelta_{i,t}| \leq 0.375$, ATM options as those with $0.375 < |Odelta_{i,t}| \leq 0.625$, and ITM options as those with $|Odelta_{i,t}| > 0.625$. Standard errors are double clustered on stock and day, and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $\boldsymbol{ODspread_{i,t}}$ is the dependent variable.

	$SHFT^{All}_{i,d}$			$SHFT^D_{i,d}$			$SHFT^S_{i,d}$		
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.02*	0.02**	0.02**	0.02**	0.02**	0.02**	-0.01	-0.01	-0.01
	(1.83)	(2.12)	(2.01)	(2.21)	(2.43)	(2.19)	(-0.89)	(-1.04)	(-0.81)
ATM	0.01*	0.01**	0.01*	0.02**	0.02**	0.02**	-0.01	-0.01	-0.01
	(1.79)	(2.09)	(1.87)	(1.99)	(2.50)	(1.98)	(-0.90)	(-0.50)	(-0.93)
ITM	0.00	0.00	0.00	0.01	0.01	0.01	-0.01*	-0.01*	-0.01*
	(0.15)	(0.13)	(0.06)	(1.49)	(1.37)	(1.55)	(-1.66)	(-1.73)	(-1.69)
OTM	0.00	0.00	0.00	0.01	0.01	0.01	-0.01	-0.01	-0.01
	(0.08)	(0.07)	(0.12)	(0.63)	(0.18)	(0.62)	(-0.86)	(-0.95)	(-0.92)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
N	18,210	18,210	18,210	18,210	18,210	18,210	18,210	18,210	18,210

Panel B: $OPspread_{i,t}$ is the dependent variable.

	$SHFT_{i,d}^{All}$			$SHFT_{i,d}^{D}$			$SHFT_{i,d}^{S}$		
	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)	OLS (1)	IV (Flash orders) (2)	IV (Average HFT) (3)
Full	0.16**	0.19**	0.13**	0.29**	0.28***	0.27**	-0.13	-0.11	-0.16
	(2.21)	(2.06)	(1.99)	(2.33)	(2.98)	(2.34)	(-0.17)	(-1.36)	(-0.25)
ATM	0.13**	0.20**	0.16**	0.26**	0.25***	0.18**	-0.12	-0.09	-0.13
	(2.12)	(2.04)	(2.33)	(2.20)	(2.78)	(2.43)	(-0.01)	(-0.06)	(-0.01)
ITM	-0.02	-0.11	-0.05	0.16	0.10	0.16	-0.21*	-0.26*	-0.25*
	(-0.29)	(-1.37)	(-0.36)	(1.51)	(1.52)	(1.51)	(-1.75)	(-1.88)	(-1.82)
OTM	0.05	-0.04	0.00	0.13	0.12	0.12	-0.20	-0.19	-0.21
	(0.33)	(-0.45)	(0.08)	(0.74)	(1.08)	(0.72)	(-1.08)	(-1.21)	(-1.07)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
N	18,210	18,210	18,210	18,210	18,210	18,210	18,210	18,210	18,210

Figure 1. The evolution of trading volume in US equity and options markets
This figure reports the evolution of trading volume in US equity and options markets. The grey (dark) bar corresponds to the number of shares (contracts) traded in US equity (options) markets. The sample contains all stocks traded between January 1, 1996 and December 31, 2020 on the US exchanges. The data is obtained from CRSP and OptionMetrics.

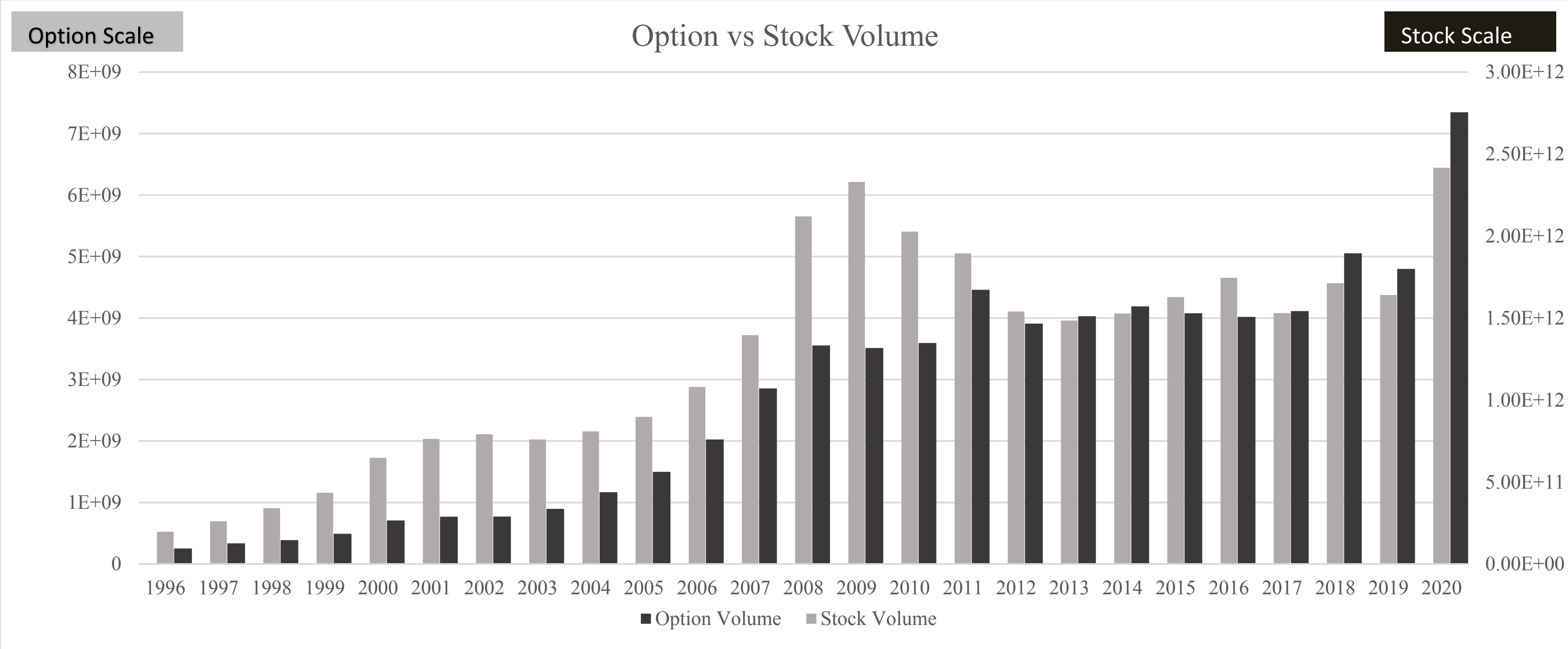

Figure 2. The impact of HFT on options spread – MIDAS sample: Year-by-year estimation
This figure reports the results from year-by-year OLS regressions of the association between HFT and the options spread. The grey (dark) bar corresponds to the impact of HFT on the options dollar (proportional) spread. $QT_{i,d}$ is used as an HFT measure (γ_1). The sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. All estimations include stock and time (day) fixed effects. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

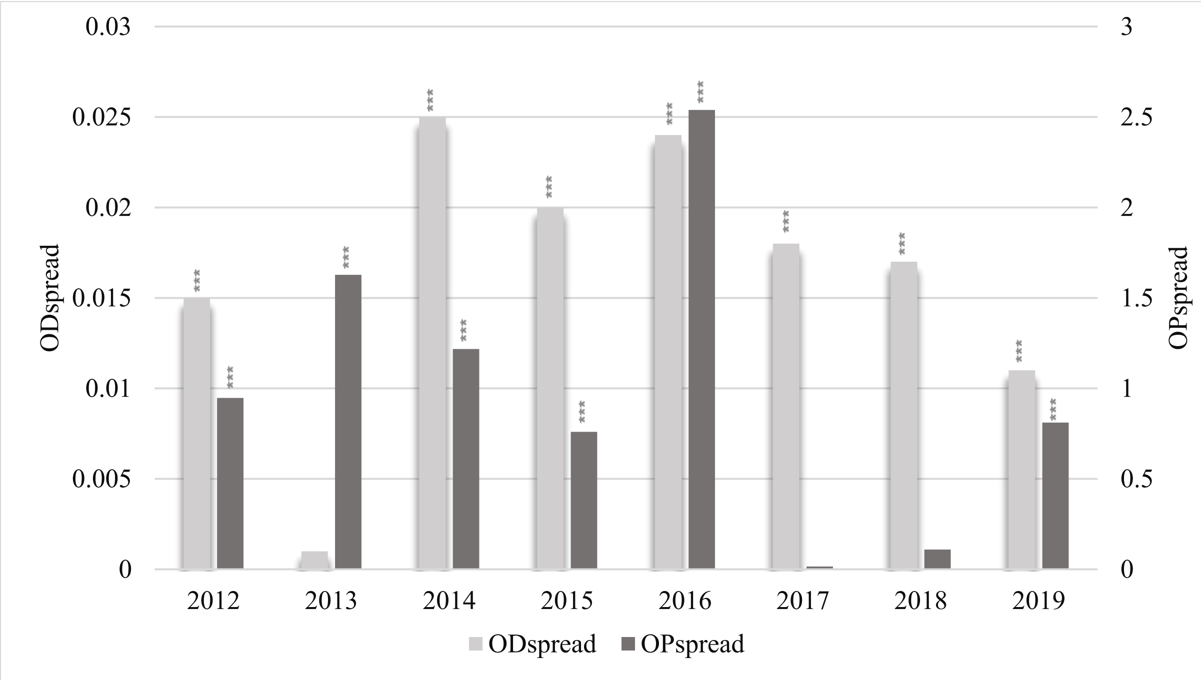

Table A.1. The impact of HFT on options spread controlling for the volatility of stock spread – MIDAS sample

This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT_{i,d}} + \sum_{k=1}^{8} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$
$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{8} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$ and the stock market variables are $SPspread_{i,d}$ and $SPspreadvol_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ and $SDspreadvol_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. $SPspreadvol_{i,d}$ and $SDspreadvol_{i,d}$ are the volatility of proportional and dollar spreads. For the definitions and computation methods of all the other variables, see Table 1. Three specifications of the model are estimated. In Column 1, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 3, we use the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks, and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $ODspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.02***	0.02***	0.03***
		(26.51)	(9.81)	(25.99)
	$CT_{i,d}$	0.01***	0.02***	0.01***
		(13.72)	(10.01)	(9.25)
	$OR_{i,d}$	0.07***	0.02***	0.11***
		(57.68)	(6.71)	(64.80)
	$OV_{i,d}$	0.05***	0.02***	0.09***
		(53.38)	(8.49)	(62.55)
	$ITS_{i,d}$	0.05***	0.02***	0.08***
		(41.03)	(6.22)	(44.36)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Panel B: $OPspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.89***	1.46***	1.31***
		(18.67)	(11.73)	(17.34)
	$CT_{i,d}$	0.60***	1.35***	0.96***
		(11.84)	(11.54)	(11.58)
	$OR_{i,d}$	2.76***	1.67***	3.59***
		(32.03)	(5.86)	(27.21)
	$OV_{i,d}$	1.89***	1.50***	2.61***
		(25.48)	(6.53)	(24.75)
	$ITS_{i,d}$	1.13***	1.54***	2.53***
		(12.79)	(5.06)	(19.68)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Table A.2. First stage instrumental variable (IV) regression results – MIDAS sample

This table presents the results for the estimation of the impact of the selected instruments on HFT measures:

$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

Where $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SDspread_{i,d}$ and $SVolatility_{i,d}$. Standard errors are double clustered on stock and day. For the definitions and computation methods of all the variables, see Table 1. In Columns 1, 3, 5 and 7, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Columns 2, 4, 6 and 8, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1, 3, 5 and 7, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. For Columns 2, 4, 6 and 8, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. We follow Bollen and Whaley (2004) and define OTM options as those with $|Odelta_{i,t}| \leq 0.375$, ATM options as those with $0.375 < |Odelta_{i,t}| \leq 0.625$, and ITM options as those with $|Odelta_{i,t}| > 0.625$. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

		Full sample		ATM		ITM		OTM	
	Dependent variable	IV (Tick size pilot) (1)	IV (Average HFT) (2)	IV (Tick size pilot) (3)	IV (Average HFT) (4)	IV (Tick size pilot) (5)	IV (Average HFT) (6)	IV (Tick size pilot) (7)	IV (Average HFT) (8)
$HFT_{i,t}$	$QT_{i,t}$	-0.30***	0.01***	-0.31***	0.01***	-0.31***	0.001*	-0.30***	0.05***
		(-98.58)	(2.89)	(-83.20)	(4.95)	(-74.90)	(1.83)	(-89.06)	(3.78)
	$CT_{i,t}$	-0.35***	0.04***	-0.35***	0.05***	-0.35***	0.04***	-0.34***	0.04***
		(-127.24)	(45.76)	(-107.74)	(45.34)	(-98.34)	(32.84)	(-113.69)	(43.11)
	$OR_{i,t}$	-0.09***	0.38***	-0.08***	0.36***	-0.08***	0.36***	-0.08***	0.36***
		(-66.38)	(307.8)	(-51.43)	(269.9)	(-43.73)	(250.9)	(-58.37)	(284.2)
	$OV_{i,t}$	-0.14***	0.32***	-0.14***	0.31***	-0.14***	0.29***	-0.14***	0.31***
		(-85.74)	(266.18)	(-68.49)	(233.25)	(-61.23)	(212.26)	(-76.18)	(246.03)
	$ITS_{i,t}$	-0.07***	0.21***	-0.07***	0.22***	-0.07***	0.21***	-0.07***	0.21***
		(-63.56)	(193.69)	(-49.34)	(181.89)	(-46.70)	(166.31)	(56.75)	(186.26)
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock and time FEs		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N		640,306	2,967,095	402,656	2,264,040	343,460	2,077,881	480,324	2,540,617

Table A.3. The impact of HFT on the realized volatility – MIDAS sample

This table presents the results for the estimation of the impact of HFT on the realized volatility:

$$SRVolatility_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT_{i,d}} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$
$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $SRVolatility_{i,d}$ is computed as the variance of second midpoint stock returns, $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), and $SVolatility_{i,d}$. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Column 1, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 3, we use the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks, and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

$SRVolatility_{i,d}$ is the dependent variable.				
	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	-0.001***	-0.002***	-0.002***
		(-45.22)	(-28.99)	(-44.14)
	$CT_{i,d}$	-0.002***	-0.002***	-0.003***
		(-56.77)	(-28.18)	(-39.99)
	$OR_{i,d}$	0.01***	-0.01***	0.02***
		(121.30)	(-21.88)	(173.34)
	$OV_{i,d}$	0.01***	-0.002***	0.01**
		(96.29)	(-20.82)	(140.18)
	$ITS_{i,d}$	0.01***	-0.002***	0.01***
		(82.27)	(-21.26)	(124.06)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Table A.4. The impact of HFT on options spread after controlling for the realized volatility – MIDAS sample

This table presents the results for the estimation of the impact of HFT on the options spread:

$$OSpread_{i,d} = \alpha_i + \beta_d + \gamma_1 \widehat{HFT_{i,d}} + \sum_{k=1}^{8} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$
$$HFT_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{8} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $OSpread_{i,d}$ corresponds to either the proportional spread ($OPspread_{i,d}$) or the dollar spread ($ODspread_{i,d}$), $HFT_{i,d}$ corresponds to one of the five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$), and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$, and the stock market variables are $SPspread_{i,d}$ (when we employ $OPspread_{i,d}$ as the dependent variable), $SDspread_{i,d}$ (when we use $ODspread_{i,d}$ as the dependent variable), $SVolatility_{i,d}$ and quote-based realized volatility ($SRVolatility_{i,d}$). $SRVolatility_{i,d}$ is computed as the variance of second midpoint stock returns. For the definitions and computation methods of all the variables, see Table 1. Three specifications of the model are estimated. In Column 1, we estimate the model by using OLS with stock and fixed effects. In Columns 2 and 3, we use the 2SLS IV approach. In Column 2, the level of HFT is instrumented with the changes in tick size; for this specification, $IV_{i,d}$ is a dummy variable that takes the value of one after the change (from October 3, 2016 to September 28, 2018), and zero before (from October 1, 2014 to October 2, 2016) for the treatment stocks, and $IV_{i,d}$ takes the value of zero in the entire period for the control stocks. In Column 3, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of five HFT proxies ($QT_{i,d}$, $CT_{i,d}$, $OR_{i,d}$, $OV_{i,d}$, $ITS_{i,d}$) in all other stocks in the corresponding size quintile. For Columns 1 and 3, the sample contains 2,746 stocks traded between January 1, 2012 and December 31, 2019 on the US exchanges. For Column 2, the sample contains 1,235 stocks (617 control stocks and 618 treated stocks) implemented in the SEC's Tick Size Pilot Program from October 1, 2014 to September 28, 2018. Standard errors are double clustered on stock and day and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

Panel A: $ODspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.02***	0.02***	0.03***
		(29.40)	(11.03)	(26.12)
	$CT_{i,d}$	0.01***	0.02***	0.01***
		(16.81)	(11.20)	(11.43)
	$OR_{i,d}$	0.06***	0.02***	0.10***
		(51.40)	(6.95)	(50.97)
	$OV_{i,d}$	0.05***	0.02***	0.08***
		(48.67)	(8.61)	(52.06)
	$ITS_{i,d}$	0.05***	0.02***	0.07***
		(35.65)	(6.40)	(35.39)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Panel B: $OPspread_{i,t}$ is the dependent variable.

	Variable	OLS (1)	IV (Tick size pilot) (2)	IV (Average HFT) (3)
$HFT_{i,d}$	$QT_{i,d}$	0.93***	1.48***	1.37***
		(20.24)	(11.34)	(18.66)
	$CT_{i,d}$	0.64***	1.37***	1.04***
		(12.87)	(15.30)	(12.51)
	$OR_{i,d}$	2.57***	1.72***	3.18***
		(29.67)	(5.96)	(23.40)
	$OV_{i,d}$	1.76***	1.54***	2.33***
		(25.29)	(7.31)	(21.74)
	$ITS_{i,d}$	0.93***	1.59***	2.18***
		(10.00)	(5.33)	(16.39)
	Controls	Yes	Yes	Yes
	Time and stock FEs	Yes	Yes	Yes
	N	2,969,829	640,306	2,967,095

Table A.5. Summary statistics for NASDAQ sample

This table reports the descriptive statistics for the variables used in our analysis. Panel A provides the descriptive statistics for all options-related variables separately for the full sample and three groups based on moneyness. Panel B shows the descriptive statistics for all variables from the underlying stock market. For the definitions and computation methods of the variables, see Table 1. We follow Bollen and Whaley (2004) and define OTM options as those with absolute option delta $|\Delta| \leq 0.375$, ATM options as those with $0.375 < |\Delta| \leq 0.625$, and ITM options as those with $|\Delta| > 0.625$. The sample contains 103 stocks traded between January 1, 2009 and December 31, 2009 on the NASDAQ.

Panel A. Equity market variables				
	Variable	Mean	Median	St dev
Full sample	$SHFT_{i,d}^{All}$	0.49	0.48	0.21
	$SHFT_{i,d}^{D}$	0.33	0.31	0.16
	$SHFT_{i,d}^{S}$	0.25	0.25	0.17
	$SPspread_{i,d}$	0.11	0.12	0.23
	$SDspread_{i,d}$ ($)	0.03	0.02	0.04
	$SVolatility_{i,d}$	0.41	0.35	6.65
	$SOIB_{i,d}$	0.09	0.07	0.09
Panel B. Option market variables				
Full sample	$ODspread_{i,t}$	0.136	0.100	0.137
	$OPspread_{i,t}$ (%)	5.61	4.52	5.25
	$Ovolume_{i,t}$	5.892	6.276	2.762
	$Oimplied_{i,t}$	0.407	0.359	0.245
	$\lvert Odelta_{i,t} \rvert$	0.539	0.530	0.247
	$Ogamma_{i,t}$	0.11	0.09	0.14
	$Ovega_{i,t}$	2.403	1.910	2.330
ATM	$ODspread_{i,t}$	0.144	0.117	0.178
	$OPspread_{i,t}$ (%)	0.056	0.046	0.089
	$Ovolume_{i,t}$	6.900	7.154	2.455
	$Oimplied_{i,t}$	0.398	0.344	0.244
	$\lvert Odelta_{i,t} \rvert$	0.476	0.469	0.159
	$Ogamma_{i,t}$	0.13	0.09	0.16
	$Ovega_{i,t}$	4.041	3.505	3.736
ITM	$ODspread_{i,t}$	0.207	0.163	0.216
	$OPspread_{i,t}$ (%)	0.037	0.031	0.073
	$Ovolume_{i,t}$	4.295	4.444	2.467
	$Oimplied_{i,t}$	0.502	0.468	0.286
	$\lvert Odelta_{i,t} \rvert$	0.732	0.703	0.182
	$Ogamma_{i,t}$	0.10	0.09	0.18
	$Ovega_{i,t}$	2.308	2.001	2.316
OTM	$ODspread_{i,t}$	0.103	0.098	0.134
	$OPspread_{i,t}$ (%)	0.075	0.069	0.091
	$Ovolume_{i,t}$	5.331	5.578	2.499
	$Oimplied_{i,t}$	0.206	0.202	0.202
	$\lvert Odelta_{i,t} \rvert$	0.110	0.107	0.088
	$Ogamma_{i,t}$	0.08	0.07	0.12
	$Ovega_{i,t}$	2.193	1.691	2.478

Table A.6. First stage instrumental variable (IV) regression results – NASDAQ sample

This table presents the results for the estimation of the impact of the selected instruments on HFT measures:

$$SHFT^{D}_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

$$SHFT^{S}_{i,d} = \alpha_i + \beta_d + \vartheta_1 IV_{i,d} + \sum_{k=1}^{7} \delta_k C_{k,i,d} + \varepsilon_{i,d}$$

where $SHFT^{D}_{i,d}$ and $SHFT^{S}_{i,d}$ are the measures of HFTs' liquidity-demanding and -supplying activities respectively, and α_i and β_d are stock and time (day) fixed effects. The $C_{k,i,d}$ is a set of k control variables, including variables from both the options and underlying markets. The options market variables are $Ovolume_{i,d}$, $Oimplied_{i,d}$, $|Odelta_{i,d}|$, $Ogamme_{i,d}$ and $Ovega_{i,d}$ and the stock market variables are $SDspread_{i,d}$ and $SVolatility_{i,d}$. Standard errors are double clustered on stock and day. For the definitions and computation methods of all the variables, see Table 1. Two specifications of the model are estimated. In Columns 1 and 3, $IV_{i,d}$ is a dummy variable that takes the value 1 during the flash-orders period (from June 5, 2009 to August 31, 2009) initiated by the NASDAQ. In Columns 2 and 4, the level of HFT is instrumented with the average level of HFT on that day in all other stocks in the corresponding size quintile; for this specification, $IV_{i,d}$ is the average level of two HFT proxies ($SHFT^{D}_{i,d}$ and $SHFT^{S}_{i,d}$) in all other stocks in the corresponding size quintile. The sample contains 103 stocks traded between January 1, 2009 and December 31, 2009 on the NASDAQ. Standard errors are double clustered on stock and day, and t-statistics are reported in parentheses. *, ** and *** denote significance at 10%, 5% and 1%.

	$SHFT^{All}_{i,d}$		$SHFT^{D}_{i,d}$		$SHFT^{D}_{i,d}$	
	IV (Flash orders) (1)	IV (Average HFT) (2)	IV (Flash orders) (3)	IV (Average HFT) (4)	IV (Flash orders) (5)	IV (Average HFT) (6)
$IV_{i,d}$	3.52*** (2.91)	0.27*** (9.89)	3.12*** (2.70)	0.20*** (8.37)	2.24** (2.52)	0.18*** (7.21)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes	No	Yes

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Economic and Social Research Council

Systemic Risk Centre

The London School of Economics
and Political Science
Houghton Street
London WC2A 2AE
United Kingdom

tel: +44 (0)20 7405 7686
systemicrisk.ac.uk
src@lse.ac.uk