Withdrawal of High-Frequency Traders and Intraday ETF Volatility during the Covid-19 Crisis *

Rajesh K. Aggarwal †

D'Amore-McKim School of Business – Northeastern University

Da Huang ‡

David Eccles School of Business – University of Utah

December 22, 2021

Abstract

Does high-frequency trade increase or decrease volatility in financial markets during crises? We introduce a novel intraday volatility measure for ETFs, and find that during the Covid-19 crisis period, the withdrawal of high-frequency trade from large stock ETFs increases intraday ETF volatility *net of the fundamental shock from Covid itself* by over 30%. The speed of arbitrage activities slows down during the Covid-19 period as high-frequency traders reduce the intensity of their trading. While high-frequency traders may serve as de facto market makers during normal times, they withdraw from the market during a crisis, precisely when they are needed most.

Keywords: Volatility, high-frequency trading, Covid-19.

*We thank Jonathan Brogaard, Mike Cooper, Felipe Cortes, Weiling Liu, Ali Sharifkhani, KC Zheng, and the seminar audience at Northeastern University for helpful comments. All errors are our own. Send correspondence to Rajesh Aggarwal.

†360 Huntington Avenue, 413 Hayden Hall, Boston, MA 02115. Email: r.aggarwal@northeastern.edu

‡1655 Campus Drive, Salt Lake City, UT 84112. Email: da.huang@eccles.utah.edu

1 Introduction

During the onset of the Covid-19 pandemic in the United States, volatility in U.S. equity markets increased dramatically. **Figure 1** plots the time series of the VIX volatility index from January 2019 to December 2020. The shaded region covers February 24, 2020, to April 30, 2020. During this period, there was a dramatic increase and then decrease in the VIX index. This time period corresponds to when the Covid-19 crisis was impounded into securities prices in the United States. During this time period, there was a large fundamental shock to both cash flows and discount rates, which could account for the large increase in volatility. **Figure 1** also shows that exchange traded funds (ETFs) exhibited a parallel volatility pattern to the VIX index. We argue that the withdrawal of high-frequency traders (HFT) from the ETF market increased volatility during the Covid-19 crisis. By focusing on ETFs, we are able to control for the general increased volatility in financial markets due to the fundamental shock, supporting a causal interpretation of our results.

The VIX began to increase on Friday, February 21, and then sharply increased starting on Monday, February 24, which Ramelli and Wagner (2020) attribute to the potential ramifications of Covid-19 beginning to be impounded into securities prices. They note that the first lockdowns in Italy in Lombardy occur on Sunday, February 23. Volatility is elevated and increasing through March 23, and then reverses and returns to its pre-Covid level by April 30. We call the period between February 24 and April 30 the Covid-19 period for U.S. equity markets, consistent with the timing in Pagano, Wagner, and Zechner (2020) in their study of cross-sectional equity returns. Because Covid-19 was expected to affect cash flows as well as access to financing and market risk premia, there were multiple sources of fundamental shocks, as discussed in Gormsen and Koijen (2020) and Landier and Thesmar (2020).

Prior literature has documented that high-frequency traders can provide liquidity to financial markets, and in so doing, stabilize securities prices (see Brogaard, et al. (2017)). By contrast, Anand and Venkataraman (2016) argue that in times of crises, high-frequency traders (or endogenous liquidity providers) may exacerbate volatility by trading in the general direction of other traders. In a similar vein, Korajczyk and Murphy (2019) argue that high-frequency trade reduces market quality. The Covid-19 crisis allows us to examine the impact of high-frequency trade during a period of market stress, but also introduces an additional complication. Because of the general increase in volatility due to the fundamental Covid shock, we must disentangle any changes in volatility due to the presence or withdrawal of high frequency traders from increases in volatility due to the fundamental shock itself.

We first show cross-sectionally that ETFs with less HFT intensity exhibit greater volatility during the Covid-19 period. This cross-sectional result implies that a reduction in high-frequency trade is associated with increased volatility, but need not be causal. For example, the fundamental shock may lead to both an increase in volatility and a reduction in high-frequency trade during the Covid-19 period.

To address the concern of endogeneity from the fundamental Covid shock, we synthetically construct the basket of securities held by each ETF into a fundamental ETF portfolio–that is, the underlying portfolio of stocks is the fundamental for the ETF. We calculate fundamental ETF portfolio prices by multiplying a vector of weights of ETF securities holdings by a vector of the underlying securities' prices. We then compute the percentage return on this fundamental portfolio held by the ETF. By a standard no-arbitrage argument, this fundamental portfolio return should equal the return based on the ETF's traded price.[1]

Empirically, the fundamental ETF portfolio return and the traded ETF return are highly correlated, but need not be exact due to trading effects. Specifically, the ETF and its fundamental portfolio are identical except that one is traded and the other is not. We next subtract the volatility of the fundamental ETF portfolio from the actual ETF volatility. This difference, or volatility gap, measures how much the actual ETF volatility differs from its underlying securities' portfolio volatility.

In theory, there should be no difference between the fundamental portfolio volatility and the actual ETF volatility–a volatility gap of zero. In practice, trading effects (e.g., from high-frequency traders) could cause the volatility gap to differ from zero, and variation in the volatility gap can only be attributed to trading.

Figure 2 plots the average volatility gap of the largest equity ETFs in the market from January 2019 to December 2020. We see a sharp increase during the Covid-19 crisis. Note that our focus on

[1] Note that we first calculate the fundamental portfolio price (similar to a NAV) and then calculate the fundamental portfolio return based on that price. Therefore, the return based on the fundamental portfolio price incorporates all return correlations between underlying securities in the portfolio. The return correlations between underlying securities would be ignored if the portfolio return were calculated as the weighted average of the return of the underlying securities.

ETFs specifically allows us to separate volatility due to fundamental shocks from volatility effects due to the presence or absence of high-frequency traders. The fundamental portfolio volatility captures all fundamental shocks, leaving the volatility gap to capture any volatility due to trading effects. We find that during the Covid period, lower high-frequency trade intensity is associated with a larger volatility gap. This result is consistent with high-frequency traders withdrawing from the market and thereby increasing volatility.

We next sort ETFs into those exposed to large amounts of high-frequency trading relative to those with small amounts of high-frequency trading prior to the Covid-19 crisis. The sorting is done based on the average HFT intensity as of the month of January 2018, more than two years prior to the onset of Covid-19 in the United States. Then, using a differences-in-differences specification, we show that ETFs exposed to high levels of ex ante HFT have larger increases in volatility gaps during the Covid period than those exposed to low levels of HFT.

Importantly, the relative ordering of HFT intensity across ETFs did not change from before to after the onset of the Covid-19 crisis, but the absolute level of HFT intensity in all securities decreases significantly, suggesting that high-frequency traders withdrawing from the market contributes to volatility. We extend this point by showing that only high ex ante HFT ETFs have a negative association between contemporaneous HFT and the volatility gap during the Covid-19 period. High-frequency traders withdraw the most from the high ex ante high-frequency trade ETFs, and these ETFs exhibit a 30% increase in volatility gaps. This is consistent with the reduction in high-frequency trade causing the increase in the volatility gap.

It is possible that high-frequency traders prefer to trade liquid or less volatile ETFs, and sorting ETFs into their ex ante high-frequency trade intensity is actually proxying for selection based on factors like volatility and illiquidity. We address this possibility by directly controlling for selection based on ex ante volatility gaps and ex ante illiquidity. Our results are unaffected. We conclude that selection does not explain our results. Instead, by withdrawing from the ETF market during the Covid-19 crisis, high-frequency traders contribute to volatility, as they abandon the role they may play as shadow or *de facto* market makers.

While econometrically the volatility gap is a novel way to address causal inference in the face of endogenous fundamental shocks, the volatility gap also has a natural economic interpretation of measuring the breakdown of arbitrage and the speed with which arbitrage and convergence occurs during a crisis. In non-crisis times, arbitrage between an ETF and its underlying fundamental portfolio happens very quickly. Using returns at the five second frequency to calculate volatility at the five minute frequency, we show that there is essentially no volatility gap at all during normal times, suggesting that arbitrage is happening more quickly than five seconds for returns. During the Covid period, the volatility gap expands out to the 20 minute frequency, representing arbitrage opportunities present for up to 20 seconds in the large, liquid ETFs in our sample.[2] These results are consistent with a subset of high-frequency traders being arbitrage capital that withdraws from the market during a shock.

We also find that high-frequency traders return to the market and close the volatility gap within five weeks after the Federal Reserve announces its liquidity support programs on March 23. It is worth emphasizing that the ETFs in our sample are the largest, most liquid stock ETFs in the United States market, thus depicting arbitrage breakdowns in normally well-functioning markets after a shock. Our analysis of volatility gap for ETFs is, to the best of our knowledge, a novel way to examine the impact of high-frequency trade, and more broadly trading in general, on the functioning of markets.

This paper is organized as follows. We discuss related literature in Section 2. We describe our data and variable construction in Section 3. Section 4 presents our econometric specifications and results for ETF volatility as well as the volatility gap. It also addresses our identification strategy relative to several endogeneity and selection concerns. Finally, Section 5 concludes.

2 Literature Review

We examine whether high-frequency trade increases or decreases volatility during periods of market crisis. Others have examined volatility and market fragility during periods of market stress (see, for example, Menkveld and Yueshen (2019), Brogaard, et al. (2017), and Anand and Venkataraman (2016)).

A number of papers have examined the effect of high-frequency trade on markets. Kirilenko and Lo (2013) provide a good overview of the history and issues associated with high-frequency and algorithmic trading. Hendershott, Jones, and Menkveld (2011) examine the effect of high-

[2]Because ETF sponsors publish net asset values (NAV) every fifteen seconds, beyond 15 seconds, high-frequency traders no longer have an advantage relative to any other traders in seeing and exploiting the arbitrage opportunity.

frequency and algorithmic trade on market liquidity. Brogaard, Hendershott, and Riordan (2014) examine the effect of high-frequency trade on price discovery, and Menkveld and van Kervel (2019) examine how high-frequency traders trade around large orders. Menkveld (2013) argues that high-frequency traders provide a substantial amount of liquidity to the market and can be viewed as effective market makers. Weller (2018) shows that high-frequency and algorithmic trading can reduce incentives to acquire information. Korajczyk and Murphy (2019) argue that high-frequency trade reduces market quality. In a similar vein, Bessembinder, Hao, and Zheng (2020) demonstrate that designated market makers serve an important role in improving market quality and question whether high-frequency traders do so as well. Hao and Zheng (2018) examine market concentration and liquidity using the failure of Knight Capital as a shock.

There is also a burgeoning literature examining the effect of Covid-19 on financial markets. Ramelli and Wagner (2020) provide a nice overview of events in equity markets during the period up to and including the Covid-19 period. Haddad, Moreira, and Muir (2020) study disruptions in the debt markets and show that Federal Reserve Bank actions on March 23, 2020, quickly began to reverse disruptions in targeted markets. Gormsen and Koijen (2020) extract growth expectations for equities over the Covid-19 period based on dividends, while Landier and Thesmar (2020) reach somewhat different conclusions based on earnings expectations. Pagano, Wagner, and Zechner (2020) examine cross-sectional heterogeneity in stock returns over the Covid-19 period based on firm's disaster resilience. This literature is rapidly growing.

We also contribute to the literature examining the effect of ETFs on securities markets. Ben-David, Franzoni, and Moussawi (2018) argue that ETFs increase volatility in their underlying securities. Our paper is related, but examines how high-frequency trade may induce a wedge between the volatility of the ETF and the underlying portfolio of securities. Our paper is also related to the literature on excess volatility (Shiller (1981), Marsh and Merton (1986)).

Overall, our contributions are fourfold. First, we are able to demonstrate that the withdrawal of high-frequency trade contributes to market volatility during a very recent period of market stress–the Covid-19 crisis. Second, we can separate the effect of high-frequency trade from the effect of fundamental shocks by examining both ETFs and underlying stocks (the fundamental ETF portfolio) held by the ETFs. Third, more generally, we show that ETF volatility is different from the volatility of the underlying portfolio of securities at higher frequencies, thus placing bounds on the speed of arbitrage. Fourth, in showing that high-frequency traders withdraw from markets during a period of crisis, our results further call into question whether high-frequency traders truly are effective or shadow market makers.

3 Data

The sample period is between January 1, 2019, and December 31, 2020. We designate February 24 through April 30, 2020 as the Covid period. The rest of the sample period (January 1, 2018, to February 21, 2020, and May 1, 2020, to December 31, 2020) is the control period. We use NYSE Trade and Quote (TAQ) data to calculate the odd-lot ratio (OLR) as our primary measure of high-frequency trade at an intraday frequency. We supplement this measure with a composite high-frequency trade measure at a daily frequency that incorporates popular high-frequency trade intensity measures in the literature (Weller (2018)). These data are from the Securities and Exchange Commission (SEC) Market Information Data Analytics System (MIDAS).[3]

We hand collect holdings of ETFs on a weekly basis throughout our sample period from Bloomberg.[4] We also use price and asset under management data from the Center for Research in Securities Prices (CRSP). To construct our primary volatility variables, we first compute the ETF returns every five seconds. We then compute the standard deviation of returns for each ETF using five minute intervals. Each standard deviation (volatility) observation is therefore computed from 60 returns calculated every five seconds over the five minute interval. During a 6.5 hour trading day, there are 78 five-minute intervals. Following the standard practice of the literature, we discard the first and last three five-minute intervals during the trading day (the first and last fifteen minutes of trade) so that our results do not depend on and are not contaminated by beginning-of-trade and close-auction effects. During the Covid period, there were four suspensions of trade due to the market being limit down (Market-Wide Circuit Breakers, or MWCBs). In those cases, we also discard the first fifteen minutes of trade after the market reopens.[5]

We emphasize that there is a significant difference between the intraday volatility we construct and conventional volatility measures using daily or even lower frequency returns. Our intraday

[3]The SEC MIDAS data are available at: https://www.sec.gov/marketstructure/downloads.html

[4]ETF holdings are highly correlated ($\rho = 0.99$) week-to-week in our sample.

[5]The results are qualitatively the same whether we include the first and last fifteen minutes of the trading session and/or the first fifteen minutes after the market reopens following MWCBs.

volatility only captures return variability during a trading session, whereas conventional volatility captures the total variability in close-to-close returns. As a result, conventional volatility will capture volatility that occurs outside of trading hours, and will be substantially larger than our intraday volatility measure.

Table 1 presents summary statistics for volatility and other variables in the sample. **Figure 1** plots average volatility by day for the ETFs in the sample, alongside the VIX index. The Covid period is shaded. We average across all ETFs at each five-minute interval during the day, and then across all five-minute intervals during the day:

$$\sigma_d = \frac{1}{72}\frac{1}{N}\sum_{t=4}^{t=75}\sum_{i=1}^{N}\sigma_{i,d,t}. \tag{1}$$

Here, $\sigma_{i,d,t}$ is the volatility based on the five-second returns of ETF i on day t during five-minute interval t. As previously noted, **Figure 1** shows there is an increase in ETF volatility starting around February 24, that then becomes more substantial over the month of March before declining in April, similar to the increase and then decrease in the VIX. Our definition of the sample period from February 24 to April 30 as the Covid-19 crisis period is consistent with Ramelli and Wagner (2020) and Pagano, Wagner, and Zechner (2020), who argue that the news about Covid-19 started to be incorporated into securities' prices on Monday, February 24. They end their analyses on Friday, March 20, as the Federal Reserve announced its liquidity measures for the bond markets–the Primary Market Corporate Credit Facility (PMCCF) and the Secondary Market Corporate Credit Facility (SMCCF)–on Monday, March 23. We extend the definition of Covid crisis period to April 30 as volatility remains elevated.

We focus on U.S. equity ETFs that are not levered or inverse so that we can later synthetically construct a fundamental portfolio of the ETF using its holdings data, which plays a crucial part of the identification strategy. We use the universe of 64 U.S equity ETFs with assets under management greater than $5 billion at the end of 2019 that meet the above restrictions. We exclude five ETFs due to missing data, leaving a final sample of 59 ETFs.[6]

Table 1 shows that there is a notable increase in ETF volatility during the Covid-19 period

[6] The list of ETFs used and the detailed selection procedure are in **Appendix Table A1**.

(mean of 4.59 basis points per five minute interval) relative to the control period (mean of 1.28 basis points per five minute interval). To separate the effects of general market volatility from intraday ETF volatility induced by trading effects, we introduce the concept of the ETF volatility gap. The ETF volatility gap is the difference between the actual ETF volatility and the volatility of the fundamental portfolio that consists of the actual holdings of the ETF. In order to construct the fundamental portfolio of stocks held by each ETF, we collect the weekly holdings data (from the prior Friday) from Bloomberg. We then compute the price every five seconds for the fundamental portfolio using the weekly weightings $\boldsymbol{w}$ of securities in the ETF's basket of holdings. To do so, we take the associated vector of stock prices $\boldsymbol{P}$ at five second trading intervals during the subsequent week for all securities held in an ETF's basket.[7] The inner product of the weights and stock prices yields a sequence of prices for fundamental portfolio i: $\hat{p}_{i,d,t} = \boldsymbol{w} \cdot \boldsymbol{P}$. This is equivalent to the net asset value (NAV) of the ETF (up to linear scaling) computed every five seconds during the trading day.[8]

As an illustration, **Figure 3** plots the price trend of the SPDR S&P 500 (SPY) ETF and its fundamental portfolio on one day–March 5, 2020. That the ETF market price and the portfolio fundamental price can deviate from each other induces the ETF volatility gap. We calculate a series of returns ($\hat{r}_{i,d,t}$) for the fundamental portfolio every five seconds within the five-minute interval t, and compute volatility ($\hat{\sigma}_{i,d,t}$) for the fundamental portfolio every five minutes using 60 five-second return observations. The difference between the volatility based on the ETF market returns and the volatility based on fundamental portfolio returns is the volatility gap: $\Delta\sigma_{i,d,t} = \sigma_{i,d,t} - \hat{\sigma}_{i,d,t}$. Similarly, the difference between the ETF market return and the fundamental portfolio return is the return gap: $\Delta r_{i,d,t} = r_{i,d,t} - \hat{r}_{i,d,t}$.[9]

The volatility gap measures how much the actual ETF volatility differs from its underlying securities' portfolio volatility. The underlying portfolio is the fundamental for the ETF. In theory, there should be no difference between the fundamental portfolio volatility and the actual ETF volatility–a volatility gap of zero–because there is no uncertainty about fundamentals for the ETF

[7]One may be concerned by the difference in the frequencies of the holdings data and price data. In reality, the ETF holdings are strikingly persistent as the week-to-week holdings weight correlation is on average above 0.99 in the sample.

[8]We do not use the ETF's index or benchmark, as each ETF chooses a different subset of stocks in its index in order to mimic or minimize tracking error relative to its benchmark's performance. Our fundamental portfolios are more reflective of actual ETF holdings than is the benchmark or index.

[9]Note that, importantly, volatility gap, $\Delta\sigma_{i,d,t}$, is *not* calculated as Standard Deviation($\Delta r_{i,d,t}$).

relative to its underlying portfolio of securities. In practice, trading effects, and in particular, the effects of high-frequency traders could cause the volatility gap to differ from zero.

Both ETFs and each of their component stocks have two sources of shocks: shocks due to fundamentals and shocks due to trading. Because the fundamental for the ETF is the portfolio of stocks it holds, when we compute the volatility of the portfolio of stocks held by the ETF, we have directly incorporated both shocks to the fundamentals of the stocks and shocks due to trading for the stocks. Each stock price itself already incorporates trading and fundamental shocks to the stock. The fundamental portfolio price of the ETF (NAV) is just the weighted average of the stock prices in the portfolio. Any deviation of the actual ETF price from the fundamental portfolio price of the ETF is a reflection of shocks due to trading of the ETF itself, or a failure of arbitrage, as parity should always be maintained between the ETF market trading price and the fundamental portfolio price. Therefore, in focusing on the ETF volatility gap, we have eliminated fundamental shocks to the ETF (and, by extension, fundamental and trading shocks to the stocks comprising the ETF) and left only shocks due to trading of the ETF.

Table 1 shows that the mean volatility gap over the entire sample period is 0.44 basis points per five minute interval. However, the difference between the control period and Covid-19 period is striking. During the control period, the mean volatility gap is 0.30 basis points (median of 0.14 basis points, essentially zero). During the Covid period, the volatility gap increases to 1.85 basis points, so the increase in the volatility gap is 1.55 basis points (1.85-0.30 basis points) from the control period to the Covid period.

Total intraday volatility for the ETFs in the sample increases by 3.31 basis points from 1.28 basis points during the control period to 4.59 basis points during the Covid period. The 1.55 basis points increase in the volatility gap accounts for nearly half of the increase in total intraday volatility for the ETF sample from the control period to the Covid period. Others have documented divergences between bond ETF prices and NAVs due to lack of liquidity in those markets (see Haddad, Moreira, and Muir (2020)) during the Covid-19 period. To our knowledge, we are the first to study the volatility gap in stock ETFs during a crisis. We do this intraday for the largest, most liquid ETFs in the US market.

To further illustrate the extraordinary nature of the increase in volatility gap during the Covid period, **Figure 2** plots the average volatility gap over the time period January 2019 to December

2020 for our ETFs. The sharp increase in the volatility gap during the Covid period is evident.

Not surprisingly, the mean return gap is zero for both the control period and the Covid period. This is because of two effects. First, ETF sponsors force end-of-day convergence in ETF prices and the value of the underlying portfolio of stocks held by the ETF. Second, sponsors publish the ETF NAV intraday every 15 seconds.[10] As a practical matter, any trader can see divergences between an ETFs traded price and its NAV at the 15 second frequency. Thus, positive and negative return gaps will mechanically cancel out not only at the end of the day but also throughout the day. Instead of presenting summary statistics for the return gap, we present statistics for the absolute value of the return gap. The absolute return gap exhibits a similar pattern to the volatility gap statistics. In our subsequent tests, we focus on the cross-sectional heterogeneity in the volatility gap.

We hypothesize that the increase in the volatility gap is related to the amount of high-frequency trade in the ETFs. To define the amount of high-frequency trade, we use two measures from the literature. Weller (2018) uses a combination of four proxies: odd-lot ratio, trade-to-order ratio, cancel-to-trade ratio, and average trade size. As noted by Weller (2018), odd-lot ratio and cancel-to-trade ratio are associated with higher HFT, while trade-to-order ratio and average trade size are associated with lower HFT. We construct a composite HFT intensity measure by summing up the z-scores of odd-lot ratio and cancel-to-trade ratio and the negative z-scores of trade-to-order ratio and average trade size. We compute the level of HFT intensity ($\text{HFT}_{i,d}$) for the ETFs on a daily basis using the MIDAS data.

A disadvantage of using this measure is that it can only be computed on a daily basis, yet we measure volatility gaps at the five-minute frequency. We are more interested in changes in high-frequency trading intraday, rather than a composite daily average of high-frequency trading. Therefore, we focus specifically on one component of high-frequency trade intensity–the odd-lot ratio–as it is the cleanest of the measures of HFT and can be calculated on an intraday frequency using readily available TAQ data. High-frequency traders frequently use odd-lot orders to execute their trades, as shown by O'Hara, Yao, and Ye (2014). We initially present results for both measures, but then focus on the odd-lot ratio to identify high-frequency trade intraday.

The composite high-frequency trade intensity is high for ETFs during the control period (mean of 0.65), and declines during the Covid period (mean of 0.47). The difference is even more substan-

[10]For more details see the investor bulletin from the SEC: https://www.sec.gov/investor/alerts/etfs.pdf

tial at the median. We see similar effects for the odd-lot ratio ($\text{OLR}_{i,d,t}$). Odd-lot trades decline by 33 percent during the Covid period from 0.18 to 0.12 of all trades for our sample ETFs. This is our first indication that high-frequency traders withdraw from the market during the Covid-19 shock period. To see this graphically, **Figure 4** plots HFT intensity over the entire sample period. There is a sharp reduction in high-frequency trade in the March 6 to March 23 period, followed by a substantial rebound. In what follows, we examine whether high-frequency trade in ETFs is positively or negatively correlated with volatility and volatility gaps during the Covid-19 crisis period relative to the control period.

We focus on returns measured every five seconds and volatility measured every five minutes to approximate the frequency with which high-frequency traders trade (see Menkveld (2018)). Consistent with this, the volatility gap is more pronounced at the five second/five minute level than they are at longer frequencies. To see this, **Figure 5** graphs the distribution of volatility gap calculated from $\sigma_{i,d,t}$ and $\hat{\sigma}_{i,d,t}$ measured at five different intervals: five second returns with five minute volatility, 10 second returns with 10 minute volatility, 15 second returns with 15 minute volatility, 20 second returns with 20 minute volatility, and 30 second returns with 30 minute volatility. In all cases, we use 60 return observations to calculate volatility.

Panel A plots the volatility gap density during the control period, while Panel B plots the volatility gap density during the Covid period. In Panel A (control period), all of the peaks are quite sharp and centered around a zero volatility gap. The 20 minute and 30 minute volatility gaps are more sharply peaked around zero than is the 15 minute volatility gap than is the 10 minute volatility gap than is the five minute volatility gap. In other words, as the volatility gap is measured at longer frequencies, volatility of the ETF tends to converge to the volatility of its underlying fundamental portfolio of securities.

Panel B (Covid period) shows a somewhat different pattern–the densities are much more spread out. Further, the volatility gap exhibits a right skew, and the peak of the density for the five minute volatility gap is further from zero than the peak of the density is for the 10 minute and 15 minute ones, which in turn are further than the 20 minute and 30 minute ones.[11] Because the mean volatility gap is positive during the Covid period, on average the volatility of the ETFs is greater than the volatility of the underlying portfolio. These patterns suggest that, during the Covid period,

[11] We find a similar pattern (unreported) when we examine the absolute return gaps for the sample ETFs.

arbitrage works in the ETF market at the level of 30 seconds/30 minutes rather than at the level of five seconds/five minutes as during our control period. This is indicative that high-frequency traders may not close arbitrage gaps quickly during periods of heightened market volatility. This observation is one motivation for our study and we will examine this issue systematically.

In our subsequent empirical work, one of our primary variables for addressing selection concerns is the level of high-frequency trade in each ETF averaged over the month of January 2018, two years prior to the onset of the Covid-19 crisis in the United States. High-frequency trade in January 2018 is not likely to be affected by volatility in securities in February-April 2020. We sort 59 ETFs in the sample into three terciles based on their HFT intensity for the ETFs in January 2018. Note that we sort ETFs based on the level of high-frequency trade in the ETF itself rather than on the level of high-frequency trade in the underlying securities held by each ETF.

A key assumption of the analysis is that the assignment of ETFs to high-frequency trade categories is stable over time. To see whether this is true, we compute daily average high-frequency trade intensity for each tercile from January 2019 through December 2020. **Figure 6** graphs each terciles' average high-frequency trade intensity over time for both the composite HFT measure and the odd-lot ratio. It is important to emphasize that the assignment to terciles occurs in January 2018 and is then held fixed–we do not re-sort every month. While one can see changes in HFT intensity over time, the ordering of the terciles does not change–the high ex ante HFT intensity tercile persistently has higher average HFT intensity, and HFT intensity moves in parallel over time across the three terciles. Also notably, there is a marked drop in HFT intensity starting in February 2020 for all three terciles, but more so for the high ex ante HFT intensity tercile.

4 Results

4.1 Cross-Sectional ETF Volatility

As previously noted, we are interested in whether volatility and the volatility gap are increasing in high-frequency trade during the Covid-19 crisis period relative to the control period. We use several econometric specifications. The first examines the cross-sectional correlation between volatility (and volatility gap) and the contemporaneous amount of high-frequency trade in the ETF:

$$\text{Volatility}_{i,d,t} = \beta \times \text{HFT} + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t}, \tag{2}$$

where κ_i is an ETF fixed effect, η_d is a day fixed effect, τ_t is a time of day fixed effect, and $\mathbf{X}$ is a vector of covariates. The dependent variable $\text{Volatility}_{i,d,t}$ is the volatility of ETF i on day d at five-minute interval t. High-frequency trade intensity, HFT, is measured in two ways–the composite HFT measure at the ETF-day level and the odd-lot ratio measured at the ETF-day-five-minute level. The odd-lot ratio is likely to be the better measure for this study as it matches the frequency with which we measure volatility and the volatility gap. Throughout, we cluster standard errors by ETF, date, and time of the day.

Many studies of the cross-sectional determinants of idiosyncratic volatility (Irvine and Pontiff (2009), Bekaert, Hodrick, and Zhang (2012), Dennis and Strickland (2002)) include variables such as firm size, leverage, return on assets, market concentration, book-to-market, and momentum as security characteristics that are associated with idiosyncratic volatility. Many of these variables are measured at the monthly or quarterly frequency, and many are hard to define at the ETF level. Because volatility is measured in five-minute increments, the traditional controls above will largely be subsumed in ETF fixed effects, time of day fixed effects, and day fixed effects, and so we do not include such variables in our specifications. In addition, liquidity, as measured by the effective spread is also often included as a control. Our specifications do control for liquidity (volume-weighted percentage effective spread) at the daily level for each ETF. In addition, we control for ETF price on a daily basis, and ETF size (assets under management) on a monthly basis.

We first examine the contemporaneous correlation between ETF volatility and high-frequency trade intensity during the control period. **Table 2**, Panel A, shows that the association is significantly negative–less high-frequency trade is associated with greater volatility over this time period. Because we include time fixed effects, we can interpret our results as the relation between high-frequency trade and idiosyncratic volatility (relative to the market). Columns 1 and 3 use the composite measure of HFT, while Columns 2 and 4 use the odd-lot ratio. Columns 1 and 2 include the entire day of trade, while Columns 3 and 4 exclude the first and last 15 minutes of trade. Columns 3 and 4 show that the magnitude of the coefficients is not affected if we remove the first and last 15 minutes of trade during the trading day.[12] While the coefficients are statistically significant, they are small in magnitude. For example, using the coefficient in Column 3 for high-frequency trade, a one standard deviation increase in high-frequency trade decreases volatility during the control period by 0.007 standard deviations, or 0.007·1.63 bps = 0.01 basis points per five minute interval.

Panel B examines the contemporaneous correlation between volatility and high-frequency trade during the Covid period. Again we find that the association is negative–less high-frequency trade is associated with greater volatility. However, relative to the control period, the association between high-frequency trade and ETF volatility is much stronger. The coefficients are much larger in magnitude, and the R^2 are dramatically greater. Using the coefficient from Column 3, a one standard deviation increase in high-frequency trade decreases volatility during the Covid period by 0.18 standard deviations, or 0.18*4.09 basis points = 0.74 basis points per five minute interval. The effect of high-frequency trade on volatility is statistically and economically much larger during the Covid period than the control period–as much as 74 times as large as shown in Column 3.

4.2 Volatility Gaps

While we are generally interested in how high-frequency trade affects contemporaneous volatility, we know that contemporaneous high-frequency trade and volatility are likely to suffer from substantial endogeneity and reverse causality concerns. For example, high market-wide volatility (e.g., the Covid shock) could result in traders withdrawing while increasing ETF volatility.

Second, a mechanical correlation concern may apply to our ETFs. For the high ex ante HFT ETFs, high-frequency traders may be trading in the ETFs to hedge their holdings in the underlying securities. If the ETF holdings are high volatility securities, then the effect of high-frequency trade on volatility in ETFs may be mechanically driven by the volatility of the underlying securities. In other words, there may be no unique information in the ETFs relative to the underlying portfolio of securities.

To address both market-wide volatility and mechanical correlation concerns, we compute ETF volatility gaps, and use volatility gaps as our dependent variable. The volatility gap is defined

[12]The results are qualitatively the same for all the analyses whether or not we include the first and last 15 minutes of trade. We only report results without the first and last 15 minutes of trade in subsequent analyses.

as $VolGap_{i,d,t} = \sigma_{i,d,t} - \hat{\sigma}_{i,d,t}$ where $\sigma_{i,d,t}$ is the ETF volatility and $\hat{\sigma}_{i,d,t}$ is the volatility of its fundamental portfolio. The logic here is that the ETF volatility gap purges any market-wide or fundamental volatility in the portfolio of stocks held by the ETF. In subtracting the volatility of the fundamental portfolio from the volatility of the ETF, we separate out the fundamental shock from the effect of volatility induced by trade in the ETF. By construction, we also eliminate any mechanical correlation between the ETF and its underlying holdings.

What remains is volatility associated with trade of the ETF itself, independent of any shocks or volatility in the underlying stocks. Put differently, due to the otherwise identical nature of the ETF and its fundamental portfolio, trading is the only difference between the two, and therefore must drive variation in the volatility gap.

In **Table 3**, we estimate **Equation 2** using the volatility gap as the dependent variable. Columns 1 and 2 report results for the control period, and Columns 3 and 4 report results for the Covid period. Columns 1 and 3 use the composite high-frequency trade measure, and Columns 2 and 4 use the odd-lot ratio. In the control period, the effects of the composite high-frequency trade measure on the volatility gap are relatively weak, and are insignificant for the odd-lot ratio. This result is consistent with the idea that during normal times, there is little divergence of ETF prices, returns, and volatilities from those of the underlying portfolios, and thus little scope for impact from high-frequency trade.

By contrast, during the Covid period, there is a strong negative association between high-frequency trade and the volatility gap. Our primary specification is in Column 4 where we use the five minute frequency odd-lot ratio as our measure of high-frequency trade. Less high-frequency trade in the form of fewer odd-lot trades is associated with larger volatility gaps. We emphasize that in this specification, the dependent variable is the volatility gap measured at the five-minute frequency and our independent variable is the lagged odd-lot ratio also measured at the five-minute frequency. We saturate this specification with high dimensional ETF, day, and time-of-day fixed effects. Each of these features of the specification make it difficult to envision sources of endogeneity and reverse causality that affect our results. Because our specification is occurring at the five-minute frequency, fundamental information arrival seems unlikely to be driving the result. The dependent variable–volatility gap–by construction will eliminate endogenous sources of variation other than trading effects. The fixed effects further control for many other sources of variation, including any

ETF time-invariant characteristics.

As our previous figures showed, high-frequency trade withdraws from the market during the onset of the Covid period. Our results in **Table 3** suggest that the withdrawal of high-frequency trade during the Covid period increased ETF volatility gaps.

4.3 Heterogeneity across ETFs

Different ETFs are likely to have varying degrees of high-frequency trade. We next examine heterogeneity in high-frequency trade. We sort our 59 ETFs into three terciles based on their level of high-frequency trade in January, 2018. We do this for both the composite high-frequency trade measure and the OLR intensity for our ETFs. We define high, middle, and low HFT terciles of ETFs. Our logic here is that we can identify ETFs with high levels of high-frequency trade long before there are any changes in volatility for these ETFs. We calculate the average level of high-frequency trade in each ETF averaged over the month of January 2018, two years prior to the onset of the Covid-19 crisis in the United States. High-frequency trade in January 2018, is presumably not affected by volatility in securities in February-April 2020. As **Figure 6** shows, the ordering of the terciles does not change–the ETFs in the high tercile persistently have higher average high-frequency trade than do those in the middle tercile, and than do those in the low tercile. The fact that the assignment to terciles is stable over time demonstrates that relative interest in an ETF by high-frequency traders is a characteristic of the ETF. We next examine how much volatility gaps increase for ETFs with ex ante (January 2018) high levels of high-frequency trade relative to other ETFs during the Covid-19 period.

We estimate the following equation:

$$VolGap_{i,d,t} = \boldsymbol{\beta} \times (HighOLR_i + LowOLR_i) \times Covid_d + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t}, \tag{3}$$

where $VolGap_{i,d,t}$ is the difference between volatility of ETF i and the synthetically constructed portfolio using its holdings on day d at the five-minute interval t. The main independent variables of interest are $HighOLR_i$ and $LowOLR_i$. The odd-lot ratio ($OLR_{i,d,t-1}$) is our intraday high-frequency trade intensity measured at the five minute frequency. We do an ex ante OLR sort, and $HighOLR_i$ and $LowOLR_i$ are indicator variables that equal 1 if ETF i is in the high (low) OLR tercile as of January 2018 and 0 otherwise. The middle tercile is omitted to avoid collinearity. $Covid_d$ is an indicator variable that equals 1 if day d is between February 24, 2020 and April 30, 2020, and 0 otherwise. ETF fixed effects and day fixed effects absorb the standalone tercile and Covid indicators.

Table 4, Column 1 presents the difference-in-difference coefficients for the high OLR tercile and the low OLR tercile relative to the middle tercile from the Covid period to the control period. The high OLR tercile shows a positive, significant, and large in magnitude increase in volatility during the Covid-19 period relative to the middle tercile.

To assess magnitudes, consider the coefficient of 0.53 on $HighOLR_i \times Covid_d$ in Column 1. The standard deviation for the volatility gap during the Covid period in our sample is 3.41 basis points. Therefore, high ex ante OLR ETFs have volatility gaps that are $3.41 \cdot 0.53 = 1.81$ basis points higher than middle tercile OLR ETFs, with both relative to the control period. Recall from **Table 1** that the mean ETF during the control period has a volatility gap of 0.30 basis points, which then increases to 1.85 basis points during the Covid period. Almost all of the volatility gap increase during the Covid period comes from the high ex ante OLR ETFs. This result suggests that it is those ETFs with the most high-frequency trade ex ante that show large increases in volatility gaps. We emphasize that volatility gaps are, by construction, unrelated to volatility attributable to fundamental shocks.

We note that the Covid-19 crisis affects all securities and ETFs. Thus, the result in **Table 4**, Column 1, is not a classical natural experiment. Our hypothesis is that the Covid-19 crisis differentially affects high ex ante high-frequency trade ETFs versus low ex ante high-frequency trade ETFs. In order for this to be a valid experiment, the Covid-19 crisis should not change the categorization of the ETFs into terciles, which we demonstrate in **Figure 6**.

Given that ETFs with high levels of ex ante high-frequency trade are the ones that show the greatest increases in volatility gaps during the Covid period, these ETFs should also show a larger reduction in high-frequency trade during the Covid period. We augment **Equation 3** by fully interacting the ex ante OLR sort and the $Covid_d$ indicator with the contemporaneous odd-lot ratio ($OLR_{i,d,t-1}$) measured the same day as the volatility gap but five minutes earlier. Our specification is:

$$
\begin{aligned}
VolGap_{i,d,t} = \boldsymbol{\beta} \times \Big[& (HighOLR_i + LowOLR_i) \times Covid_d \\
& + (HighOLR_i + LowOLR_i) \times Covid_d \times OLR_{i,d,t-1} \\
& + (HighOLR_i + LowOLR_i) \times OLR_{i,d,t-1} + Covid_d \times OLR_{i,d,t-1} \\
& + OLR_{i,d,t-1} \Big] + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t}.
\end{aligned} \tag{4}
$$

The ETF fixed effects absorb the standalone tercile variables, and the time fixed effects absorb the standalone $Covid_d$ variable. The coefficient on $HighOLR_i \times Covid_d \times OLR_{i,d,t-1}$, the main variable of interest, explains the mechanics of how high-frequency trade behaves for high ex ante OLR ETFs during the Covid period. The coefficient on $HighOLR_i \times \text{Covid}_d$ should continue to be positive and significant as in Column 1.

Table 4, Column 2, presents the results. We continue to find that the high ex ante OLR tercile shows a large increase in the volatility gap during the Covid-19 period (the first coefficient), consistent with the results in Column 1. However, the key result is in the third row. The interaction of high OLR tercile with the Covid period indicator and the contemporaneous level of high-frequency trade (OLR) is negative and significant. This demonstrates that less high-frequency trade within high ex ante OLR ETFs is associated with a significant increase in the volatility gap. This result is crucial, as the effect of the withdrawal of high-frequency trade on the volatility gap cannot be attributed to increased volatility due to fundamental shocks. The coefficient of -0.23 says that for ex ante high odd-lot ratio (i.e., high ex ante HFT) ETFs during the Covid period, a one standard deviation decrease in the odd-lot ratio (high-frequency trade) is associated with an increase of 0.43 basis point (-0.23 times a standard deviation of 1.85 basis point) in volatility gap per five minute interval. This is a sizable effect–the mean volatility gap for the whole sample period is 0.44 basis point, i.e., a one standard deviation decrease in high-frequency trade doubles the volatility gap. As shown in **Table 1**, the odd-lot ratio declines by one third of a standard deviation from control to Covid period, which translates into a 30% increase in volatility gap solely attributable to the withdrawal of high-frequency traders. This is one of the key findings of our study.

We conclude that, while high-frequency trade could provide market stability, its absence increases the volatility gap. In crisis times, volatility increases when high-frequency traders withdraw from the ETFs in which they usually trade the most. ETFs that get the most benefit from shadow market making from high-frequency traders during normal times are also harmed the most when high-frequency traders withdraw during crises, while those ETFs who are helped less by high-frequency trading during normal times are not harmed as much.

A potential concern with these results is that they could be generated primarily on days when the market as a whole was either up or down. High-frequency traders could withdraw from the market on down days, exacerbating drawdowns and volatility, and then return on up days. Another possibility is that on down days, the volatility gap decreases or even becomes negative as volatility in the underlying fundamental portfolio increases more than volatility of the ETF itself.

To address this, Columns 3 and 4 separate our analysis into up day and down day subsamples. We find that our results are not generated specifically on either up or down days–both show a negative coefficient on the triple interaction of the high ex ante OLR tercile, the Covid period, and the continuous odd-lot ratio (continuous high-frequency trade). Less high-frequency trade during the Covid period for ETFs that were previously heavily traded by high-frequency traders increases the volatility gap (volatility net of fundamentals) on both up and down days. This result also further confirms the point that trade in the ETFs themselves is the source of the volatility, not their underlying holdings.

To further establish the robustness of this result, we conduct a placebo test by examining the volatility gap over the months February to April, 2019–one year prior to our Covid-19 period. We choose February 24, 2019, as the onset of the placebo treatment to mirror February 24, 2020, as the onset of the actual Covid-19 period. We replicate our results for Column 2 for this placebo period in Column 5. None of the coefficients are significant and all are close to zero in magnitude. There is essentially no volatility gap between the traded ETFs and their fundamental portfolios, and high-frequency trade therefore has no effect on the volatility gap. This result shows two things. First, in non-crisis times such as our placebo period, high-frequency trade (as a subset of arbitrage capital) either does not encounter volatility gaps or has completely arbitraged away volatility gaps by the time we measure them at the five second/five minute frequency. Second, the sudden increase in volatility during the Covid-19 period is strongly correlated with the substantial reduction in high-frequency trading in precisely those ETFs that were, prior to Covid-19, heavily traded by high-frequency traders.

4.4 Robustness

A key assumption in our previous analysis is that volatility gaps exist due to trading or microstructure effects. We next relax this assumption. Suppose for some reason independent of trading effects that high volatility ETFs are also high volatility gap ETFs. For example, suppose that volatility gaps simply scale with volatility–greater volatility mechanically increases the volatility gap. Suppose further that high-frequency traders do not cause volatility gaps by withdrawing their trade, but do dislike volatility. If high-frequency traders dislike volatility, then empirically, it will appear that they avoid ETFs with high volatility gaps. This will be pronounced in the Covid period, when volatility and volatility gaps increase. In this possibility, there is no causal relation between reduced HFT in an ETF and a larger volatility gap.[13] Along these same lines, high-frequency traders may dislike illiquidity, and if illiquidity is correlated with volatility gaps, then it may appear that high-frequency traders avoid large volatility gaps. Again, there would be no causal relation between reduced HFT in an ETF and a larger volatility gap.

If these alternatives are true, then when a large shock to fundamentals such as Covid-19 occurs, it should lead to decreased high-frequency trade in ETFs with ex ante greater volatility gaps or ex ante less liquidity, and result in greater ex post volatility gaps. In other words, if these alternatives are true, then selection based on ex ante volatility gaps or ex ante illiquidity drives the association between high-frequency trade and the volatility gap during the Covid-19 period. To address these concerns, we augment our specification in **Equation 4** by including additional sorts into terciles based on the ex ante volatility gap and illiquidity as of January 2018.

For ease of comparison across specifications, **Table 5**, Column 1 replicates our results for **Table 4**, Column 2. We only report the triple interaction coefficients. **Table 5**, Column 2 includes the additional tercile sort based on the ex ante volatility gap and Column 3 includes the additional tercile sort based on ex ante illiquidity (effective spread). The specification parallels our treatment of ex ante HFT intensity terciles. These additional categorical variables are interacted with the Covid_d indicator variable as well as the continuous odd-lot ratio measure to control for selection based on ex ante volatility gaps or ex ante illiquidity. Column 4 includes all three sorts–ex ante high-

[13] To be clear, empirically, high volatility ETFs are high volatility gap ETFs, and do have less high-frequency trade. The alternative explanation described here could explain this result without being causally due to the withdrawal of HFT in the Covid period.

frequency trade intensity, ex ante volatility gaps, and ex ante illiquidity. In all four specifications, only the ex ante high-frequency trade intensity triple interaction is significant. The coefficient again implies that after controlling for the possibility of selection based on ETF characteristics, the effect of high-frequency trade on volatility gaps is negative and significant during the Covid period for those ETFs that high-frequency traders usually trade. This coefficient is quite stable across specifications, and suggests that withdrawal of high-frequency trade hurts those ETFs that usually receive the most high-frequency trade.

Our results together suggest that high-frequency traders differentially withdraw more from their ex ante high trade ETFs while still maintaining the relative ordering of how much they trade in different ETFs. In other words, high-frequency traders may withdraw from trading both high and low high-frequency trade ETFs, but withdraw relatively more from high HFT ETFs. This withdrawal increases volatility in those ETFs that ex ante have high HFT. As shown, all of these results hold even after controlling for the possibility that high-frequency traders trade less in ex ante high volatility or low liquidity securities.

4.5 Intraday, How Quickly do Volatility Gaps Close?

Our previous results show that volatility had dissipated in the aggregate by the end of the Covid period, April 30, 2020.[14] It is also useful to see how quickly within the day the volatility gap dissipates. This analysis is interesting for two reasons. First, we know that there are no volatility gaps on average during our control period–high-frequency traders eliminate any gaps that might exist in less than five seconds/five minutes for returns/volatilities. During the Covid period, volatility gaps do emerge at the five second/five minute frequency. For how much longer beyond five seconds/five minutes do these gaps persist? Second, we demonstrate that high-frequency traders withdraw from the market during the Covid period, which we argue increases the volatility gap. Analyzing the speed with which volatility gaps close intraday during the Covid period could be informative as to when or even whether high-frequency traders return to the market during the day.

In order to address this, we examine different frequencies of volatility gaps: 1) returns every five

[14]The period from the peak of volatility gap to volatility gap dissipating roughly corresponds to March 18 to April 30. On March 23, the Federal Reserve announced its liquidity support to markets. Even so, it took another five weeks after Fed action for volatility gap to return to pre-Covid levels. While the Fed may have precipitated the return to the market of high-frequency traders, we still find that the speed of closing of arbitrage opportunities (volatility gap) was not immediate.

seconds and volatilities every five minutes, as in our standard specification, 2) returns measured every 10 seconds, with 60 return observations used to calculate volatility every 10 minutes, 3) returns measured every 15 seconds, with 60 return observations used to calculate volatility every 15 minutes, 4) returns measured every 20 seconds, with 60 return observations used to calculate volatility every 20 minutes, and 5) returns measured every 30 seconds, with 60 return observations used to calculate volatility every 30 minutes. For each of the five different frequencies, we then calculate the ETF volatility gap. We repeat our specification from **Equation 4**, using each of the five different frequencies. Our goal here is to see at what frequency high-frequency trade intensity ceases to influence the volatility gap.

The results are in **Table 6**. For the interaction of Covid with the high ex ante OLR tercile and the contemporaneous HFT measure (odd-lot ratio), the coefficients are negative (the first row of coefficients). This is consistent with the withdrawal of HFT increasing volatility gaps in the highest ex ante HFT ETF group during the Covid-19 period. This is true at all frequencies. However, the magnitude of these coefficients shrinks as the time interval increases. Beyond 20 minutes for volatility gaps (which is calculated based on 20-second returns), the coefficient on the triple interaction has lost statistical significance. This implies that at the 15 second/15 minute to 20 second/20 minute frequency, volatility gaps that could be arbitraged have been closed.

Intuitively, this makes sense. Because ETF sponsors publish NAVs for their ETFs every 15 seconds, high-frequency traders' advantage relative to other traders should disappear after 15 seconds. Put differently, HFTs should close volatility gaps in less than 15 seconds. After 15 seconds, we expect other traders to arrive and close these gaps, and this appears to be the case during the Covid period. **Figure 4** shows that HFTs withdraw from the market during the Covid period. As a result, they do not close volatility gaps within 15 seconds, as demonstrated in **Table 6**. Because volatility gaps were essentially zero prior to the Covid-19 shock at the five second/five minute interval, our results suggest that within the day, arbitrage slows from faster than the five second/five minute to as much as the 20 second/20 minute frequency during the Covid-19 period.

5 Conclusion

We find that the withdrawal of high-frequency trade increases volatility during the Covid-19 crisis. This increase in volatility is over and above the volatility induced by the fundamental shock. Indeed, we find that 30% of the increase in volatility for ETFs net of the fundamental Covid-19 shock can be attributed to the withdrawal of high-frequency trade. This increase in volatility exhibits substantial heterogeneity, and is concentrated in those ETFs that prior to the Covid shock had substantial high-frequency trade. In effect, those ETFs that benefited the most from high-frequency trade prior to Covid had the greatest reduction in high-frequency trade during Covid, and as a result had larger increases in volatility.

We also demonstrate that the speed with which arbitrage works in the ETF market slows down during the Covid-19 crisis. While in normal non-crisis times, arbitrage happens at return frequencies less than five seconds and volatility frequencies measured at less than five minutes, during the Covid shock, arbitrage takes up to 20 seconds/20 minutes for returns/volatilities. The ETFs we examine are the largest, most liquid stock ETFs in the United States with essentially no limits or restrictions on trade in either the ETFs or their underlying securities.

Our results suggest that high-frequency trade does not stabilize markets during crisis periods. To the extent that high-frequency traders serve as effective or shadow market makers, our analysis is consistent with the view that shadow market makers are not a good substitute for designated market makers during periods of extreme market stress. Shadow market makers are more likely to withdraw from markets than to stabilize them precisely when they are needed most.

References

Anand, Amber and Kumar Venkataraman. 2016. Market conditions, fragility, and the economics of market making. *Journal of Financial Economics* 121: 327-349.

Bekaert, Geert, Robert Hodrick, and Xiaoyan Zhang. 2012. Aggregate Idiosyncratic Volatility. *Journal of Financial and Quantitative Analysis* 47:1155-1185.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2018. Do ETFs Increase Volatility? *Journal of Finance* 73: 2471-2535.

Bessembinder, Hendrik, Jia Hao, and Kuncheng Zheng. 2017. Liquidity Provision Contracts and Market Quality: Evidence from the New York Stock Exchange. *Review of Financial Studies* 33:44-74.

Brogaard, Jonathan, Terrence Hendershott and Ryan Riordan. 2014. High frequency trading and price discovery. *Review of Financial Studies* 27:2267-2306.

Brogaard, Jonathan, Ryan Riordan, Andriy Shkilko, Konstantin Sokolov, Allen Carrion, and Thibaut Moyaert. 2017. High frequency trading and extreme price movements. *Journal of Financial Economics* 128:253-265.

Dennis, Patrick and Deon Strickland. 2002. Who Blinks in Volatile Markets, Individuals or Institutions? *Journal of Finance* 57:1923-1949.

Gormsen, Niels Joachim, and Ralph S.J. Koijen. 2020. Coronavirus: Impact on stock prices and growth expectations. *The Review of Asset Pricing Studies* 10.4 (2020): 574-597.

Hao, Jia and Kuncheng Zheng. 2018. Market Concentration in the Liquidity Provision Industry and Its Effect on the Cost of Liquidity. D'Amore-McKim School of Business Working Paper.

Haddad, Valentin, Alan Moreira, and Tyler Muir. 2020. When Selling Becomes Viral: Disruptions in Debt Markets in the COVID-19 Crisis and the Fed's Response. NBER Working Paper 27168.

Hendershott, Terrence, Charles Jones, and Albert Menkveld. 2011. Does algorithmic trading improve liquidity? *Journal of Finance* 66:1-33

Irvine, Paul and Jeffrey Pontiff. 2009. Idiosyncratic Return Volatility, Cash Flows, and Product Market Competition. *Review of Financial Studies* 22:1149-1177.

Kirilenko, Andrei and Andrew Lo. 2013. Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents. *Journal of Economic Perspectives* 27:51-72.

Korajczyk, Robert and Dermot Murphy. 2019. High Frequency Market Making to Large Institutional Trades. *Review of Financial Studies* 32:1034-1067.

Landier, Augustin and David Thesmar. 2020. Earnings Expectations in the Covid Crisis. *The Review of Asset Pricing Studies* 10.4 (2020): 598-617.

Marsh, Terry and Robert Merton. 1986. Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices. *American Economic Review* 76:483-98.

Menkveld, Albert. 2013. High Frequency Trading and the New-Market Makers. *Journal of Financial Markets*. 16:712-740.

Menkveld, Albert. 2018. High-Frequency Trading as Viewed Through an Electron Microscope. *Financial Analysts Journal* 74: 24-31.

Menkveld, Albert and Vincent van Kervel. 2019. High-Frequency Trading Around Large Institutional Orders. *Journal of Finance* 74:1091-1137.

Menkveld, Albert and Bart Zhou Yueshen. 2019. The Flash Crash: A Cautionary Tale about Highly Fragmented Markets. *Management Science* 65: 4470-4488.

O'Hara, Maureen, Chen Yao, and Mao Ye. 2014. What's not there: Odd lots and market data. *Journal of Finance* 69.5 2014: 2199-2236.

Pagano, Marco, Christian Wagner, and Josef Zechner. 2020. Disaster Resilience and Asset Prices. CEPR Discussion Paper No. DP14773.

Ramelli, Stefano and Alexander F. Wagner. 2020. Feverish stock price reactions to Covid-19. Swiss Finance Institute Research Paper (20-12).

Shiller, Robert. 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71: 421-436.

Weller, Brian M. 2018. Does Algorithmic Trading Reduce Information Acquisition? *Review of Financial Studies* 31:2184-2226.

Table 1: **Summary Statistics**

	Mean	StDev	P10	Median	P90	#Obs.
Volatility (bps)	1.59	2.21	0.48	1.04	2.85	2,243,056
Control period	1.28	1.63	0.46	0.96	2.19	2,035,439
Covid period	4.59	4.09	1.52	3.35	8.86	207,617
Volatility Gap (bps)	0.44	1.85	-0.43	0.17	1.19	2,243,056
Control period	0.30	1.54	-0.45	0.14	0.94	2,035,439
Covid period	1.85	3.41	-0.18	0.88	4.67	207,617
|Return Gap| (bps)	0.97	0.63	0.40	0.79	1.78	2,243,056
Control period	0.67	0.31	0.36	0.60	1.08	2,035,439
Covid period	1.37	0.73	0.64	1.19	2.37	207,617
Odd-Lot Ratio	0.18	0.19	0.03	0.12	0.38	2,243,056
Control period	0.18	0.19	0.03	0.13	0.39	2,035,439
Covid period	0.12	0.12	0.02	0.09	0.25	207,617
Composite HFT Intensity	0.63	0.93	-0.51	0.68	1.73	2,243,056
Control period	0.65	0.94	-0.51	0.70	1.75	2,035,439
Covid period	0.47	0.78	-0.48	0.49	1.43	207,617
Price ($)	130.91	72.64	52.84	117.86	228.60	2,243,056
Effectlive Spread (bps)	4.51	10.81	1.18	2.57	7.88	2,243,056
Asset under Management ($Mil)	29,448	47,275	5,466	13,120	52,039	2,243,056

The table presents the summary statistics of the variables used in the paper. Volatility is measured at the five-minute frequency using five-second returns. Volatility gap is the difference between volatility of an ETF's returns of its market price and the volatility of the returns of a synthetically constructed fundamental portfolio using its holdings. Return gap is the difference between the return of an ETF's market price and the return of the price of its fundamental portfolio. High-frequency trade intensity (HFT) is the sum of the z-scores of odd-lot ratio and cancel-to-trade ratio and the negative z-scores of trade-to-order ratio and average trade size. Effective spread is volume-weighted and measured as a percentage of ETF's price (reported in basis point). Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETFs that are not levered or inverse and have at least $5 billion assets under management as of the end of 2019 (59 unique ETFs in total), from Jan 2019 to Dec 2020.

Table 2: **ETF volatility and high-frequency trade intensity**

Panel A: Control Period

	(1) $\text{Volatility}_{i,d,t}$	(2) $\text{Volatility}_{i,d,t}$	(3) $\text{Volatility}_{i,d,t}$	(4) $\text{Volatility}_{i,d,t}$
HFT	-0.008**	-0.005**	-0.007*	-0.003***
	(0.004)	(0.002)	(0.004)	(0.001)
$\text{Price}_{i,d-1}$	0.04*	0.04*	0.04*	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)
Effective $\text{Spread}_{i,d-1}$	0.02***	0.02***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Asset under $\text{Management}_{i,m-1}$	-0.09***	-0.09***	-0.09***	-0.09***
	(0.03)	(0.03)	(0.02)	(0.03)
HFT Measure	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$
Sample Period	Non-Covid	Non-Covid	Non-Covid	Non-Covid
First/Last 15 Min	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
Time of Day FE	Yes	Yes	Yes	Yes
Observations	2,035,439	2,035,439	1,904,505	1,904,505
Adj. R-squared	0.222	0.222	0.208	0.208

Panel B: Covid Period

	(1) $\text{Volatility}_{i,d,t}$	(2) $\text{Volatility}_{i,d,t}$	(3) $\text{Volatility}_{i,d,t}$	(4) $\text{Volatility}_{i,d,t}$
HFT	-0.17**	-0.08***	-0.18**	-0.07***
	(0.07)	(0.03)	(0.07)	(0.02)
$\text{Price}_{i,d-1}$	0.44	0.47	0.47	0.50
	(0.36)	(0.37)	(0.37)	(0.38)
Effective $\text{Spread}_{i,d-1}$	-0.00	-0.00	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Asset under $\text{Management}_{i,m-1}$	-0.27	-0.26	-0.29	-0.28
	(0.25)	(0.26)	(0.26)	(0.27)
HFT Measure	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$
Sample Period	Covid	Covid	Covid	Covid
First/Last 15 Min	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
Time of Day FE	Yes	Yes	Yes	Yes
Observations	207,617	207,617	194,308	194,308
Adj. R-squared	0.580	0.579	0.591	0.590

This table presents regressions of ETF volatility and high-frequency trade intensity:

$$\text{Volatility}_{i,d,t} = \beta \times \text{HFT} + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t},$$

where $\text{Volatility}_{i,d,t}$ is the standard deviation of ETF i's five-second returns during the five-minute interval d on day t. The main independent variable of interest HFT is the high-frequency trading intensity, measured as a composite of four popular HFT measure at ETF-day level and odd-lot-ratio measured at ETF-day-five-minute level. $\mathbf{X}$ is a vector of covariates that includes price of ETF i measured as of last trading day, volume-weighted percentage effective spread of ETF i measured as of last trading day, and asset under management of ETF i measured as of the end of last month. All continuous variables are standardized to have a standard deviation of 1. Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETF that is not levered or inverse and has at least \$5 billion asset under management as of the end of last month, from Jan 2019 to Dec 2020. Standard errors in parentheses are robust and clustered by ETF, day, and time of the day. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3: **ETF volatility gap and high-frequency trade intensity**

	(1)	(2)	(3)	(4)
	Control Period		Covid Period	
	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$
HFT	-0.01*	0.00	-0.23***	-0.07**
	(0.01)	(0.00)	(0.08)	(0.03)
$\text{Price}_{i,d-1}$	-0.04	-0.05	1.46**	1.51**
	(0.07)	(0.07)	(0.55)	(0.56)
$\text{Effective Spread}_{i,d-1}$	0.01*	0.01*	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
$\text{Asset under Management}_{i,m-1}$	-0.18***	-0.18***	-0.85**	-0.84**
	(0.05)	(0.05)	(0.36)	(0.38)
HFT Measure	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$	$\text{HFT}_{i,d}$	$\text{OLR}_{i,d,t-1}$
Day FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
Time of Day FE	Yes	Yes	Yes	Yes
Observations	1,904,505	1,904,505	194,308	194,308
Adj. R-squared	0.149	0.149	0.455	0.453

This table presents regressions of ETF volatility gap and high-frequency trading intensity:

$$\text{Volatility Gap}_{i,d,t} = \beta \times \text{HFT} + \boldsymbol{\chi}\boldsymbol{X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t},$$

where Volatility $\text{Gap}_{i,d,t}$ the difference between volatility of ETF i and a synthetically constructed portfolio using its holding, on day d at the five-minute interval t. The main independent variable of interest HFT is the high-frequency trading intensity, measured as a composite of four popular HFT measure at ETF-day level and odd-lot-ratio measured at ETF-day-five-minute level. $\mathbf{X}$ is a vector of covariates that includes price of ETF i measured as of last trading day, volume-weighted percentage effective spread of ETF i measured as of last trading day, and asset under management of ETF i measured as of the end of last month. All continuous variables are standardized to have a standard deviation of 1. Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETF that is not levered or inverse and has at least $5 billion asset under management as of the end of last month, from Jan 2019 to Dec 2020. Standard errors in parentheses are robust and clustered by ETF, day, and time of the day. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 4: **ETF volatility gap and high-frequency trade intensity: withdrawal effect**

	(1) $\text{VolGap}_{i,d,t}$	(2) $\text{VolGap}_{i,d,t}$	(3) $\text{VolGap}_{i,d,t}$	(4) $\text{VolGap}_{i,d,t}$	(5) $\text{VolGap}_{i,d,t}$
High $\text{OLR}_i \times \text{Covid}_d$	0.53**	0.48**	0.43**	0.53**	-0.07
	(0.20)	(0.20)	(0.21)	(0.21)	(0.05)
Low $\text{OLR}_i \times \text{Covid}_d$	-0.15	-0.14	-0.17	-0.11	0.03
	(0.12)	(0.13)	(0.13)	(0.12)	(0.03)
High $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$		-0.23**	-0.20*	-0.26**	0.02
		(0.10)	(0.11)	(0.10)	(0.02)
Low $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$		-0.05	-0.07	-0.03	-0.01
		(0.08)	(0.08)	(0.07)	(0.02)
Sample	All	All	Up Days	Down Days	All
Treatment	True	True	True	True	Placebo
Controls	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes	Yes
Time of Day FE	Yes	Yes	Yes	Yes	Yes
Observations	2,098,813	2,098,813	1,229,365	869,448	2,098,813
Adj. R-squared	0.263	0.264	0.237	0.297	0.257

Column 1 of this table presents the regression of ETF volatility gap and the interaction of the sorting based on ex ante HFT intensity and the Covid-19 period indicator:

$$VolGap_{i,d,t} = \boldsymbol{\beta} \times (\text{High OLR}_i + \text{Low OLR}_i) \times \text{Covid}_d + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t},$$

Columns 2 through 4 of this table present regressions of ETF volatility gap and the factorial interaction of the sorting based on ex ante HFT intensity, the Covid-19 period indicator, and the contemporaneous HFT intensity:

$$\begin{aligned} VolGap_{i,d,t} = \boldsymbol{\beta} \times \Big[& (\text{High OLR}_i + \text{Low OLR}_i) \times \text{Covid}_d \\ & + (\text{High OLR}_i + \text{Low OLR}_i) \times \text{Covid}_d \times \text{OLR}_{i,d,t-1} \\ & + (\text{High OLR}_i + \text{Low OLR}_i) \times \text{OLR}_{i,d,t-1} + \text{Covid}_d \times \text{OLR}_{i,d,t-1} + \text{OLR}_{i,d,t-1} \Big] \\ & + \boldsymbol{\chi X} + \kappa_i + \eta_d + \tau_t + \zeta_{i,d,t}, \end{aligned}$$

Only the estimates of interactions of interest is reported. Volatility $\text{Gap}_{i,d,t}$ the difference between volatility of ETF i and a synthetically constructed portfolio using its holding, on day d at the five-minute interval t. The main independent variables of interest, High OLR_i and Low OLR_i, are indicator variables that equal 1 if ETF i is in the high (low) odd-lot ratio tercile as of January 2018 and 0 otherwise, and the continuous variable $\text{OLR}_{i,d,t-1}$ that represents the odd-lot ratio of ETF i on day d during the five-minute interval $t-1$. We drop the mid tercile to avoid collinearity. Covid_d is an indicator variable that equals 1 if day d is between February 24, 2020 and April 30, 2020, and 0 otherwise. $\mathbf{X}$ is a vector of covariates that includes price of ETF i measured as of last trading day, volume-weighted percentage effective spread of ETF i measured as of last trading day, and asset under management of ETF i measured as of the end of last month. Columns (3) and (4) repeat analysis in column (2) with subsamples of days in which market return is positive or negative. Under column (5), Covid_d is re-defined as an indicator variable that equals 1 if day d is between February 24, 2019 and April 30, 2019, and 0 otherwise. All continuous variables are standardized to have a standard deviation of 1. Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETF that are not levered or inverse and have at least $5 billion asset under management as of the end of last month, from Jan 2019 to Dec 2020. Standard errors in parentheses are robust and clustered by ETF, day, and time of the day. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 5: **ETF volatility gap and high-frequency trade intensity: selection**

	(1) $\text{VolGap}_{i,d,t}$	(2) $\text{VolGap}_{i,d,t}$	(3) $\text{VolGap}_{i,d,t}$	(4) $\text{VolGap}_{i,d,t}$
High $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$	-0.23**	-0.25**	-0.29**	-0.22**
	(0.10)	(0.10)	(0.13)	(0.11)
Low $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$	-0.05	-0.00	0.00	0.06
	(0.08)	(0.12)	(0.11)	(0.11)
High $\text{VolGap}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$		0.11		0.07
		(0.09)		(0.09)
Low $\text{VolGap}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$		0.03		-0.01
		(0.04)		(0.07)
High $\text{Espread}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$			0.01	-0.01
			(0.07)	(0.11)
Low $\text{Espread}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$			-0.11	0.00
			(0.11)	(0.10)
Controls	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
Time of day FE	Yes	Yes	Yes	Yes
Observations	2,098,813	2,098,813	2,098,813	2,098,813
Adj. R-squared	0.264	0.271	0.265	0.272

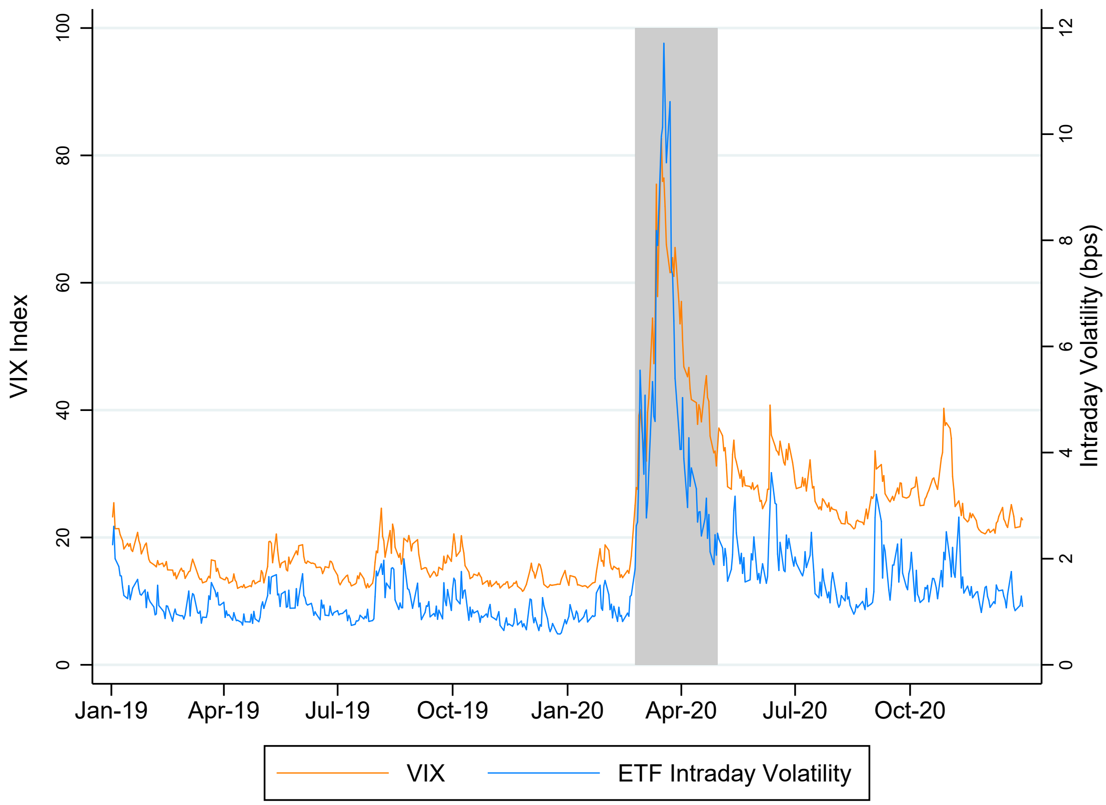
This table repeats the analysis in **Table 4** column 2. Column 1 is identical to **Table 4** column 2 to establish a baseline. Columns 2 and 3 add to the factorial interaction the tercile sorting based on ex ante volatility gap and the tercile sorting based on ex ante illiquidity, separately. Column 4 adds both. High VolGap_i and Low VolGap_i are indicator variables that equal 1 if ETF i if ETF i is in the high (low) volatility gap tercile as of January 2018 and 0 otherwise. High Espread_i and Low Espread_i are indicator variables that equal 1 if ETF i if ETF i is in the high (low) effective spread tercile as of January 2018 and 0 otherwise. $\text{OLR}_{i,d,t-1}$ is the odd-lot ratio of ETF i on day d during the five-minute interval $t-1$. Covid_d is an indicator variable that equals 1 if day d is between February 24, 2020 and April 30, 2020, and 0 otherwise. All continuous variables are standardized to have a standard deviation of 1. Only the estimates of the triple interactions of tercile sorts, the Covid-19 period indicator, and the contemporaneous odd-lot ratio are reported. Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETF that is not levered or inverse and has at least $5 billion asset under management as of the end of last month, from Jan 2019 to Dec 2020. Standard errors in parentheses are robust and clustered by ETF, day, and time of the day. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6: **Absence of arbitrage by high-frequency trade during Covid-19 crisis**

	(1)	(2)	(3)	(4)	(5)
	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$	$\text{VolGap}_{i,d,t}$
High $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$	-0.23**	-0.25**	-0.24**	-0.21*	-0.17
	(0.10)	(0.12)	(0.11)	(0.10)	(0.09)
Low $\text{OLR}_i \times \text{Covid}_d \times \text{OLR}_{i,d,t-1}$	-0.05	-0.02	-0.01	0.03	0.00
	(0.08)	(0.09)	(0.09)	(0.10)	(0.09)
Frequency	5Sec/5Min	10Sec/10Min	15Sec/15Min	20Sec/20Min	30Sec/30Min
Controls	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes	Yes
Time of day FE	Yes	Yes	Yes	Yes	Yes
Observations	2,098,813	1,051,134	701,017	525,967	350,731
Adj. R-squared	0.264	0.302	0.264	0.217	0.198

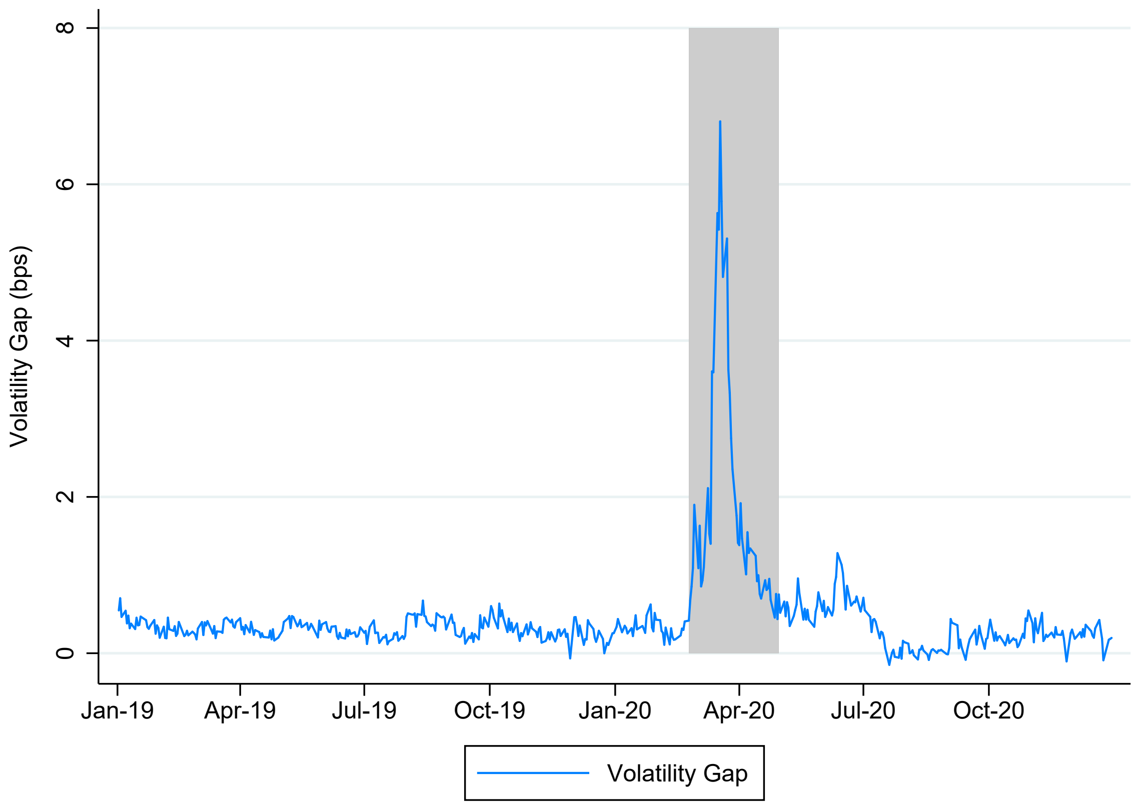
This table repeats the analysis of column (2) of **Table 4** with dependent variable Volatility $\text{Gap}_{i,d,t}$ and independent variable $\text{OLR}_{i,d,t-1}$ calculated at progressively lower frequencies: returns/volatility gaps computed at 5 second/5 minute, 10 second/10 minute, 15 second/15 minute, 20 second/20 minute, and 30 second/30 minute frequencies. All continuous variables are standardized to have a standard deviation of 1. Only the estimates of the triple interactions of tercile sorts, the Covid-19 period indicator, and the contemporaneous odd-lot ratio are reported. Variable definitions can be found in **Appendix Table A2**. The sample consists of all U.S. equity ETF that is not levered or inverse and has at least $5 billion asset under management as of the end of last month, from Jan 2019 to Dec 2020. Standard errors in parentheses are robust and clustered by ETF, day, and time of the day. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Figure 1: **Volatility during the Covid-19 crisis**



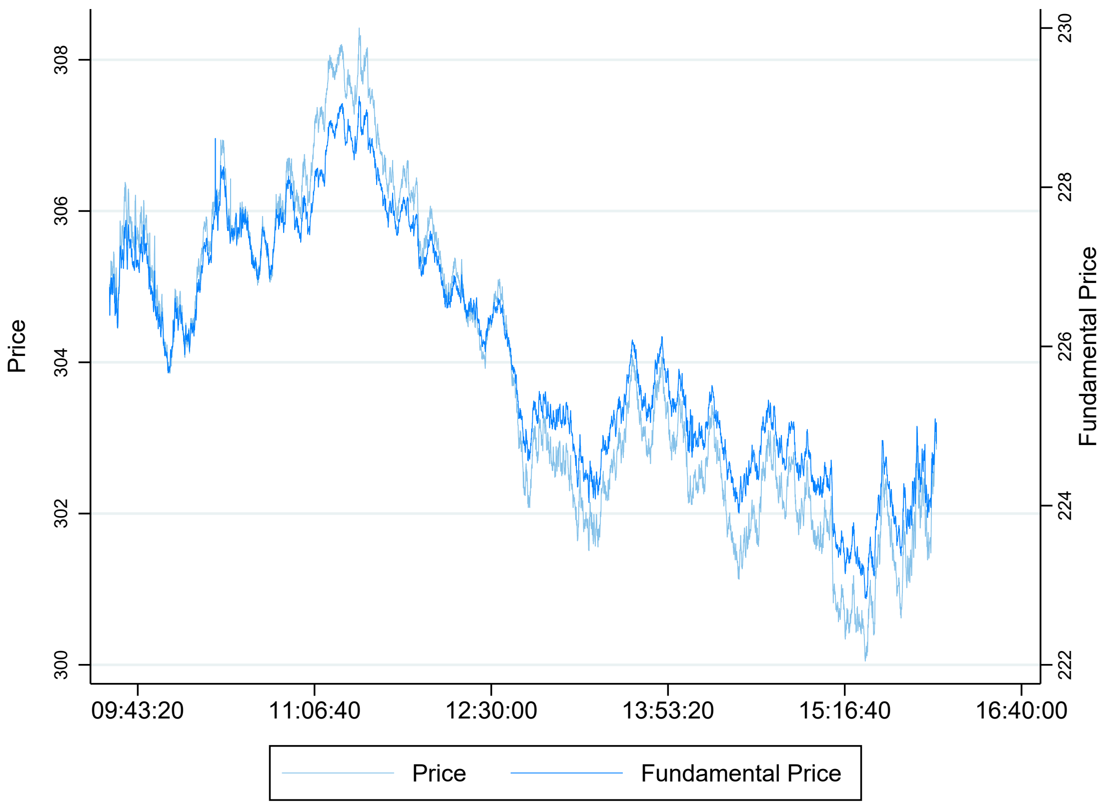
The figure plots the value of the VIX index and the average five-minute level volatility of ETFs in the sample from January 2019 to December 2020. The shaded area marks the Covid-19 crisis period between February 24, 2020 and April 30, 2020. ETF volatility is calculated as the standard deviation of five-second returns for each of the 78 five-minute intervals during the trading day. Variable definitions can be found in **Appendix Table A2**.

Figure 2: **Volatility Gap throughout Time**



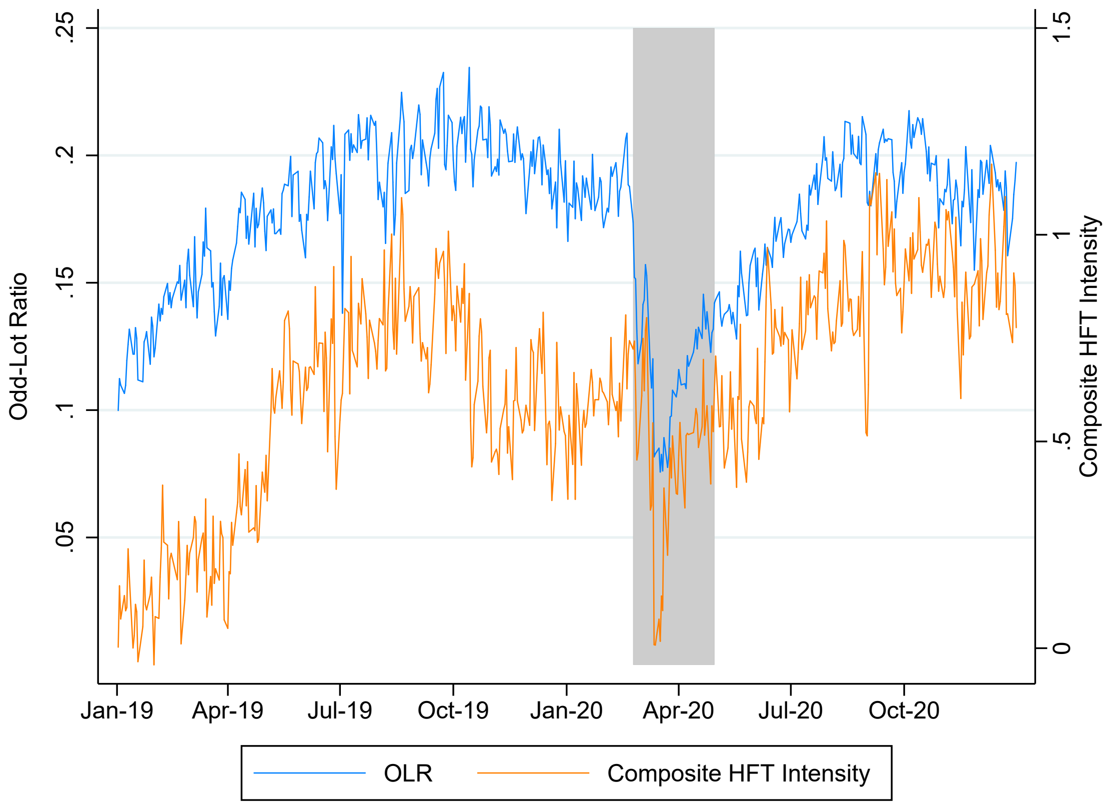
The figure plots the average volatility gap of the ETFs in the sample from January 2019 to December 2020. Volatility gap is the difference between volatility of ETF i and a synthetically constructed portfolio using its holding, on day d at the five-minute interval t. The shaded area marks the Covid-19 crisis period between February 24, 2020 and April 30, 2020. Variable definitions can be found in Appendix Table A2.

Figure 3: **ETF Price and Fundamental Price**



The figure plots the market price and the price of the fundamental portfolio of SPDR S&P 500 ETF (SPY) every five second on March 5, 2020.

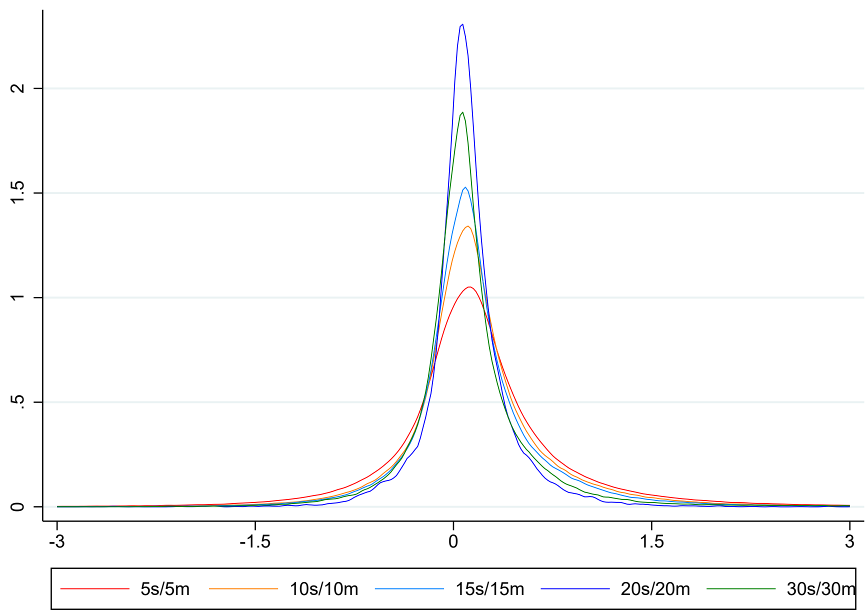
Figure 4: **High-frequency trading Trend throughout Time**



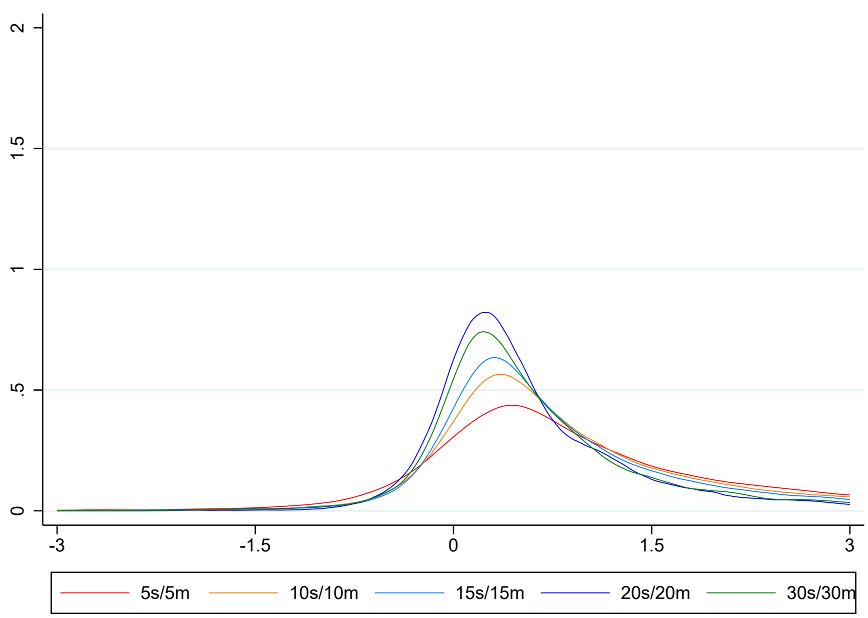
The figure plots the trend of the daily average of the odd-lot-ratio, $\text{OLR}_{i,d,t}$, and the composite HFT intensity, $\text{HFT}_{i,d}$ of the ETFs in the sample throughout time. The shaded area marks the Covid-19 crisis period between February 24, 2020 and April 30, 2020. Variable definitions can be found in **Appendix Table A2**.

Figure 5: **Volatility Gaps Measured at Different Frequencies**

(a) Control Period



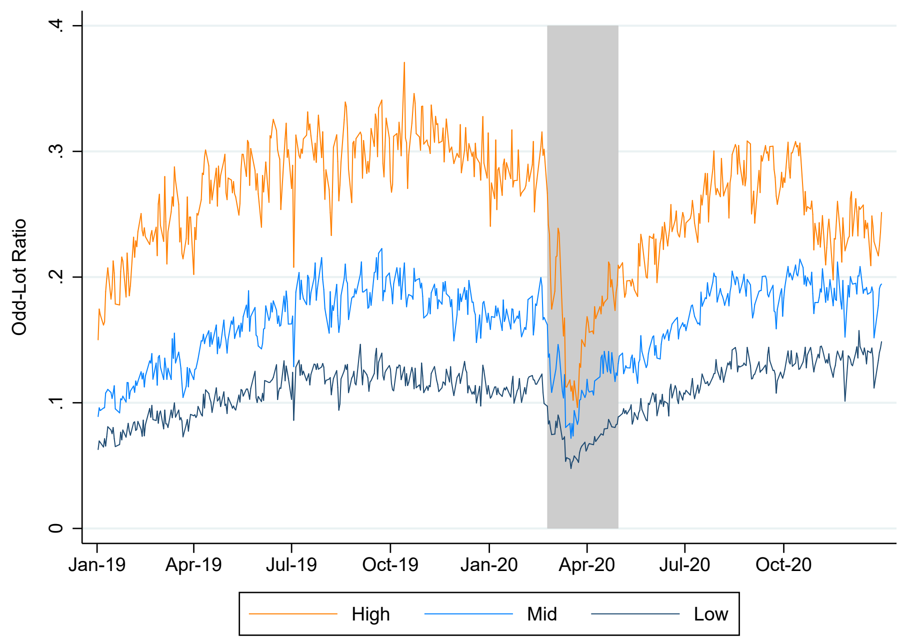
(b) Covid Period



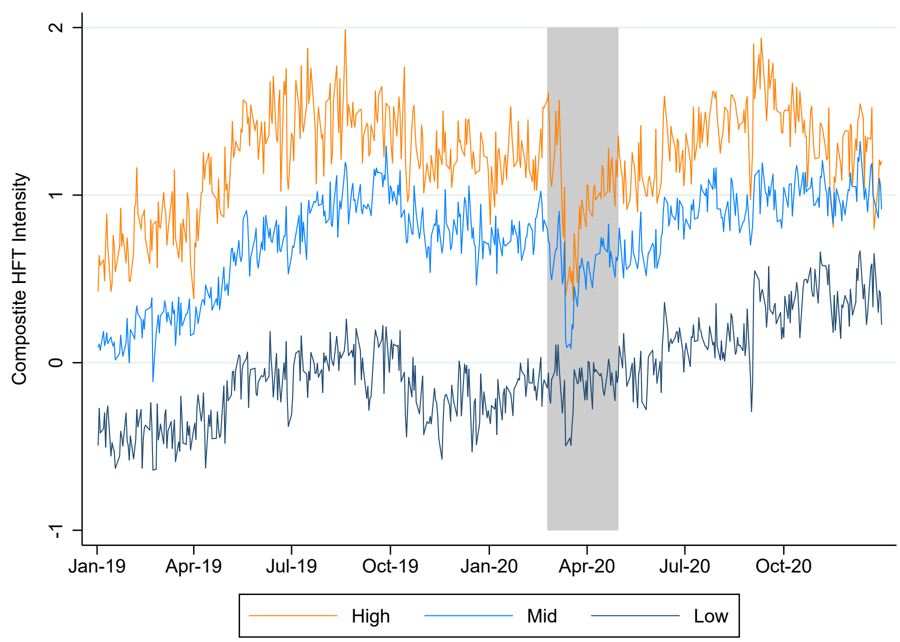
Panel A of the figure plots the kernel densities of the volatility gaps measured at different time frequencies during the control period and Panel B plots the kernel densities of the volatility gaps during the Covid period.

Figure 6: **High-frequency trading Trend throughout Time by Tercile**

(a) Odd-Lot Ratio Sort



(b) Composite HFT Intensity Sort



Panel A of the figure plots the trend of the composite HFT intensity, $\text{HFT}_{i,d}$, of the ETFs in the sample by terciles sorted based on the average HFT intensity as of January 2018. Panel B plots the trend of the daily average of the odd-lot ratio, $\text{OLR}_{i,d,t}$, based on odd-lot ratio sort as of January 2018. The shaded area marks the Covid-19 crisis period between February 24, 2020 and April 30, 2020. Variable definitions can be found in **Appendix Table A2**.

Table A1: ETF sample

Ticker	Name
DGRO	iShares Core Dividend Growth ETF
DIA	SPDR Dow Jones Industrial Average ETF
DVY	iShares Select Dividend ETF
FNDX	Schwab Fundamental US Large Company Index ETF
FVD	First Trust Value Line Dividend Index Fund
GSLC	Goldman Sachs ActiveBeta US Large Cap Equity ETF
HDV	iShares Core High Dividend ETF
IJH	iShares Core S&P Mid-Cap ETF
IJJ	iShares S&P Mid-Cap 400 Value ETF
IJK	iShares S&P Mid-Cap 400 Growth ETF
IJR	iShares Core S&P Small-Cap ETF
IJS	iShares S&P Small-Cap 600 Value ETF
IJT	iShares S&P Small-Cap 600 Growth ETF
ITOT	iShares Core S&P Total US Stock Market ETF
IUSG	iShares Core S&P US Growth ETF
IUSV	iShares Core S&P US Value ETF
IVE	iShares S&P 500 Value ETF
IVV	iShares Core S&P 500 ETF
IVW	iShares S&P 500 Growth ETF
IWB	iShares Russell 1000 ETF
IWD	iShares Russell 1000 Value ETF
IWF	iShares Russell 1000 Growth ETF
IWM	iShares Russell 2000 ETF
IWN	iShares Russell 2000 Value ETF
IWO	iShares Russell 2000 Growth ETF
IWP	iShares Russell Mid-Cap Growth ETF
IWR	iShares Russell Mid-Cap ETF
IWS	iShares Russell Mid-Cap Value ETF
IWV	iShares Russell 3000 ETF
MDY	SPDR S&P MidCap 400 ETF
MGK	Vanguard World Funds: Vanguard Mega Cap Growth ETF
OEF	iShares S&P 100 ETF
PRF	Invesco FTSE RAFI US 1000 ETF
QQQ	Invesco QQQ ETF
RSP	Invesco S&P 500 Equal Weight ETF
SCHA	Schwab US Small-Cap ETF
SCHB	Schwab US Broad Market ETF
SCHD	Schwab US Dividend Equity ETF
SCHG	Schwab US Large-Cap Growth ETF
SCHM	Schwab US Mid-Cap ETF
SCHV	Schwab US Large-Cap Value ETF
SCHX	Schwab US Large-Cap ETF
SDY	SPDR S&P Dividend ETF

Continued on next page

Table A1 – continued from previous page

Ticker	Name
SPLV	Invesco S&P 500 Low Volatility ETF
SPY	SPDR S&P 500 ETF Trust
SPYG	SPDR S&P 500 Growth ETF
VB	Vanguard Small-Cap ETF
VBK	Vanguard Small-Cap Growth ETF
VBR	Vanguard Small-Cap Value ETF
VIG	Vanguard Dividend Appreciation ETF
VO	Vanguard Mid-Cap ETF
VOE	Vanguard Mid-Cap Value ETF
VOO	Vanguard 500 ETF
VOT	Vanguard Mid-Cap Growth ETF
VTI	Vanguard Total Stock Market ETF
VTV	Vanguard Value ETF
VUG	Vanguard Growth ETF
VV	Vanguard Large-Cap ETF
VYM	Vanguard High Dividend Yield ETF

The sample is selected based on ETF's market capitalization and holdings. We select ETFs with market capitalization of $5 billion or higher as of December 2019 and with only domestic stocks in its holdings (CRSP objective code starting with ED). This yields 64 ETFs. VXF is excluded due to incomplete data on holdings from Bloomberg and MTUM, NOBL, QUAL, and USMV are excluded due to incomplete data on high-frequency trade intensity from SEC MIDAS database. This results in the final sample of 59 ETFs listed in the table.

Table A2: **Variable Definitions**

Variable Names	**Description**
$\text{Volatility}_{i,d,t}$	The volatility of ETF i on day d during five-minute interval t, calculated as the standard deviation of the 60 5-second returns of ETF i on day d during five-minute interval t. using TAQ.
$\widehat{\text{Volatility}}_{i,d,t}$	The volatility of the fundamental portfolio of ETF i on day d during five-minute interval t, calculated as the standard deviation of the 60 5-second returns of ETF i on day d during five-minute interval t. ETF fundamental portfolio price is calculated every five seconds using the five-second prices of the held securities from TAQ and their weights in the ETF holdings from Bloomberg.
$\text{VolGap}_{i,d,t}$	The difference between $\text{Volatility}_{i,d,t}$ and $\widehat{\text{Volatility}}_{i,d,t}$.
$\lvert\text{RetGap}\rvert_{i,d,t}$	The average of the absolute values of the 60 five-second return differences between ETF i and its fundamental portfolio on day d across five-minute interval t using data from TAQ.
$\text{OLR}_{i,d,t}$	The odd-lot ratio of ETF i on day d during five-minute interval t, calculated as the ratio of the volume of odd-lot trade and the volume of total trade. Odd-lot trade is identified using sale condition (I) in TAQ.
$\text{HFT}_{i,d}$	The composite high-frequency trade intensity of ETF i on day d, calculated as the sum of the z-score of odd-lot ratio, cancel-to-trade ratio, negative trade-to-order ratio, and negative average trade size using MIDAS. See Weller (2018) for details on the four individual components.
$\text{Price}_{i,d}$	The close price of ETF i on day d from CRSP.
$\text{Effective Spread}_{i,d}$	The share-weighted percentage effective spread of ETF i on day d from WRDS IID.
$\text{Asset under Management}_{i,m}$	The asset under management of ETF i at the end of month m, calculated as the product of month-end price times month-end shares outstanding from CRSP.
Covid_d	An indicator variable that equals 1 if day d is between February 24, 2020 and April 30, 2020.
High OLR_i	An indicator variable that equals 1 if ETF i is in the high average odd-lot ratio tercile during January 2018 and 0 otherwise.

Continued on next page

Table A2 – continued from previous page

Variable Names	**Description**
Low OLR_i	An indicator variable that equals 1 if ETF i is in the low average odd-lot ratio tercile during January 2018 and 0 otherwise.
High VolGap_i	An indicator variable that equals 1 if ETF i is in the high average volatility gap tercile during January 2018 and 0 otherwise.
Low VolGap_i	An indicator variable that equals 1 if ETF i is in the low average volatility gap tercile during January 2018 and 0 otherwise.
High Espread_i	An indicator variable that equals 1 if ETF i is in the high average effective spread tercile during January 2018 and 0 otherwise.
Low Espread_i	An indicator variable that equals 1 if ETF i is in the low average effective spread tercile during January 2018 and 0 otherwise.