DOI: 10.1111/fima.12395

ORIGINAL ARTICLE

Financial Management | WILEY

Diving into dark pools

Sabrina Buti[1] | **Barbara Rindi**[2,3] | **Ingrid M. Werner**[4,5]

[1]Université Paris Dauphine-PSL, CNRS, DRM, Paris, France

[2]Bocconi University and IGIER, Milan, Italy

[3]Baffi-Carefin, Milan, Italy

[4]Fisher College of Business, The Ohio State University, Columbus, Ohio, USA

[5]CEPR, London, United Kingdom

Correspondence
Ingrid M. Werner, Fisher College of Business, Ohio State University, 2100 Neil Ave., Columbus, OH 43210, USA.
Email: werner.47@osu.edu

Abstract

We study 2009 and 2020 dark trading for U.S. stocks. Dark trading is lower when volume is low, volatility high, and in periods of markets stress. Dark pools are more active for large caps, while internalization is more common for small caps. Traders use dark pools to jump the queue for large caps in 2009, and to avoid crossing the spread for small caps in both years. Internalization is higher when spreads are wide and depth is high. Dark pool trading improves spreads in 2009, but worsens market quality for large caps in 2020. We discuss explanations for the change.

1 | INTRODUCTION

In its *Concept Release on Equity Market Structure* (SEC, 2010), the U.S. Securities and Exchange Commission (SEC) raises concerns about the consequences of a rising dark pool market share on public order execution quality and price discovery. More recently, in its *Staff Report on Equity and Options Market Structure Conditions in Early 2021* (SEC, 2021), the SEC raises concerns about the large amount of volume that is executed away from the lit markets by internalizing over-the-counter (OTC) market makers, particularly during periods of market stress. To help inform the regulatory debate, we use data from both 2009 and 2020 to study dark trading. Specifically, we examine two main research questions. What factors influence order routing to dark pools and internalizing OTC market makers? Does dark trading affect market quality?

There are several reasons for why institutional traders may want to avoid displaying their orders in continuous limit order markets. First of all, dark trading offers market participants additional liquidity that allows them to minimize market impact costs and also to trade inside the lit market spread.[1] In addition, order display invites imitation, potentially reducing the alpha of the underlying investment strategy. Displayed orders also invite front running by broker-dealers as well as by opportunistic traders, resulting in higher trading costs. Moreover, institutional traders

[1] See Mittal (2008) for a discussion of dark pool characteristics.

worry about the risk of trading against informed order flow, especially order flow from proprietary trading desks. *Regulation National Market System (NMS)* (SEC, 2005) opened the door for broker-operated nondisplayed liquidity venues, so called dark pools. Dark pools have limited or no pretrade transparency reducing the problems of imitation and front running. They also control access, potentially reducing the risks of facing informed order flow.

Despite these advantages, regulators and exchanges have expressed concerns that the migration of orders flows to dark markets may have two main detrimental effects on the quality of the lit markets: first, it may divert volume away from lit markets thus discouraging liquidity provision and harming liquidity; second, having dark markets no or little pretrade transparency, it may harm price discovery. In addition, to increase executed volumes, dark pool operators have gradually opened to their own proprietary trading desks and to High Frequency Traders (HFTs), thus exacerbating the potential conflict of interest with their clients.[2] Finally, in some dark pools the execution price is derived from the primary lit markets. This derived pricing rule may therefore lead to transactions based on stale reference prices, thus exposing investors trading in the dark to adverse selection costs (Aquilina et al., 2021).

While our focus is on dark pools, which executed less than 10% of U.S. share volume in 2009 and about 14% in 2020, these venues are not the only way in which trading occurs away from lit exchanges.[3] OTC market makers internalize roughly 24% of U.S. share volume in 2020. This is mainly orders routed to OTC market makers by retail brokerage firms, but OTC market makers also interact with institutional order flow. We incorporate internalization by OTC market makers throughout the analysis. Moreover, we control for trading in lit venues that compete with the listing exchange. This allows us to study the complex ways in which dark trading may affect market quality.

Our 2009 data come from a survey conducted by the Securities Industry and Financial Market Association (SIFMA) on our behalf. SIFMA solicited daily stock-level dark pool share-volume data from all their members operating dark pools. Participation was voluntary, and SIFMA in the end obtained data from 11 dark pools. The SIFMA sample allows us to examine dark pool activity for over 3000 stocks. For 2009, the only publicly available data on dark trading are the trades reported to one of the Trade Reporting Facilities (TRFs) thar aggregates dark pools and trades internalized by OTC market makers. We use this information to proxy for 2009 internalized trades by subtracting SIFMA-reported volume from TRF-reported volume, recognizing that this measure includes the dark pools that chose not to report their trades to SIFMA. We supplement the 2009 self-reported data with weekly Financial Industry Regulatory Authority (FINRA) OTC Transparency data for 2020. The 2020 data are comprehensive, and we are able to study trading activity in Alternative Trading Systems (ATSs) and internalized by OTC market makers (Non-ATSs) for over 2900 stocks.

Our samples each include a dramatic decline in the stock market and elevated uncertainty, associated with the Great Financial Crisis in 2009 and the COVID pandemic in 2020, followed by a steady recovery. This allows us not only to study if and how dark trading has changed over time, but also to examine the role of dark trading during periods of market stress. The increase in the CBOE Volatility Index (VIX) was particularly dramatic in March 2020, and we find that the results are sensitive to including this period. Therefore, we report results for 2020 both overall and excluding the period February 15 to April 15, 2020 (ex-COVID sample).

Dark trading overall is relatively stable in each sample year despite the tumultuous stock market, but it does exhibit an increasing trend in 2009. We find that dark pools are more active for large capitalization firms than for small capitalization firms, while OTC market makers internalize more for small capitalization than for large capitalization firms. Consequently, we control for firm by quarter fixed effects throughout our analyses.

We first examine how order routing to dark venues depends on market conditions, such as price, volume, and volatility, as well as on instruments for order book characteristics. For small caps, consistently across our sample periods, order routing to dark pools increases when the book is less competitive (proxied by depth in 2009 and by both

[2] In recent years, the SEC has brought several enforcement actions against dark pool operators due to the improper management of conflicts of interest. See, for example, SEC (2014a, 2014b, 2015). Also notice that despite being sophisticated market participants, institutional traders may struggle to detect issues with brokers' execution quality due to the lack of transparency, potentially paying higher transaction costs (Anand et al., 2021).

[3] Rosenblatt Securities, Inc. started tabulating monthly share volume for dark pools in its Trading Talk publication in 2008 and TABB Group started its Liquidity Matrix publication in 2007. Since 2014, the Financial Industry Regulatory Authority (FINRA) collects Alternative Trading System level data on trading volume on a weekly basis, and from 2016 the data include trades by OTC market makers.

spread and depth in 2020), the relative tick size is large, and trading activity is high, suggesting that traders value the advantage of sourcing liquidity on dark venues when the lit market is illiquid due to either low depth or wide spreads.

For large caps, how traders route orders to dark pools changes over our sample periods. In 2009, dark pool market share for large caps is higher when the book is more competitive, suggesting the ability to jump the long queue at the inside of the order book is particularly valuable for large caps, especially when the tick size is large. By contrast, in 2020 order routing to dark pools for large caps is unrelated to the state of the book, except during the COVID period when traders seem to use dark pools to avoid crossing the wide spread and the opportunity to execute within the quoted spread is more important.

We find that OTC market makers tend to internalize more orders when depth is high. In 2009, we find evidence that the market share of OTC market makers is also higher when spreads are wide. This evidence suggests that payments for order flow arrangements are more profitable for internalizing market makers when spreads are wide, and a higher depth in the lit market means that OTC market makers can more easily offset order imbalances. Taken together, these results show that the effect of market conditions and order book characteristics on market shares differs between categories of dark trading, as well as across stocks within a particular category.

We investigate how dark trading affects market quality using a simultaneous equation system where we instrument for dark trading to account for the fact that market quality and dark trading are jointly determined. We document that aggregate dark trading leads to lower spreads in 2009, but does not significantly affect our market quality measures in 2020. Separating the two forms of dark trading, we find that both higher dark pool market share and more internalization by OTC market makers lead to lower spreads in 2009. This is generally true for subsamples by firm size as well as overall. By contrast, we find no effect of dark trading on short-term volatility in 2009. In the later sample, more dark pool trading leads to higher short-term volatility overall, and both wider spreads and higher short-term volatility for the ex-COVID sample. By examining the subsamples by size, we show that the negative effect of dark pool trading on market quality derives from large capitalization stocks. We find no evidence that more internalization affects market quality in 2020.

Finally, we study whether dark trading plays a different role during periods of market stress, defined as the first 6 months of each year, and days (weeks) with low returns, high selling pressure, or high volatility. The market shares of dark pools and OTC market makers are generally lower during periods of market stress. In 2009, higher dark trading (of either type) during days when markets are under stress leads to narrower spreads and lower short-term volatility. By contrast, higher dark pool trading leads to wider spreads and higher short-term volatility during weeks with low returns and high volatility in 2020.

These differences in the effects of dark trading—between types of dark trading, across stocks, and between sample periods—highlight the complex ways that dark trading affects market quality for U.S. stocks. They illustrate that it can be misleading to focus on one form of fragmentation, on one part of the cross-section of stocks, or on one specific time period. By comparing the 2009 and the 2020 sample periods, we see that changing market conditions, the development of new venues, as well as the practices of market participants can significantly affect inference regarding the role dark trading plays in markets. We speculate that the difference between the two samples arises because more proprietary order flow, HFTs, and informative retail order imbalances reach dark pools in 2020, and that these venues have become less attractive for institutional traders as a result.

This paper contributes to the literature on dark pools in several ways.[4] Our study is the first broad cross-sectional study documenting dark trading in U.S. equity markets. The evidence we present suggests that studying aggregate dark trading instead of its components can lead to very different conclusions (Degryse et al., 2015). Prior studies have emphasized that different types of dark pool pricing mechanisms (Foley & Putniņš, 2016) and dark pool trade sizes (Comerton-Forde & Putniņš, 2015) may have different effects on market quality, and we show that there are also differences between trades internalized by OTC market makers and dark pool trades (Kwan et al., 2015). Prior work has focused on the effect of dark pools on market quality for Large capitalization securities, missing the important

[4] We discuss the extensive existing literature on fragmentation in Section 3 of the Supporting Information.

cross-sectional variation documented in our paper (e.g., Comerton-Forde & Putniņš, 2015; Degryse et al., 2015; Foley & Putniņš, 2016). Finally, we show that the nature of dark trading and its effect on market quality have changed significantly in recent years. However, we also find that a few effects persist across our sample periods. For small stocks, when the lit market is less competitive with either low depth or large spread but trading activity is high and volatility is low, traders hunt liquidity in the dark venues and therefore dark order routing increases. We also find that, overall, both dark trading and market makers internalization tend to be lower during periods of market stress.

The paper proceeds as follows. Section 2 describes our samples and provides descriptive statistics. How firm and order book characteristics influence order routing is discussed in Section 3. Section 4 studies the relationship between dark trading and measures of market quality. Market stress is the focus of Section 5. We discuss the results in Section 6, and Section 7 concludes.

2 | DATA AND DESCRIPTIVE STATISTICS

SIFMA solicited daily data on stock-level dark pool share volume for the 2009 calendar year from all their members operating dark pools. The reporting was voluntary, and SIFMA collected data on daily single-counted share volume from 11 dark pools. The data are daily share volume per security for each of the 11 dark pools, but the data include no names of the dark pools. Our agreement with SIFMA precludes us from study the data for individual (or groups of) dark pools. Therefore, we are unable to report results for individual dark pools, or results divided into groups of dark pools by the type of ownership or by the execution algorithm.[5] Figure SA1 shows that our SIFMA raw data represent between 47% and 60% of dark volume as reported by Rosenblatt, Inc. in their *Let There Be Light* publication (Gawronski & Schack, 2010). We screen the SIFMA dark pool data as described in Section 1 of the Supporting Information, which results in a sample with a cross-section of 3098 securities. We aggregate the daily share volume across reporting venues into a stock-day series (*DP*). We use DTAQ to calculate daily total dark share volume reported to one of the TRFs (*TRF*), and lit competing share volume (*COMP*) as share volume reported to one of the transparent venues that compete with the listing exchange.[6] All registered exchanges can trade all U.S. stocks through unlisted trading privileges. Hence, for each listing exchange, there were as many as a dozen lit competing venues. There is no data source for internalized trades in 2009, so we proxy for internalized trades (*INT*) by subtracting dark pool share volume from TRF share volume for each stock day. Note that while this measure includes primarily internalized trades, it also includes the dark pools that did not report to SIFMA. Finally, we express each measure of fragmentation as a fraction of consolidated share volume.

To compare our original 2009 SIFMA data to a more recent period, we download data from FINRA for 2020. Since 2014, FINRA has been publishing security-level weekly OTC Transparency data for ATSs, and FINRA augmented the data to include weekly data for individual OTC market makers grouped together under the Non-ATS header in 2016.[7] We screen the FINRA data as described in Section 2 of the Supporting Information, which results in a 52-week sample with a cross section of 2902 securities. For each stock and week, we aggregate the ATS share volume into a variable *ATS*, and the volume reported by OTC market makers into a variable *Non-ATS* and label the sum of these two as *FINRA*. The advantage of the FINRA OTC Transparency data is that all dark pools and all internalizing OTC market makers are covered. The drawback is that data are only available weekly, and we expect to have less power as a result. We use the SEC's Market Information Data Analytics System (MIDAS) to calculate 2020 weekly share volume for lit competing venues as share volume reported to one of the transparent venues that compete with the listing exchange, and express each measure as a fraction of consolidated share volume.

[5] Foley and Putniņš (2016) using Canadian data find that dark pools that enable traders to supply two-sided liquidity inside the lit market spread improve lit market quality, while those that execute at the midpoint have no effect on market quality. By contrast, Comerton-Forde et al. (2018) and IIROC (2015) find no significant effect of either type of dark pool activity on market quality using the same data.

[6] TRF trades appear with exchange code "D" in DTAQ data.

[7] For details, see https://www.finra.org/filing-reporting/otc-transparency.

Figure 1 illustrates the variation in the three daily dark market shares averaged across stocks for 2009 (Panel A), and in the three weekly market shares averaged across stocks for 2020 (Panel B). We superimpose the VIX and the S&P 500 index values in each panel. Both years display a very large stock market decline followed by a rapid recovery in the first 6 months, and a calmer stock market in the second half of the year. Volatility spikes during the rapid stock market decline in the first half, particularly during 2020. Therefore, we also analyze a restricted 2020 ex-COVID sample where we exclude 9 weeks between February 15 and April 15, 2020. Dark fragmentation is much less volatile.[8] It increases gradually for the 2009 sample, but there is no noticeable secular trend in 2020.

Figure 1 hides significant cross-sectional variation in dark trading. We visualize this variation in Figure 2 where we plot time-series average measures of dark trading against the natural logarithm of previous year-end market capitalization, *log(Size)*, for 2009 (Panel A) and 2020 (Panel B). Overall dark trading is declining in *log(Size)* regardless of sample period. The next two plots in each panel show that there is a clear difference between dark pool trading and internalization. While internalization (*INT* and *Non-ATS*) is declining in *log(Size)*, dark pool trading (*DP* and *ATS*) is increasing in *log(Size)*. Hence, internalization is higher for small capitalization stocks than for large capitalization stocks on average, and the opposite is true for dark pool trading. The overall cross-sectional patterns are very similar across the 2 years, despite the fact that the data sources are very different and that we do not have comprehensive dark pool data for 2009. Figures 1 and 2 suggest that we should follow Degryse et al. (2015) and use stock-by-quarter fixed effects in our panel regressions to control for the slow-moving trend and the significant cross-sectional variation in fragmentation.

We draw information on size and daily market conditions from CRSP, including market capitalization, share volume, closing stock price, and volatility (defined as [high–low]/high based on quotes). We also compute daily market quality measures from DTAQ for 2009. We draw daily market quality measures from the WRDS Intraday Indicators for 2020. To match our weekly FINRA data for 2020, we average the daily market condition and market quality data to create a weekly panel. To reduce the influence of outliers, we impose further screens on the data. We exclude stock-days where there is no reported consolidated volume in CRSP, where there are fewer than 20 trades per day in TAQ or WRDS Intraday Indicators, and we exclude early closing days around holidays. Finally, we drop stock-days (stock-weeks) where the SIFMA (FINRA) reported dark volume exceeds the consolidated volume as reported in CRSP.

Our market quality measures include stock-level daily time-weighted National Best Bid Offer (NBBO) quoted spreads, share-weighted effective half-spreads, and the standard deviation of mid-quote returns measured over 15-min (quote-update) intervals for 2009 (2020).[9] Short-term volatility is a measure of trading frictions, and a market with lower volatility is more efficient. We multiply this variable by 10,000 so it is in basis points. We express quoted and effective spreads as a percentage of the mid-quote. We measure depth computed as the time-weighted average bid and offer depths in shares at the NBBO. To address the significant skewness in the data, we follow the literature and take the natural logarithm of both market conditions and market quality measures in our regression analyses.[10]

Table 1 summarizes the descriptive statistics for our stock-day 2009 sample in Panel A, for our stock-week 2020 overall sample in Panel B, and for our stock-week 2020 ex-COVID sample in Panel C. The median 2009 firm size is $483 million, volume is 270 thousand shares, stock price is $13.71, and volatility is 6.18%. The median 2020 firm size for the overall (ex-COVID) sample is $1.0 ($1.1) billion, volume is 1.9 (1.8) million shares, stock price is $22.34 ($22.80), and volatility is 4.62% (3.99%). Thus, the median firm is larger and has higher volume, a higher stock price, and lower volatility in 2020 (especially for the ex-COVID sample) compared to 2009. Despite this, we find that the median firm faces worse market quality in 2020 than in 2009. The median 2009 quoted spread is 24.78 basis points, effective half-spread is 7.72 basis points, depth is 474 shares, and standard deviation of 15-min mid-quote returns is 48.11 basis points. The median 2020 quoted spread for the overall (ex-COVID) sample is 34.42 (31.42) basis points, the effective

[8] There is a large spike in early 2020 for the three market shares. Our results are robust to excluding this week.

[9] WRDS Intraday Indicators do not include 15-min measures of standard deviation based on mid-quote returns. We also repeated all our analyses for the variance ratio, and this variable is unaffected by dark trading for both years.

[10] Comerton-Forde and Putniņš (2015), Degryse et al. (2015), and Foley and Putniņš (2016) take the natural logarithm of market quality and firm characteristics. Furthermore, we Winsorize market quality measures daily at 1% and 99% to deal with significant outliers. Table SA1 reports descriptive statistics in logs.

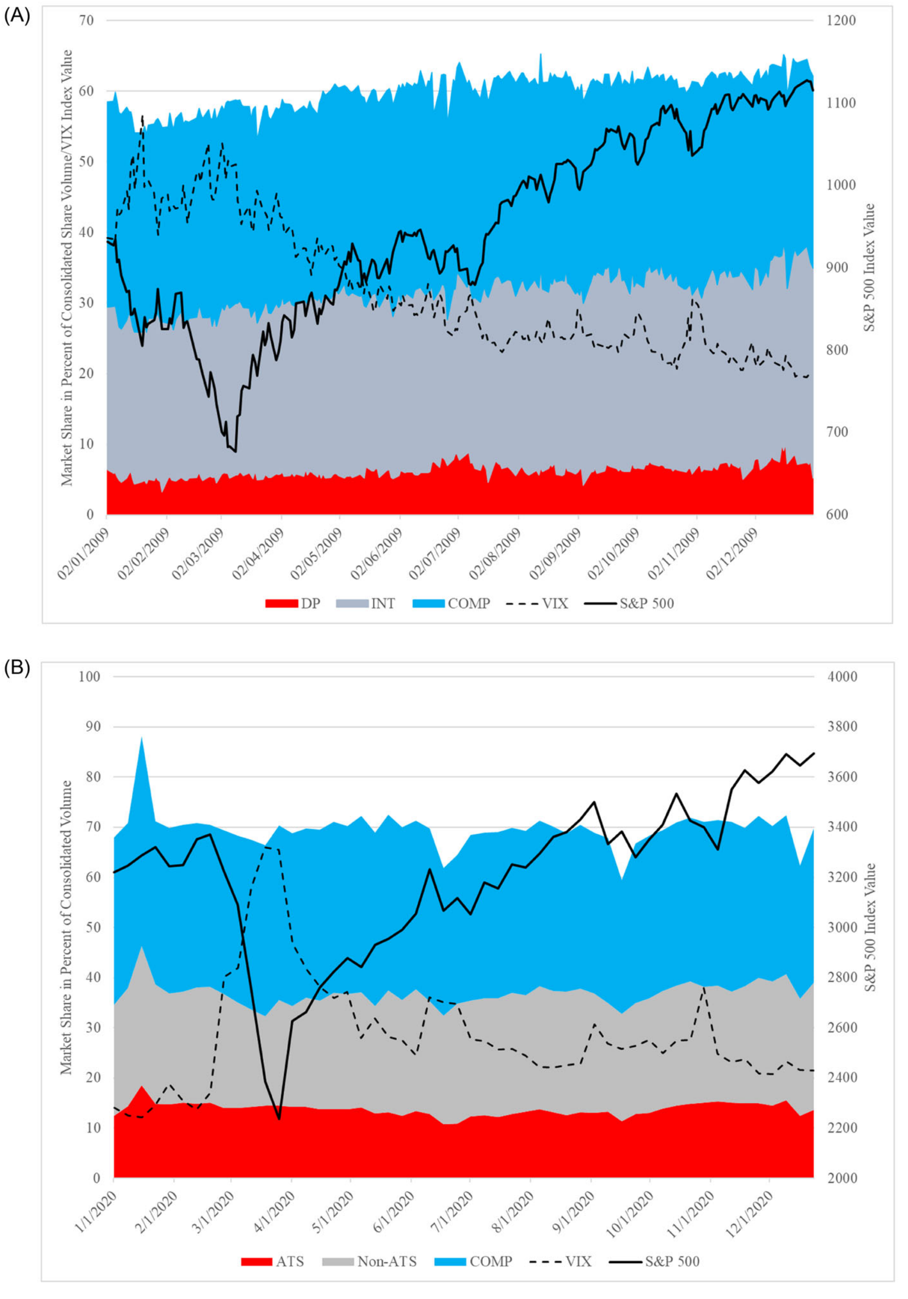

FIGURE 1 Fragmentation over time. The figure reports daily average market shares expressed in percent of consolidated share volume for three forms of fragmentation for 2009 in Panel A: *DP* is defined as SIFMA sample dark pool single-counted share volume; *INT* is defined as share volume reported to Trade Reporting Facilities (*TRF*) minus *DP*; and *COMP* is defined as share volume reported to lit venues excluding the exchange where the stock is listed. For 2020 in Panel B, *ATS* is defined as Alternative Trading Systems reporting to FINRA, *Non-ATS* are OTC market makers reporting to FINRA, and *COMP* is defined as for the 2009 sample. We plot the CBOE Volatility Index (VIX) on the left vertical axis, and the S&P 500 index on the right vertical axis [Color figure can be viewed at wileyonlinelibrary.com]

(A) 2009 sample dark fragmentation for *TRF*, *DP*, and *INT*. (B) 2020 sample dark fragmentation for *FINRA*, *ATS*, and *Non-ATS*.

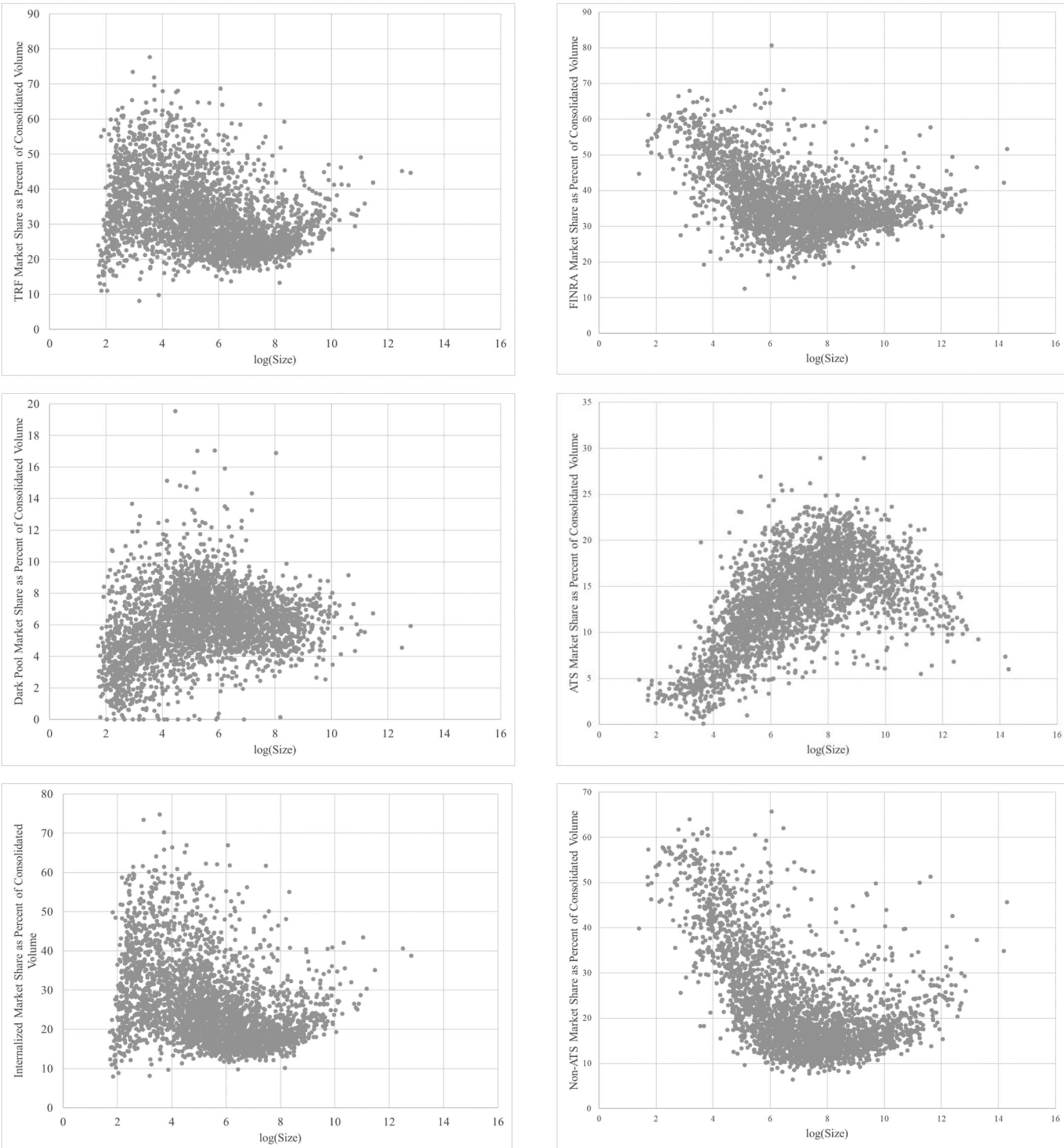

FIGURE 2 Dark fragmentation by size

half-spread is 9.68 (8.75) basis points, depth is 282 (281) shares, and standard deviation of mid-quote returns is 4.08 (3.62) basis points (not directly comparable to 2009 where we have 15-min returns). As expected, the 2020 ex-COVID sample has not only lower volatility (as measured by the intraday range) and lower standard deviation of returns, but also lower spreads than the full year sample.

It is well known that U.S. equity trading is highly fragmented (see O'Hara & Ye [2011] and the references therein). The fragmentation measures for our samples are reported in the bottom third of each panel in Table 1. *TRF* (*FINRA*) represents 30.5% (35.2%) of share volume, while trades reported to competing exchanges (*COMP*) represents 28.5%

TABLE 1 Descriptive statistics

A. 2009 Sample	*N*	Mean	*SD*	p25	p50	p75
Characteristics						
Size	3098	3.345	13.856	0.157	0.483	1.699
Volume	3098	1756	11,296	74	270	1017
Price	3098	19.07	23.62	6.63	13.71	24.95
Volatility	3098	6.77	3.13	4.60	6.18	8.37
Market quality						
Quoted spread	3098	65.11	93.71	11.14	24.78	72.21
Effective half-spread	3098	16.49	20.57	3.69	7.72	20.11
Depth	3098	884	1258	319	474	881
StD returns	3098	68.15	27.68	48.11	63.80	82.88
Fragmentation						
TRF	3098	0.324	0.098	0.250	0.305	0.380
DP	3098	0.060	0.021	0.047	0.060	0.074
INT	3098	0.263	0.102	0.187	0.236	0.312
COMP	3098	0.279	0.053	0.244	0.285	0.317
B. 2020 Sample	***N***	**Mean**	***SD***	**p25**	**p50**	**p75**
Characteristics						
Size	2902	9.565	50.210	0.243	1.021	4.126
Volume	2902	5953	12,527	589	1925	5305
Price	2902	45.57	61.74	10.04	22.34	54.87
Volatility	2902	4.99	1.85	3.63	4.62	6.13
Market quality						
Quoted spread	2902	80.06	112.20	16.98	34.42	85.73
Effective half-spread	2902	24.42	34.68	4.88	9.68	26.10
Depth	2902	711	1439	188	282	589
StD returns	2902	10.73	15.34	2.01	4.08	12.10
Fragmentation						
FINRA	2902	0.369	0.084	0.312	0.352	0.409
ATS	2902	0.137	0.047	0.108	0.142	0.170
Non-ATS	2902	0.231	0.112	0.149	0.192	0.282
COMP	2902	0.329	0.073	0.281	0.312	0.395
C. 2020 ex-COVID	***N***	**Mean**	***SD***	**p25**	**p50**	**p75**
Characteristics						
Size	2902	9.847	52.160	0.250	1.053	4.238
Volume	2902	5558	11,771	553	1798	4944
Price	2902	46.96	63.91	10.20	22.80	56.24
Volatility	2902	4.36	1.76	3.06	3.99	5.42

(Continues)

TABLE 1 (Continued)

C. 2020 ex-COVID	*N*	Mean	*SD*	p25	p50	p75
Market quality						
Quoted spread	2902	72.41	102.80	15.17	31.42	77.85
Effective half-spread	2902	21.98	31.38	4.36	8.75	23.90
Depth	2902	717	1482	189	281	584
StD returns	2902	9.60	13.87	1.75	3.62	10.63
Fragmentation						
FINRA	2902	0.372	0.085	0.313	0.355	0.415
ATS	2902	0.136	0.047	0.107	0.140	0.169
Non-ATS	2902	0.235	0.113	0.153	0.199	0.289
COMP	2902	0.327	0.074	0.278	0.309	0.394

Note: The table reports descriptive statistics based on daily stock-level data for the 2009 daily sample in Panel A, the 2020 overall weekly sample in Panel B, and the 2020 ex-COVID weekly sample where we exclude 8 weeks of observations between February 15, 2020 and April 15, 2020 in Panel C. We obtain market capitalization in billion dollars, stock price in dollars, and *Volatility* is 100* (High–Low)/High from CRSP. For 2009, we use TAQ to calculate average daily: *Volume* as consolidated share volume divided by 1000, time-weighted *Quoted spread* at the National Best Bid Offer (NBBO) and share-weighted *Effective half-spread* both in basis points of the mid-quote, *Depth* is the time-weighted shares at the NBBO, and *StD returns* is 10,000 times the standard deviation of 15-min mid-quote returns. For 2020, we use CRSP and WRDS Intraday Indicators to calculate the average daily characteristics and market quality measures for each week. Note that *StD returns* for 2020 is 10,000 times the standard deviation of mid-quote returns. For 2009, fragmentation is measured as a fraction of TAQ consolidated volume, where *COMP* are lit competing venues with the main (listing) exchange, *TRF* is volume reported with "D" in TAQ, and *DP* are the SIFMA reporting dark pools and *INT* is Internalized trades defined as *TRF-DP*. For 2020, fragmentation is measured as a fraction of WRDS Intraday Indicator reported daily consolidated volume aggregated to weekly data, where *COMP* are lit competing venues from MIDAS and *FINRA* is *ATS* and *Non-ATS* volume as reported to FINRA. Data are Winsorized at the 1st and 99th percentiles.

(31.2%) of share volume for the median firm in 2009 (2020).[11] This means that overall fragmentation has increased substantially, and the listing exchange captures a smaller fraction of trading activity in 2020 compared to 2009. Internalized trading—*INT* (*Non-ATS*)—represents 23.6% (19.2%) of share volume for the median firm in 2009 (2020). This does not mean that internalization is declining. Recall that *INT* for 2009 comprises internalized trades, but also trades executed in dark pools that did not voluntarily report to SIFMA. The dark pool market share—*DP* (*ATS*)—for the median firm is 6.0% (14.2%) in 2009 (2020). Since the SIFMA sample represents roughly half of all dark trading in 2009 (see Figure SA1), it appears that dark pool trading has increased somewhat. Finally, we note that the level of dark trading in both our samples is higher than the 25% figure reported by Degryse et al. (2015) for European stocks in 2009 (dark pools, internalization, and over-the-counter). Dark trading for Australia in 2012 is 18% (dark pools and block trades) according to Comerton-Forde and Putniņš (2015), and it was 8.5% for Canada that same year according to Foley and Putniņš (2016).

It is clear from Table 1 that both samples span stocks with very different firm characteristics, market quality, and levels of dark trading. We believe this very diverse set of stocks will help us better understand the full role of dark trading in securities markets. In our analysis of dark trading, we control for competition from lit venues (*COMP*) as suggested by Degryse et al. (2015).

[11] As the breakdown of fragmentation is very similar in the 2020 samples, we only report in parenthesis statistics for the overall 2020 sample.

3 | ORDER ROUTING

When deciding where to send an order, a smart order router takes into account asset-, order-, and market-level characteristics as inputs, and uses this information to predict the fill probability for each venue. The characteristics "include all the factors that may influence fill rates: each exchange's market share, the state of the limit order book (e.g., the depth of each market at the inside price), trading volume, price level, volatility, asset type…" (Bacidore, 2020, p. 162). Data availability necessitates that we focus on a parsimonious specification to capture the main aspects of the order-routing process. We proxy for the state of the limit order book using NBBO depth and the inside quoted spread. Clearly, order routing decisions affect the limit order book, and we therefore use the lagged NBBO depth and the lagged inside quoted spread as instruments for the endogenous contemporaneous limit order book characteristics. We capture the market conditions by including price, share volume, and the intraday range defined as the (High–Low)/High. Since the tick size is constant at one cent for all our sample stocks, the stock price maps into the relative tick size (a high price means a low relative tick size).

To examine how order routing ($OR_{i,t}$) varies with order book characteristics, we run the following IV/2SLS daily panel regressions:

$$log(Quotedspread)_{i,t} = a_i d_q + b_{1,1} log(Quotedspread)_{i,t-1} + b_{1,2} log(Depth)_{i,t-1} + c_1 X_{i,t} + e_{i,t}, \tag{1}$$

$$log(Depth)_{i,t} = a_i d_q + b_{2,1} log(Quotedspread)_{i,t-1} + b_{2,2} log(Depth)_{i,t-1} + c_2 X_{i,t} + e_{i,t}, \tag{2}$$

$$OR_{i,t} = a_i d_q + \beta_1 log(Quotedspread)_{i,t} + \beta_2 log(Depth)_{i,t} + \gamma X_{i,t} + e_{i,t}, \tag{3}$$

where $X_{i,t}$ is a vector of control variables that captures the market conditions: $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$; we control for firm-by-quarter fixed effects ($a_i\ d_q$), and standard errors are clustered by stock and day. Table 2 reports the results for the second-stage regressions in Equation (3).[12] The first-stage regressions (Table SA2.1) for 2009 show that the lagged instruments for both own-effects are positive and significant ($b_{1,1} = 0.3678^{***}$ and $b_{2,2} = 0.3424^{***}$). The cross-effects are negative and small ($b_{1,2} = -0.0107^{***}$ and $b_{2,1} = -0.0192^{***}$), but significant. For the 2020 full sample, only the own-effects are positive and significant ($b_{1,1} = 0.3306^{***}$ and $b_{2,2} = 0.2743^{***}$). For the 2020 ex-COVID sample, only the own-effect for depth is positive and significant ($b_{2,2} = 0.2325^{***}$), while for spread the own-effect is positive and significant only for Small stocks ($b_{1,1} = 0.1885^{***}$).

Our variables *TRF* (2009) and *FINRA* (2020) capture aggregate order routing to dark venues. However, as discussed in Section 2, these data consist of two broad categories of dark trades, those that execute in dark pools and trades internalized by OTC market makers. For completeness, we also report results for order routing to lit competing exchanges. Access to lit competing venues enables liquidity providers to bypass time-priority on the listing exchange.

TRF order routing in 2009 is increasing in both our instruments for quoted spread and depth. However, when the aggregate *TRF* is decomposed into dark pools and internalization, order routing to dark pool is negatively related to both our metrics of market quality, whereas internalization is positively related to both spread and depth, which is puzzling as it is unclear whether routing to dark pools or internalizing OTC market makers are related to more or less liquid books.

FINRA order routing is unrelated to the book characteristics for the 2020 sample, but it is increasing in the spread for the 2020 ex-COVID sample. However, when we decompose the aggregate *FINRA* into dark pool trading and internalization by OTC market makers, the picture changes substantially. Order routing to dark pools is higher when the

[12] The results are robust to using lagged instruments directly in Equation (3), and estimating the panel regressions using OLS (see Tables SA3.1, SA3.2, and SA3.3).

TABLE 2 Order routing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	A. 2009				B. 2020				C. 2020 ex-COVID			
	TRF	*DP*	*INT*	*COMP*	*FINRA*	*ATS*	*Non-ATS*	*COMP*	*FINRA*	*ATS*	*Non-ATS*	*COMP*
log(Quoted spread)	0.0168***	−0.0065***	0.0208***	−0.0253***	0.0061	0.0243***	−0.0187**	−0.0038	0.0806**	0.0582***	0.0213	−0.0227
	(0.0041)	(0.0020)	(0.0037)	(0.0032)	(0.0103)	(0.0086)	(0.0083)	(0.0091)	(0.0372)	(0.0206)	(0.0216)	(0.0219)
log(Depth)	0.0125***	−0.0060***	0.0196***	−0.0066***	0.0050	−0.0221***	0.0276***	−0.0006	0.0085	−0.0165***	0.0265***	0.0072
	(0.0029)	(0.0012)	(0.0024)	(0.0017)	(0.0063)	(0.0040)	(0.0064)	(0.0061)	(0.0081)	(0.0056)	(0.0075)	(0.0072)
log(Price)	−0.0196***	−0.0169***	−0.0022	0.0025	0.0095	−0.0064	0.0166**	−0.0043	−0.0306**	−0.0276***	−0.0012	0.0253**
	(0.0044)	(0.0022)	(0.0042)	(0.0034)	(0.0101)	(0.0057)	(0.0075)	(0.0082)	(0.0140)	(0.0084)	(0.0083)	(0.0098)
log(Volatility)	−0.0222***	−0.0103***	−0.0099***	0.0073***	−0.0335***	−0.0288***	−0.0034	0.0122	−0.0094	−0.0146	0.0061	−0.0143
	(0.0013)	(0.0005)	(0.0011)	(0.0008)	(0.0087)	(0.0057)	(0.0071)	(0.0076)	(0.0125)	(0.0095)	(0.0075)	(0.0110)
log(Volume)	0.0384***	0.0123***	0.0241***	−0.0118***	0.0327***	0.0215***	0.0105**	−0.0129**	0.0491***	0.0290***	0.0190***	−0.0210**
	(0.0017)	(0.0005)	(0.0015)	(0.0008)	(0.0050)	(0.0033)	(0.0041)	(0.0053)	(0.0108)	(0.0073)	(0.0065)	(0.0078)
Observations	693,453	693,453	693,453	693,453	143,774	143,774	143,774	143,774	118,055	118,055	118,055	118,055
R-squared	0.0363	0.0206	0.0181	0.0060	0.0308	−0.0003	0.0028	0.0121	0.0072	−0.0557	0.0055	0.0245
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of regressions of *TRF*, *DP* and *INT*, and *COMP* (*FINRA*, *ATS*, *Non-ATS*, and *COMP*) on order book and firm characteristics. We estimate the following panel regression of market shares ($MS_{i,t}$) using IV/2SLS:

(1) $log(Quoted\ spread)_{i,t} = \alpha_i\, d_q + \beta_{1,1}\, log(Quoted\ spread)_{i,t-1} + \beta_{1,2}\, log(Depth)_{i,t-1} + \gamma_1\, X_{i,t} + e_{1,i,t},$
(2) $log(Depth)_{i,t} = \alpha_i\, d_q + \beta_{2,1}\, log(Quoted\ spread)_{i,t-1} + \beta_{2,2}\, log(Depth)_{i,t-1} + \gamma_2\, X_{i,t} + e_{2,i,t},$
(3) $MS_{i,t} = \alpha_i\, d_q + \beta_{3,1}\, log(Quoted\ spread)_{i,t} + \beta_{3,2}\, log(Depth)_{i,t} + \gamma_3\, X_{i,t} + e_{3,i,t},$

where $X_{i,t}$ is a vector of control variables that includes $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$. Equations (1) and (2) model the endogenous variables $log(Quoted\ spread)_{i,t}$ and $log(Depth)_{i,t}$ using their lagged values as instruments. Equation (3) uses the fitted values from the first-stage regressions for the endogenous variables.
$^{*}p < 0.1$ $^{**}p < 0.05$, $^{***}p < 0.01$.

book is illiquid and the opposite is true for routing to internalizing OTC market makers. Order routing to lit competing venues is decreasing in both quoted spread and depth in 2009, whereas in both the 2020 samples it is unrelated to the book characteristics. Furthermore, we find that volume is consistently positively related to dark trading and negatively related to order routing to lit competing venues, but the effects of volatility and price change substantially between 2009 and 2020.

To understand the mechanisms that explain the aggregate results, and why they differ between 2009 and 2020, we need to investigate order routing at the most granular level available. We therefore sort our roughly 3000 stocks on previous year-end market capitalization and divide them into terciles (Small, Medium, and Large). We also separately analyze the stocks that are part of the S&P 500 index.

Table 3 shows that in 2009 for Large caps—and more generally for the S&P 500 stocks—order routing to dark pools increases when the book is liquid as reflected in NBBO spread and depth. Jumping the long queues on the limit order book motivates traders to route to dark pools for Large caps as predicted by Buti et al. (2017). For Small caps, the opposite holds, and order routing to dark pools is strongly negatively related to NBBO depth. This suggests that for Small caps shallow books motivate traders to seek liquidity in dark pools.

For the 2020 ex-COVID sample, Large caps are unrelated to the state of the book, whereas the effect for Small caps gets stronger and order routing to dark pools increases not only when depth is shallow but also when the spread is large.[13] These results confirm that the desire to avoid crossing the spread is an important driver for order flow in Small caps. Note that once we add the 9 COVID weeks back into the sample, orders are routed to dark pools when the book is illiquid not only for Small caps, but also for Large caps.

The market conditions also significantly affect order routing. Order routing to dark pools is increasing in volume and this result holds across stocks terciles and sample periods, consistent with the idea that investors seek alternative liquidity venues when trading activity is high. The effect of volatility is more complex. Order routing to dark pools is lower when volatility is high in 2009 and for the entire 2020 sample, whereas it is unrelated with volatility for the 2020 ex-COVID sample. This suggests that, while execution uncertainty and adverse selection costs due to stale trading (Aquilina at al., 2021) were a concern for dark venues during periods of high volatility in 2009, and during the 2020 COVID crash, those concerns have attenuated in 2020 overall, probably as a result of improvements in dark venue execution speeds. As a result, volatility is less of a deterrent to use dark pools in 2020 than it was in 2009. Finally, order routing to dark pools decreases in price: hence it increases with relative tick size in 2009 and—especially for Small caps—in the more recent 2020 ex-COVID sample, while it disappears if we add the turbulent weeks of the COVID period in 2020. Two possible mechanisms may be at work. First, when the relative tick size is large crossing the spread to execute a market order is more expensive and therefore opting for dark pools executing inside the NBBO may be more desirable. Second, when the relative tick size is large the queues at the top of the limit order book become longer and jumping the queues by routing orders to dark pools becomes an attractive order submission strategy.[14] Table 3 shows that dark pool order routing is positively related with depth for Large caps, whereas it is negatively related for Small caps, and we therefore conjecture that it is primarily the second mechanism that is at work for Large caps.

Turning to internalization, Tables 2 and 3 show that order routing to OTC market makers is increasing in NBBO depth and this result holds across all terciles and sample periods. This means that the main effect at work over time and across different stocks is that OTC market makers are more likely to internalize orders when they can lay off the resulting inventory against a deep book. Results for spread are more complex. In 2009, the desire to avoid crossing wide spreads is an important driving force for routing trades to OTC market makers. However, this effect virtually disappears in the 2020 samples. When including the 9 weeks of extreme COVID-related volatility, results become noisy and insignificant: for example, it is hard to explain why order routing to internalizing OTC market makers decreases in

[13] As mentioned before (Table SA2.1), this result should be taken with caution as for Large caps the lagged NBBO spread value is a weak instrument for the state of the book in the 2020 ex-COVID sample.

[14] Yao and Ye (2018) show that an increase in the relative tick size increases rents for liquidity provision and lengthens the queue at the BBO.

TABLE 3 Order routing by size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	DP				*INT*				*COMP*			
A. 2009	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500
log(Quoted spread)	−0.0025	−0.0066**	−0.0165***	−0.0120***	0.0091*	0.0392***	0.0211***	0.0103	−0.0219***	−0.0341***	−0.0223***	−0.0182**
	(0.0026)	(0.0033)	(0.0030)	(0.0035)	(0.0051)	(0.0057)	(0.0050)	(0.0070)	(0.0048)	(0.0045)	(0.0051)	(0.0075)
log(Depth)	−0.0182***	−0.0118***	0.0055***	0.0064***	0.0246***	0.0155***	0.0270***	0.0253***	−0.0090***	0.0007	−0.0130***	−0.0150***
	(0.0020)	(0.0021)	(0.0012)	(0.0013)	(0.0043)	(0.0036)	(0.0029)	(0.0034)	(0.0033)	(0.0033)	(0.0021)	(0.0026)
log(Price)	−0.0180***	−0.0242***	−0.0098***	−0.0053	−0.0150***	0.0010	0.0316***	0.0214***	0.0132***	−0.0011	−0.0151***	−0.0122
	(0.0026)	(0.0035)	(0.0031)	(0.0039)	(0.0057)	(0.0056)	(0.0057)	(0.0076)	(0.0045)	(0.0046)	(0.0057)	(0.0087)
log(Volatility)	−0.0133***	−0.0137***	−0.0056***	−0.0058***	−0.0049**	−0.0162***	−0.0115***	−0.0091***	0.0083***	0.0100***	0.0052***	0.0046***
	(0.0010)	(0.0008)	(0.0006)	(0.0007)	(0.0021)	(0.0014)	(0.0012)	(0.0015)	(0.0016)	(0.0011)	(0.0009)	(0.0012)
log(Volume)	0.0114***	0.0168***	0.0138***	0.0144***	0.0101***	0.0384***	0.0373***	0.0363***	−0.0055***	−0.0198***	−0.0166***	−0.0184***
	(0.0007)	(0.0007)	(0.0007)	(0.0009)	(0.0022)	(0.0018)	(0.0017)	(0.0023)	(0.0012)	(0.0011)	(0.0012)	(0.0016)
Observations	199,964	244,210	249,279	106,785	199,964	244,210	249,279	106,785	199,964	244,210	249,279	106,785
R-squared	0.0154	0.0291	0.0302	0.0360	0.0057	0.0417	0.0502	0.0587	0.0026	0.0142	0.0169	0.0258
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(Continues)

TABLE 3 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	ATS				*Non-ATS*				*COMP*			
B. 2020	**Small**	**Medium**	**Large**	**S&P 500**	**Small**	**Medium**	**Large**	**S&P 500**	**Small**	**Medium**	**Large**	**S&P 500**
log(Quoted spread)	0.0134**	0.0193	0.0273***	0.0238***	−0.0452***	−0.0020	0.0023	0.0139	0.0076	−0.0040	−0.0173	−0.0339*
	(0.0053)	(0.0142)	(0.0087)	(0.0086)	(0.0144)	(0.0148)	(0.0116)	(0.0197)	(0.0097)	(0.0100)	(0.0117)	(0.0174)
log(Depth)	−0.0220***	−0.0290***	−0.0142**	−0.0078	0.0098	0.0423***	0.0220***	0.0319***	−0.0013	−0.0009	0.0146	−0.0020
	(0.0050)	(0.0067)	(0.0063)	(0.0072)	(0.0101)	(0.0110)	(0.0082)	(0.0096)	(0.0088)	(0.0109)	(0.0092)	(0.0094)
log(Price)	−0.0083	−0.0078	−0.0057	−0.0040	−0.0119	0.0368***	0.0312***	0.0302	0.0131	−0.0117	−0.0256**	−0.0317*
	(0.0057)	(0.0093)	(0.0085)	(0.0090)	(0.0126)	(0.0102)	(0.0098)	(0.0182)	(0.0125)	(0.0091)	(0.0115)	(0.0174)
log(Volatility)	−0.0219***	−0.0302***	−0.0319***	−0.0309***	0.0235***	−0.0101	−0.0315***	−0.0413**	0.0004	0.0130	0.0389***	0.0515***
	(0.0037)	(0.0085)	(0.0087)	(0.0102)	(0.0085)	(0.0108)	(0.0112)	(0.0198)	(0.0074)	(0.0094)	(0.0126)	(0.0180)
log(Volume)	0.0135***	0.0264***	0.0315***	0.0310***	0.0044	0.0205***	0.0128*	0.0055	0.0010	−0.0289***	−0.0371***	−0.0357***
	(0.0023)	(0.0058)	(0.0066)	(0.0079)	(0.0048)	(0.0061)	(0.0072)	(0.0124)	(0.0048)	(0.0078)	(0.0099)	(0.0132)
Observations	38,315	50,243	54,905	21,143	38,315	50,243	54,905	21,143	38,315	50,243	54,905	21,143
R-squared	0.0016	0.0062	0.0124	0.0059	−0.0010	0.0087	0.0473	0.0806	−0.0045	0.0515	0.0656	0.0363
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(Continues)

TABLE 3 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	ATS				*Non-ATS*				*COMP*			
C. 2020 ex-COVID	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500
log(Quoted spread)	0.0198**	0.0699**	0.0847*	0.0853	0.0015	0.0326*	0.0278	0.0139	0.0022	−0.0105	−0.0569	−0.1617
	(0.0077)	(0.0317)	(0.0502)	(0.0893)	(0.0194)	(0.0192)	(0.0504)	(0.0197)	(0.0136)	(0.0198)	(0.0616)	(0.2729)
log(Depth)	−0.0211***	−0.0157*	−0.0113	−0.0245*	0.0150	0.0414***	0.0200	0.0319***	0.0057	0.0094	0.0230	0.0107
	(0.0066)	(0.0086)	(0.0094)	(0.0122)	(0.0132)	(0.0126)	(0.0135)	(0.0096)	(0.0093)	(0.0119)	(0.0150)	(0.0209)
log(Price)	−0.0205***	−0.0325**	−0.0278	−0.0310	0.0223**	0.0032	−0.0360*	−0.0413**	0.0108	0.0272**	0.0501**	0.0886
	(0.0046)	(0.0132)	(0.0191)	(0.0338)	(0.0091)	(0.0081)	(0.0191)	(0.0198)	(0.0077)	(0.0107)	(0.0240)	(0.1009)
log(Volatility)	−0.0103	−0.0160	−0.0236	−0.0275	−0.0297**	0.0126	0.0350***	0.0302	−0.0018	−0.0127	−0.0373**	−0.0553
	(0.0078)	(0.0152)	(0.0177)	(0.0186)	(0.0136)	(0.0110)	(0.0106)	(0.0182)	(0.0134)	(0.0130)	(0.0176)	(0.0371)
log(Volume)	0.0143***	0.0386***	0.0502**	0.0607	0.0119**	0.0274***	0.0241	0.0055	−0.0012	−0.0372***	−0.0575**	−0.0963
	(0.0030)	(0.0131)	(0.0221)	(0.0453)	(0.0053)	(0.0070)	(0.0182)	(0.0124)	(0.0056)	(0.0097)	(0.0234)	(0.1184)
Observations	31,402	41,213	45,173	17,406	31,402	41,213	45,173	21,143	31,402	41,213	45,173	17,406
R-squared	−0.0015	−0.0787	−0.1637	−0.2844	0.0125	0.0145	0.0120	0.0806	0.0058	0.0750	0.0371	−0.8746
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of regressions of *DP* and *INT*, and *COMP* (*ATS*, *Non-ATS*, and *COMP*) on lagged instruments for subsamples by market capitalization, and for stocks that are part of the S&P 500 index. We estimate the following panel regression of market shares ($MS_{i,t}$) using IV/2SLS:

(1) $log(Quoted\ spread)_{i,t} = \alpha_i\, d_q + \beta_{1,1}\, log(Quoted\ spread)_{i,t-1} + \beta_{1,2}\, log(Depth)_{i,t-1} + \gamma_1\, X_{i,t} + e_{1,i,t}$,
(2) $log(Depth)_{i,t} = \alpha_i\, d_q + \beta_{2,1}\, log(Quoted\ spread)_{i,t-1} + \beta_{2,2}\, log(Depth)_{i,t-1} + \gamma_2\, X_{i,t} + e_{2,i,t}$,
(3) $MS_{i,t} = \alpha_i\, d_q + \beta_{3,1}\, log(Quoted\ spread)_{i,t} + \beta_{3,2}\, log(Depth)_{i,t} + \gamma_3\, X_{i,t} + e_{3,i,t}$,

where $X_{i,t}$ is a vector of control variables that includes $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$. Equations (1) and (2) model the endogenous variables $log(Quoted\ spread)_{i,t}$ and $log(Depth)_{i,t}$ using their lagged values as instruments. Equation (3) uses the fitted values from the first-stage regressions (1) and (2).
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

spread only for Small caps in the 2020 overall sample.[15] In 2009, market conditions also matter for order routing to OTC market makers and they receive more order flow when volatility is low and trading activity is high—as was the case for dark pools, but the results are noisier in the more recent sample.

Finally, order flow to lit competing venues is generally decreasing in both the quoted spread and depth in 2009, which is consistent with the idea that bypassing time-priority is more valuable when the listing exchange's order book is more competitive (Foucault & Menkveld, 2008). In contrast to dark fragmentation, order flow to lit competing venues is decreasing in volume and increasing in volatility, while it shows no significant correlation to the relative tick size. Thus, it appears that informed traders benefit from being able to sweep the possibly stale limit order books across venues when markets are more volatile and trading activity is low (Chakravarty et al., 2012). Order book characteristics are generally unrelated to order flow to competing venues in 2020.

Taken together, our results on order routing are as follows. Consistently across our sample periods, for Small caps order routing to dark pools increases when the book is illiquid, the relative tick size is large, there is less uncertainty, and high trading activity signals to traders that they may find liquidity in dark pools. In 2009, dark pools are mainly used for Large caps to jump the long queues in deep books, especially when the relative tick size is high. By contrast, in 2020, order routing to dark pools for Large caps is unrelated to the state of the book, and traders instead use dark pools to avoid crossing the wide spread during the COVID period.[16] Similar results as for Large caps hold overall for the S&P 500 sample. Consistently across stocks and samples, we find that OTC market makers internalize more orders when depth is high, and they can more easily lay off the resulting inventory in lit markets. Results for lit competing venues are intriguing for 2009 as investors resort to lit competing venues when the book spread is narrow, possibly undercutting time priority, but no significant results persist in 2020. We discuss possible reasons for the changing pattern of fragmentation between our sample periods in Section 6.

4 | DARK TRADING AND MARKET QUALITY

A central question for regulators is whether dark trading has any detrimental effects on measures of market quality, such as quoted spreads, effective spreads, and short-term volatility. Traders decide whether to submit an order to a dark venue or to the public limit order book based on observing the depth and the quoted spread as well as their information about the value of the stock. Therefore, to investigate the effects of dark trading on market quality we have to take into account the fact that dark trading is endogenous. We estimate two-stage least squares instrumental variables (IV/2SLS) panel regressions, where we instrument for dark trading as well as competition from lit venues in an attempt to control for endogeneity following Hasbrouck and Saar (2013). We also include an instrument for market-wide market quality to control for reverse causality following, for example, Degryse et al. (2015) and Comerton-Forde and Putniņš (2015).

As the power of the IV/2SLS method depends on the quality of the instruments, we need to find good instruments for dark market shares. Hasbrouck and Saar (2013) propose using the average low latency trading in other stocks during the same time period as an instrument for low latency trading in a particular stock when evaluating the impact of low latency trading on market quality. We follow their suggestion and use average across stocks of dark market share on day t as an instrument for the dark market share in stock i. The idea is that, if dark trading has a significant market-wide component, a measure of market-wide average of dark market share will correlate with firm-level dark trading. However, to be a good instrument, we also need to ensure that our market-wide average is uncorrelated with the error

[15] Notice that for Small caps, OTC market makers internalize more orders when volatility as measured by the intraday range is high, while the sign on volatility is the opposite for all other subsamples. Therefore, we conjecture that the Small cap result on spread arises because wide quoted spreads coincide with high volatility (Table SA2.1), and this causes the IV/2SLS to load any quoted spread effect on the volatility control for Small caps. Evidence supporting this conjecture is in Table SA3.3 where we instead use OLS. Narrow lagged quoted spreads still lead to more internalizing by OTC market makers, but the magnitude is much smaller and there is no significant effect of the volatility control variable.

[16] In Section 5, we further discuss order routing in periods of market stress.

term in Equation (6) defined below. Excluding the firm itself from the market-wide average eliminates a clear source of correlation, and by also excluding firms in the same industry and firms that belong to the same index we reduce the correlation that may arise because of common industry- or index-based trading strategies. Therefore, we exclude stock *i* and require the other stocks (*Noti*) to have a market capitalization in the same size tercile (Large, Medium, Small) as stock *i*.[17] Furthermore, following Hasbrouck and Saar (2013), we exclude stocks that are in the same four-digit SIC code or in the same major index (S&P 500, Nasdaq 100) as firm *i* from *Noti*. We use the same method to create instruments for each of our market shares. Finally, we also create instruments for market quality measures using the same approach.[18] Specifically, we define $Y_{Noti,t}$ as the day *t* average market quality measure $Y_{i,t}$ across stocks in the same size tercile excluding stock *i*, but that are not in same four-digit SIC code, and not in same index (S&P 500 and/or Nasdaq 100).

The IV/2SLS panel regressions take the following form:

$$DARK_{i,t} = a_i d_q + b_1 W_{i,t} + c_1 Z_{i,t} + e_{1,i,t}, \tag{4}$$

$$COMP_{i,t} = a_i d_q + b_2 W_{i,t} + c_2 Z_{i,t} + e_{2,i,t}, \tag{5}$$

$$Y_{i,t} = \alpha_i d_q + \beta_1 DARK_{i,t} + \beta_2 COMP_{i,t} + ,Z_{i,t} + e_{4,t}, \tag{6}$$

where $Y_{i,t}$ is a market quality measure, and $DARK_{i,t}$, is the market share of *TRF* trading for 2009 and the market share of *FINRA*-reported trading for 2020, and $COMP_{i,t}$ is the market share of lit competing venues expressed as a percentage of consolidated volume. $Z_{i,t}$ is a vector of control variables that includes $Y_{Noti,t}$, *log(Price)*$_{i,t}$, *log(Volatility)*$_{i,t}$, and *log(Volume)*$_{i,t}$. $W_{i,t}$ is a vector that includes $DARK_{Noti,t}$ and $COMP_{Noti,t}$, where *Noti,t* again stands for the day *t* average across stocks in the same size group as stock *i*, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 or Nasdaq 100).

We report results from the second stage of (6) in Table 4 for each of our market quality measures: *log(Quoted spread)*, *log(Effective half-spread)*, and *log(StD returns)*. The first-stage regressions (Table SA4.1) for 2009 show that the lagged instruments for own-effects are positive and significant. For example, when $Y_{i,t}$ is the quoted spread, $b_{1,1} = 0.8981^{***}$ and $b_{2,2} = 0.9449^{***}$. The cross-effects are smaller ($b_{1,2} = -0.0320$ and $b_{2,1} = 0.0245^{**}$), and only the second one is significant. For the 2020 overall and ex-COVID samples, both the own-effect and cross-effect for $DARK_{Noti,t}$ are insignificant, but the coefficient on $COMP_{Noti,t}$ is significant both as an own-effect ($b_{2,2} = 0.9693^{***}$ and $b_{2,2} = 0.9765^{***}$, respectively) and as a cross-effect ($b_{2,1} = 0.8111^{***}$ and $b_{2,1} = 0.7295^{***}$).

Starting with 2009, results in Panel A of Table 4 show that dark trading leads to lower quoted and effective half-spreads but does not affect short-term volatility significantly. The effect of lit competition on quoted spreads and short-term volatility is insignificant. The controls are all significant: as expected, spreads and short-term volatility are increasing in relative tick size (decreasing in price), and increasing in volatility as measured by the intraday range. Quoted spreads are decreasing in volume as expected, but both effective half-spreads and short-term volatility are increasing in volume.

Turning now to the 2020 samples, results in Panel B and C show that dark trading does not significantly affect any market quality measure. By contrast, lit competition is associated with lower quoted and effective half-spreads but has no significant effect on short-term volatility. With the exception of price for the quoted spread regressions, the control

[17] The idea behind the size grouping is that we have observed that there are systematic differences in dark pool trading across subsamples.

[18] Hasbrouck and Saar (2013) were able to use the spreads for other markets quoting the same security in their analysis of low latency orders on the Nasdaq. We unfortunately cannot follow their strategy because dark pools do not disseminate quotes.

TABLE 4 Dark and lit market share and market quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. 2009			B. 2020			C. 2020 ex-COVID		
	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*
DARK	−0.8245***	−0.6014***	−0.0075	1.6147	1.9497*	0.3808	0.9722	1.0382	0.6813
	(0.1428)	(0.1346)	(0.1546)	(1.3370)	(1.1634)	(0.9149)	(1.3555)	(0.6678)	(0.6328)
COMP	−0.2933*	−0.6234***	0.1134	−2.5903**	−2.8322**	−0.9996	−2.0078*	−1.9342***	−1.0432*
	(0.1739)	(0.2136)	(0.1759)	(1.2042)	(1.0602)	(0.9030)	(1.1192)	(0.6563)	(0.5703)
log(Price)	−0.3915***	−0.4740***	−0.0407***	0.0270	−0.0837*	−0.1726***	0.0486	−0.1023**	−0.0879**
	(0.0123)	(0.0142)	(0.0130)	(0.0366)	(0.0481)	(0.0415)	(0.0416)	(0.0428)	(0.0404)
log(Volatility)	0.1933***	0.1926***	0.5130***	0.4247***	0.4279***	0.3509***	0.3374***	0.3687***	0.3037***
	(0.0044)	(0.0041)	(0.0046)	(0.0393)	(0.0209)	(0.0253)	(0.0252)	(0.0163)	(0.0162)
log(Volume)	−0.1069***	0.0140***	0.1002***	−0.3015***	−0.2737***	−0.1460***	−0.2656***	−0.2268***	−0.1508***
	(0.0059)	(0.0053)	(0.0061)	(0.0528)	(0.0456)	(0.0364)	(0.0543)	(0.0266)	(0.0268)
Y_{Noti}	0.4819***	0.3480***	0.4520***	0.9770***	0.7997***	0.8216***	0.9808***	0.6566***	0.7398***
	(0.0237)	(0.0292)	(0.0197)	(0.0716)	(0.1071)	(0.0644)	(0.0830)	(0.0666)	(0.0451)
Observations	696,594	695,821	696,653	146,667	146,665	146,647	120,948	120,947	120,934
R-squared	0.1733	0.0702	0.4539	0.4392	0.1519	0.5493	0.2187	−0.0079	0.1408
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of analyzing the relationship between market shares and market quality. We estimate the following three-equation panel regression model market quality measures ($Y_{i,t}$) using IV/2SLS:

(1) $DARK_{i,t} = a_i\, d_q + b_1\, W_{i,t} + c_1\, X_{i,t} + e_{1,i,t}$,
(2) $COMP_{i,t} = a_i\, d_q + b_2\, W_{i,t} + c_2\, X_{i,t} + e_{2,i,t}$,
(3) $Y_{i,t} = \alpha_i\, d_q + \beta_1\, DARK_{i,t} + \beta_2\, COMP_{i,t} + \gamma\, X_{i,t} + e_{3,i,t}$,

where $X_{i,t}$ is a vector of control variables that includes *log(Price)*$_{i,t}$, *log(Volatility)*$_{i,t}$, *log(Volume)*$_{i,t}$, and $Y_{Noti,t}$. $W_{i,t}$ is a vector that includes $DARK_{Noti,t}$ and $COMP_{Noti,t}$, where *Noti,t* stands for the day *t* average across stocks in the same size group, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 and/or Nasdaq 100). For 2009, *DARK* = *TRF* and for 2020, *DARK* = *FINRA*. The first-stage regressions based on Equations (1) and (2) are reported in Table SA4.1. The second-stage IV/2SLS regressions in Equation (3) use the fitted value from the first-stage regressions (1) and (2).
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

variables are significant also for 2020. As expected, spreads and short-term volatility are increasing in volatility. While quoted and effective half-spreads are decreasing in volume as expected, short-term volatility is increasing in volume.

Even if overall dark trading does not have detrimental effects on market quality, it is possible that one of the dark trading types—dark pools or internalization by OTC market makers—could adversely affect market quality. Dark pools specialize in Large caps, while OTC market makers are more active in Small caps (Figure 2), and this may lead to differences in the cross section. It is also possible that dark trading is particularly detrimental for Small caps, where fragmentation and information asymmetries are more likely to affect market quality. We therefore further explore whether the effect of dark trading on market quality varies by the type of venue (dark pool or internalized by OTC market makers) or by size.

When we decompose *DARK* into *DP* and *INT* for the 2009 sample, the IV/2SLS panel regressions take the following form:

$$DP_{i,t} = a_i d_q + b_1 W_{i,t} + c_1 Z_{i,t} + e_{1,i,t}, \tag{7}$$

$$INT_{i,t} = a_i d_q + b_2 W_{i,t} + c_2 Z_{i,t} + e_{2,i,t}, \tag{8}$$

$$COMP_{i,t} = a_i d_q + b_3 W_{i,t} + c_3 Z_{i,t} + e_{3,i,t}, \tag{9}$$

$$Y_{i,t} = \alpha_i d_q + \beta_1 DP_{ii,t} + \beta_2 INT_{ii,t} + \beta_3 COMP_{i,t} + ,Z_{i,t} + e_{4,t}, \tag{10}$$

where $Z_{i,t}$ is a vector of control variables that includes $Y_{Noti,t}$, $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$. $W_{i,t}$ is a vector that includes $DP_{Noti,t}$, $INT_{Noti,t}$, and $COMP_{Noti,t}$, and *Noti,t* again stands for the day *t* average across stocks in the same size group as stock *i*, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 or Nasdaq 100). For the 2020 samples, we replace *DP* with market share of *ATSs* and *INT* with market share of *Non-ATSs*.

The results of the second-stage estimation are in Panel A of Table 5 for the 2009 sample and in Panels B and C for the 2020 samples.[19] The first-stage results are in Table SA5.1 for quoted spread and the results are similar for the other market quality variables. They show that the instruments for the own-effects are positive and highly significant in 2009 (coefficient on $DP_{Noti,t}$ is 0.8855***, $INT_{Noti,t}$ is 0.9253***, and $COMP_{Noti,t}$ is 0.9428***). In the 2020 samples, only the instruments for the own-effects in the *ATS* and *COMP* equations are significant but the cross-effects are significant for all equations, including the *Non-ATS* equation (for the 2020 overall sample coefficient on $ATS_{Noti,t}$ is 0.9624***, $Non\text{-}ATS_{Noti,t}$ is 0.0092, and $COMP_{Noti,t}$ is 0.9652***; while for the 2020 ex-COVID sample coefficients are 0.8999***, 0.0262, and 0.9669***, respectively).

For 2009, a higher dark pool market share leads to narrower quoted and effective half-spreads. A higher internalized market share also leads to narrower quoted spreads, and there is a negative coefficient in the effective half-spread regression but it is not significant at conventional levels. Neither form of dark trading affects short-term volatility. Hence, both types of dark trading lead to improved market quality on average in 2009. By contrast, we find no evidence that either form of dark trading affects spreads in the overall 2020 sample, and a higher dark pool market share leads to higher short-term volatility in this sample. In addition, we find that dark pool trading leads to wider effective half-spread for the 2020 ex-COVID sample. The coefficients on the control variables are significant and have the same signs as they did in Table 4.

[19] Note that the results for lit competition do not change qualitatively when we split dark trading into dark pools and internalization by OTC market makers. Therefore, we do not discuss these results again.

TABLE 5 Dark pool/ATS, internalization/Non-ATS, lit market shares, and market quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. 2009			B. 2020			C. 2020 ex-COVID		
	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*
DP/ATS	−1.3099***	−1.1811***	−0.1307	2.5974	1.6542	3.4465**	4.4923	3.7893**	3.3695**
	(0.1951)	(0.2376)	(0.2156)	(2.1593)	(2.0392)	(1.5449)	(3.3272)	(1.7666)	(1.5351)
INT/Non-ATS	−0.5894***	−0.3235*	0.0525	0.3338	2.2154	−1.8781	−2.4859	−0.8410	−1.1503
	(0.1514)	(0.1771)	(0.1543)	(4.9263)	(3.8585)	(2.1085)	(7.4185)	(2.0022)	(2.1784)
COMP	−0.2393	−0.5594**	0.1231	−2.5175*	−2.7948***	−1.3087*	−2.3714	−2.4443***	−1.4832**
	(0.1641)	(0.2261)	(0.1758)	(1.3004)	(0.8478)	(0.7211)	(1.5269)	(0.7399)	(0.6111)
log(Price)	−0.3896***	−0.4718***	−0.0404***	0.0233	−0.0786	−0.2014***	0.0582	−0.1324**	−0.0964*
	(0.0122)	(0.0140)	(0.0129)	(0.0483)	(0.1020)	(0.0618)	(0.0583)	(0.0527)	(0.0497)
log(Volatility)	0.1911***	0.1895***	0.5121***	0.4202***	0.4225***	0.3714***	0.3601***	0.3980***	0.3267***
	(0.0043)	(0.0041)	(0.0046)	(0.0444)	(0.0285)	(0.0223)	(0.0298)	(0.0300)	(0.0227)
log(Volume)	−0.1070***	0.0144***	0.1004***	−0.2882***	−0.2736***	−0.1465***	−0.2506**	−0.2338***	−0.1583***
	(0.0053)	(0.0052)	(0.0059)	(0.0819)	(0.0566)	(0.0333)	(0.1145)	(0.0306)	(0.0327)
Y_{Noti}	0.4878***	0.3568***	0.4532***	0.9248***	0.8242**	0.6626***	0.9502***	0.4661**	0.6533***
	(0.0229)	(0.0290)	(0.0195)	(0.2098)	(0.3379)	(0.1516)	(0.1436)	(0.1879)	(0.1039)

(Continues)

TABLE 5 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. 2009			B. 2020			C. 2020 ex-COVID		
	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*	$Y_{i,t}$ = *log(Quoted spread)*	$Y_{i,t}$ = *log(Effective half-spread)*	$Y_{i,t}$ = *log(StD returns)*
Observations	696,594	695,821	696,653	146,667	146,665	146,647	120,948	120,947	120,934
R-squared	0.1882	0.0753	0.4540	0.5158	0.1230	0.2033	−0.7547	−0.5558	−0.2649
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.

The table reports the results of analyzing the relationship between market shares and market quality for all stocks. We estimate the following four-equation panel regression model for market quality measures ($Y_{i,t}$) using IV/2SLS:

(1) $DP_{i,t} = a_i\, d_q + b_1\, W_{i,t} + c_1\, X_{i,t} + e_{1,i,t}$,
(2) $INT_{i,t} = a_i\, d_q + b_2\, W_{i,t} + c_2\, X_{i,t} + e_{2,i,t}$,
(3) $COMP_{i,t} = a_i\, d_q + b_3\, W_{i,t} + c_3\, X_{i,t} + e_{3,i,t}$,
(4) $Y_{i,t} = \alpha_i\, d_q + \beta_1\, DP_{i,t} + \beta_2\, INT_{i,t} + \beta_3\, COMP_{i,t} + \gamma\, X_{i,t} + e_{4,i,t}$,

where $X_{i,t}$ is a vector of control variables that includes $Y_{Noti,t}$, $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$. $W_{i,t}$ is a vector that includes $DP_{Noti,t}$, $INT_{Noti,t}$, and $COMP_{Noti,t}$, where *Noti,t* stands for the day *t* average across stocks in the same size group, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 and/or Nasdaq 100). For 2020, *DP* = *ATS* and *INT* = *Non-ATS*. The first-stage regressions (1)–(3) are reported in Table SA5.1. The second-stage regressions (4) use the fitted value from the first-stage regressions (1)–(3) of the endogenous variables.
$^{*}p < 0.1$. $^{**}p < 0.05$, $^{***}p < 0.01$.

We examine whether the beneficial effects of both types of dark trading on market quality in 2009, and the negative effect on effective half-spread in the 2020 ex-COVID sample and on short-term volatility in both the 2020 samples, derive from certain groups of stocks by conducting splits by size. The results are in Table 6.[20] Panel A covers the results for 2009. The beneficial effects of dark pools for quoted spreads are evident for all subsamples. Internalization by OTC market makers is beneficial for quoted spreads for all but the subsample of Small caps where the coefficient is insignificant. Similarly, the beneficial effects of dark pools for effective half-spreads are evident for all but the S&P 500 subsample where the coefficient is insignificant. Higher internalization by OTC market makers leads to lower effective half-spreads for Large caps and S&P 500 stocks, but the coefficient is insignificant for the remaining subsamples. Finally, there is no significant effect of either higher dark pool trading or internalization on short-term volatility for any subsample. The results for the 2020 overall sample are in Panel B, and show that there are no significant effects on spreads of either higher dark pool trading or internalization by OTC market makers for any subsample. Note, however, that for the 2020 ex-COVID sample, the evidence shows that more dark pool trading leads to wider spreads for Large caps and to significantly higher short-term volatility both for Large caps and S&P 500 stocks.[21]

In sum, we find that both types of dark trading lead to lower spreads in 2009, but that dark pool trading leads to wider spreads for Large caps in the 2020 ex-COVID sample. We find no evidence that higher dark trading of either type affects short-term volatility in 2009, but we do find that higher dark pool market share leads to higher short-term volatility for Large caps in 2020. To further investigate why our results change over the two sample periods, in Section 5 we show the effects of dark fragmentation during periods of market stress and in Section 6 we provide a comprehensive discussion of our results.

5 | MARKET STRESS

Our samples cover two tumultuous periods in global financial markets, the Great Financial Crisis in 2009 and the start of the COVID pandemic in 2020. Did the dramatic market moves and the high levels of uncertainty during the first halves of 2009 and 2020 imply a different relationship between dark trading and market quality? More generally, does dark trading have detrimental effects on market quality in periods of market stress?

Figure 1 suggests that the first half of each year (*H1*) was a period of market stress. We consider several other indicators of market stress at the stock level, including the lowest tercile of individual stock returns (*ret_low*), the lowest tercile of stock-specific buy-order imbalances (*bs_low*), and the highest decile of volatility (*vol_extr*). We instead use the highest tercile of volatility (*vol_high*) for the 2020 sample due to the limited time-series. Since our indicator variables are another way of capturing days (weeks) with extreme volatility, we drop the intraday range from these regressions. Moreover, since these days (weeks) are likely associated with extreme prices, we also drop the stock price as a control variable. We continue controlling for volume and for stock fixed effects.

Table 7 reports the results of regressions of *DP* and *INT* on indicators of market stress for 2009 in Panel A, and of *ATS* and *Non-ATS* on indicators of market stress for 2020 in Panel B. We estimate the following panel regression of market shares ($MS_{i,t}$) using OLS:

$$MS_{i,t} = \alpha_i d_q + \beta I^{Stress} + ,log(Volume)_{i,t} + e_{i,t}, \tag{11}$$

where I^{Stress} is a stock-specific indicator for market stress as discussed above. Full results are in Table SA7, while Table 7 focuses on the coefficient β on our market stress indicators. Both dark pool trading and internalization are

[20] The first stage results are in Table SA6.1 for quoted spread. Results are similar for the other market quality variables.

[21] Degryse et al. (2015) argue that it is important to allow for nonlinear effects of fragmentation on measures of market quality. We replicate the analysis in Tables 4 and 5 allowing the effect of dark trading to affect market quality in a nonlinear way (by including a squared term). Overall, we do not find significant effects except for dark trading on the short-term volatility for 2009 (Tables SA8.1 and SA8.2).

TABLE 6 Dark pool/ATS, internalization/Non-ATS, lit market shares, and market quality by size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = log(Quoted spread)				$Y_{i,t}$ = log(Effective half-spread)				$Y_{i,t}$ = log(StD returns)			
A. 2009	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500
DP	−1.0866***	−1.1747***	−1.0780***	−0.9402**	−1.3539***	−1.1544***	−0.7792**	−0.9596	0.1764	0.1276	−0.7749*	−1.2325*
	(0.2591)	(0.2372)	(0.3479)	(0.4214)	(0.3490)	(0.2394)	(0.3784)	(0.6567)	(0.1966)	(0.2286)	(0.4294)	(0.6511)
INT	−0.2477	−0.7548***	−0.9697***	−1.0480***	−0.1443	−0.3940*	−0.8809***	−0.9220**	0.1922	0.1431	0.0166	−0.2456
	(0.1782)	(0.1847)	(0.2159)	(0.2881)	(0.2476)	(0.2263)	(0.2979)	(0.4396)	(0.1556)	(0.1992)	(0.2821)	(0.3582)
COMP	−0.2072	−0.3288	−0.4501*	−0.5329	−0.4029*	−0.6680**	−1.1655***	−1.6678**	−0.2742*	0.1439	0.8288**	0.8869**
	(0.1899)	(0.2540)	(0.2659)	(0.3317)	(0.2389)	(0.3073)	(0.4445)	(0.7487)	(0.1459)	(0.2586)	(0.3444)	(0.4266)
log(Price)	−0.3608***	−0.3821***	−0.4344***	−0.5392***	−0.4169***	−0.5117***	−0.4992***	−0.5366***	−0.0226*	−0.0267*	−0.0643***	−0.0846***
	(0.0162)	(0.0190)	(0.0241)	(0.0340)	(0.0195)	(0.0161)	(0.0227)	(0.0314)	(0.0118)	(0.0153)	(0.0216)	(0.0300)
log(Volatility)	0.3153***	0.1516***	0.0831***	0.0531***	0.2888***	0.1594***	0.0867***	0.0528***	0.5985***	0.4917***	0.4318***	0.4009***
	(0.0066)	(0.0063)	(0.0044)	(0.0057)	(0.0074)	(0.0047)	(0.0053)	(0.0087)	(0.0055)	(0.0061)	(0.0073)	(0.0100)
log(Volume)	−0.1272***	−0.0980***	−0.0702***	−0.0469***	−0.0095	0.0158**	0.0840***	0.1120***	0.0931***	0.0882***	0.1366***	0.1763***
	(0.0060)	(0.0078)	(0.0077)	(0.0100)	(0.0065)	(0.0071)	(0.0108)	(0.0170)	(0.0040)	(0.0092)	(0.0128)	(0.0168)
Y_{Noti}	0.3909***	0.5494***	0.5754***	0.4323***	0.2947***	0.3141***	0.4496***	0.5157***	0.3929***	0.5001***	0.4600***	0.4468***
	(0.0291)	(0.0320)	(0.0351)	(0.0430)	(0.0374)	(0.0335)	(0.0489)	(0.0735)	(0.0201)	(0.0241)	(0.0253)	(0.0296)
Observations	201,001	245,243	250,350	107,237	200,614	245,030	250,177	107,168	201,002	245,250	250,401	107,265
R-squared	0.2595	0.1751	0.2039	0.2620	0.0548	0.1098	0.1224	0.1506	0.4086	0.4565	0.4918	0.5068
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(Continues)

TABLE 6 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = *log(Quoted spread)*				$Y_{i,t}$ = *log(Effective half-spread)*				$Y_{i,t}$ = *log(StD returns)*			
B. 2020	**Small**	**Medium**	**Large**	**S&P 500**	**Small**	**Medium**	**Large**	**S&P 500**	**Small**	**Medium**	**Large**	**S&P 500**
ATS	−1.2035	2.5242	4.5543*	14.0634	−2.3550	−1.8407	3.1264*	7.9742	−0.0107	4.5230	3.5207**	5.7772
	(6.8377)	(7.8653)	(2.5428)	(28.2960)	(10.5478)	(15.6181)	(1.7185)	(9.3659)	(4.2221)	(4.1793)	(1.6390)	(5.1784)
Non-ATS	2.9283	0.4072	−6.2617	−26.1819	4.9635	11.6387	−2.4194	−10.5860	2.0785	−4.7182	−4.8788	−7.9586
	(8.9757)	(30.7526)	(8.8547)	(57.6560)	(9.4925)	(40.6304)	(4.7445)	(14.5756)	(2.6624)	(10.1632)	(4.3774)	(7.6472)
COMP	−3.2092	−2.5054	−3.1643**	−5.8480	−5.2094	−4.2456	−2.6305***	−4.3078	−2.5924	−0.8540	−1.6501	−2.4466
	(7.3808)	(7.2960)	(1.4958)	(8.9162)	(6.7964)	(6.7920)	(0.8941)	(3.6900)	(2.1968)	(2.5247)	(1.0819)	(2.3172)
log(Price)	0.0204	0.0243	0.1952	0.3326	0.0844	0.0755	0.0021	−0.1068	−0.1389	−0.1364	−0.0822	−0.1805
	(0.3477)	(0.2802)	(0.2047)	(0.6526)	(0.5474)	(0.7746)	(0.0534)	(0.2120)	(0.1319)	(0.1123)	(0.1033)	(0.1610)
log(Volatility)	0.4023***	0.3584*	0.2450	−0.3384	0.3430	0.2849	0.3556***	0.2312	0.2550***	0.3525***	0.3544***	0.3188*
	(0.1404)	(0.2099)	(0.2478)	(1.5659)	(0.2567)	(0.3603)	(0.0852)	(0.2475)	(0.0749)	(0.0368)	(0.1044)	(0.1592)
log(Volume)	−0.2544***	−0.3421	−0.2878**	−0.4497	−0.2480***	−0.5738	−0.2280***	−0.2874	−0.1458***	−0.0888	−0.1639**	−0.2190
	(0.0894)	(0.8950)	(0.1217)	(0.4773)	(0.0697)	(1.0425)	(0.0623)	(0.1896)	(0.0198)	(0.2730)	(0.0660)	(0.1404)
Y_{Noti}	0.7067	1.0126	1.0947***	1.0110	0.9493	1.6853	0.7649***	0.3670	0.8057***	0.6124	0.6403***	0.5143
	(0.5093)	(0.9821)	(0.2014)	(0.9660)	(0.9270)	(3.5374)	(0.2523)	(0.7212)	(0.2066)	(0.6079)	(0.1851)	(0.3487)
Observations	39,099	51,279	55,989	21,560	39,097	51,279	55,989	21,560	39,085	51,272	55,989	21,560
R-squared	−0.9116	0.5206	−0.2372	−10.0671	−3.3721	−7.5268	0.1859	−2.2935	−0.3291	−0.9594	−0.0481	−0.7139
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(Continues)

TABLE 6 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = *log(Quoted spread)*				$Y_{i,t}$ = *log(Effective half-spread)*				$Y_{i,t}$ = *log(StD returns)*			
C. 2020 ex-COVID	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500
ATS	2.2321	5.1721	5.6124**	13.3753	3.0410	4.0300	4.0840**	7.9931	0.4224	4.8592	3.2276***	3.6788**
	(2.1161)	(8.2829)	(2.1125)	(15.2771)	(2.0525)	(2.5933)	(1.6008)	(6.5106)	(2.4659)	(4.1648)	(0.9663)	(1.6255)
Non-ATS	−0.7537	−3.8992	−5.6417	−13.5405	−0.0713	−1.3031	−2.5961	−5.1088	1.2744	−4.7981	−1.9805	−1.1528
	(1.7209)	(20.6311)	(8.9817)	(23.9465)	(0.5469)	(3.3606)	(3.4256)	(6.9409)	(1.0993)	(9.4063)	(2.8166)	(1.9221)
COMP	−0.6109	−2.5195	−3.5137**	−6.5292	−1.7885***	−2.5556**	−2.7296***	−4.2902*	−1.9599**	−1.3490	−1.5920**	−1.9234***
	(1.3884)	(2.4534)	(1.4933)	(5.0722)	(0.5691)	(0.9991)	(0.9209)	(2.5443)	(0.9674)	(1.5336)	(0.5967)	(0.6227)
log(Price)	−0.1040	0.1256	0.3560	0.7037	−0.1793**	−0.0988	0.0471	0.0289	−0.1330*	−0.0045	0.0609	0.0189
	(0.1062)	(0.1330)	(0.3570)	(1.0705)	(0.0832)	(0.0720)	(0.0842)	(0.1672)	(0.0779)	(0.0971)	(0.0990)	(0.0752)
log(Volatility)	0.4354***	0.3524**	0.1332	−0.1988	0.4506***	0.3853***	0.2488***	0.1253	0.2371***	0.3811***	0.3106***	0.3047***
	(0.0427)	(0.1352)	(0.2800)	(0.8746)	(0.0326)	(0.0692)	(0.0897)	(0.2084)	(0.0400)	(0.0914)	(0.0811)	(0.0649)
log(Volume)	−0.2137***	−0.2672	−0.3254*	−0.6229	−0.2050***	−0.2695***	−0.2331***	−0.3268*	−0.1281***	−0.1115	−0.1862***	−0.1857***
	(0.0193)	(0.4938)	(0.1792)	(0.3879)	(0.0180)	(0.0709)	(0.0648)	(0.1748)	(0.0181)	(0.2191)	(0.0574)	(0.0569)

(Continues)

TABLE 6 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = *log(Quoted spread)*				$Y_{i,t}$ = *log(Effective half-spread)*				$Y_{i,t}$ = *log(StD returns)*			
C. 2020 ex-COVID	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500	Small	Medium	Large	S&P 500
Y_{Noti}	0.6218***	1.0990***	1.0405***	0.9140	0.5309***	0.5864**	0.4081*	−0.0891	0.7577***	0.7577***	0.5271***	0.4145***
	(0.0554)	(0.2720)	(0.2061)	(0.6475)	(0.0758)	(0.2877)	(0.2058)	(0.6081)	(0.0544)	(0.2271)	(0.0995)	(0.1094)
Observations	32,184	42,251	46,254	17,823	32,183	42,251	46,254	17,823	32,174	42,247	46,254	17,823
R-squared	0.1376	−1.7095	−1.8620	−8.9688	−0.1803	−0.9490	−1.0150	−2.3090	−0.2684	−2.7206	−0.3494	−0.1514
Firm#Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of analyzing the relationship between market shares and market quality for subsamples by market capitalization. We estimate the following four-equation panel regression model for market quality measures ($Y_{i,t}$) using IV/2SLS:

(1) $DP_{i,t} = a_i\, d_q + b_1\, W_{i,t} + c_1\, X_{i,t} + e_{1,i,t}$,
(2) $INT_{i,t} = a_i\, d_q + b_2\, W_{i,t} + c_2\, X_{i,t} + e_{2,i,t}$,
(3) $COMP_{i,t} = a_i\, d_q + b_3\, W_{i,t} + c_3\, X_{i,t} + e_{3,i,t}$,
(4) $Y_{i,t} = \alpha_i\, d_q + \beta_1\, DP_{i,t} + \beta_2\, INT_{i,t} + \beta_3\, COMP_{i,t} + \gamma\, X_{i,t} + e_{4,i,t}$,

where $X_{i,t}$ is a vector of control variables that includes $Y_{Noti,t}$, $log(Price)_{i,t}$, $log(Volatility)_{i,t}$, and $log(Volume)_{i,t}$. $W_{i,t}$ is a vector that includes $DP_{Noti,t}$, $INT_{Noti,t}$, and $COMP_{Noti,t}$, where *Noti,t* stands for the day *t* average across stocks in the same size group, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 and/or Nasdaq 100). For 2020, *DP* = *ATS* and *INT* = *Non-ATS*. The first-stage regressions (1)–(3) are reported in Table SA6.1. The second-stage regressions (4) use the fitted value from the first-stage regressions (1)–(3) of the endogenous variables.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

TABLE 7 Market stress and dark and lit market shares

	(1)	(2)	(3)	(4)	(5)	(6)
	A. 2009			B. 2020		
	DP	*INT*	*COMP*	*ATS*	*Non-ATS*	*COMP*
$I = H1$	−0.0124***	−0.0283***	0.0017	0.0029	−0.0136***	0.0168***
	(0.0011)	(0.0019)	(0.0013)	(0.0038)	(0.0048)	(0.0054)
$I = ret\_low$	−0.0037***	−0.0059***	0.0006	−0.0029**	−0.0072***	−0.0006
	(0.0007)	(0.0013)	(0.0007)	(0.0013)	(0.0017)	(0.0018)
$I = bs\_low$	−0.0010***	0.0042***	−0.0019***	−0.0044***	−0.0036***	0.0111***
	(0.0003)	(0.0007)	(0.0004)	(0.0012)	(0.0011)	(0.0019)
$I = vol\_extr$	−0.0116***	−0.0155***	−0.0022**			
	(0.0006)	(0.0013)	(0.0009)			
$I = vol\_high$				0.0047**	−0.0104**	0.0065*
				(0.0020)	(0.0047)	(0.0034)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of regressions of *DP* and *INT* on indicators of market stress for 2009 in Panel A and of *ATS* and *Non-ATS* on indicators of market stress for 2020 in Panel B. We estimate the following panel regression of market shares ($MS_{i,t}$) using OLS:

$$MS_{i,t} = \alpha_i\, d_q + \beta\, I^{Stress} + \gamma\, log(Volume)_{i,t} + e_{i,t},$$

where I^{Stress} is a stock-specific indicator for the first two quarters (*H1*), the lowest tercile of individual stock returns (*ret_low*), and the lowest tercile of stock-specific buy-order imbalances (*bs_low*). For 2009 (2020), we sample the highest decile (tercile) of individual stock volatility, *vol_extr (vol_high)*. Full results are provided in Table SA7.
$^*p < 0.1$. $^{**}p < 0.05$, $^{***}p < 0.01$.

lower during periods of market stress in 2009 for all our indicators of market stress except selling pressure. For 2020, we find that both types of dark trading are lower when returns are low. Trading reported by OTC market makers is lower also in the first half of the year when the pandemic caused the stock market to roil and when short-term volatility is high. By contrast, there is more dark pool trading when short-term volatility is high. This result is consistent with Zhu (2014): when volatility becomes extreme thus worsening substantially the lit spread, investors optimally choose to go dark rather than crossing the spread in the lit market. Results for selling pressure are more nuanced: while in 2009 high selling pressure induced traders to route orders to internalizing OTC market makers instead of to dark pools or lit competing venues to manage their inventories, in 2020 they seem to opt for the lit competing venues.[22] This might be because in 2020 we capture the peak of the COVID volatility period (see the VIX pattern in Figure 1). With the exception of weeks with high short-term volatility in 2020, these results show that traders route fewer orders to dark pools and that OTC market makers internalize less when markets are under stress.

Even though overall less volume executes in dark venues during periods of market stress, it is possible that the fact that any orders leave the lit market during these periods is harmful, and could worsen market quality exactly when needed the most. To examine whether different forms of dark trading, controlling for lit competition, have a harmful effect on spreads and short-term volatility, we estimate the panel regression in Equations (7)–(10) above for periods

[22] For example, Chiyachantana et al. (2004) show that traders face asymmetric price impact for their buy and sell orders in bull versus bear markets, thus presumably increasing concerns about inventory management costs when facing imbalance pressure.

of market stress. We drop the intraday range and stock price as control variables as discussed above, but continue controlling for volume and use stock fixed effects (but not quarter fixed effects). The results are in Table 8.

Regardless of market stress subsample, in 2009 more dark trading—through dark pools or internalizing OTC market makers—leads to lower quoted and effective half-spreads, and to lower short-term volatility even if the coefficient of *vol_extr* is not always significant.[23] A higher market share of lit competing venues also generally contributes to better market quality. Hence, there is no evidence based on our 2009 data that more dark trading, whether through dark pools or OTC market makers, leads to worse market quality in periods of market stress.

The picture is quite different for 2020. While there is no detrimental effect of dark pool trading during the first half of the year, we find that more dark pool trading leads to wider quoted spreads when returns are low and in particular when short-term volatility is high. This result is in line with Zhu (2014) where high volatility diverts traders to the dark pools to avoid crossing the large spread thus worsening the spread on the lit market. More dark pool trading during periods of high volatility also leads to wider effective half-spreads and higher short-term volatility. By contrast, more trading by OTC market makers leads to lower effective half-spreads and lower volatility during weeks when volatility is high. More lit competing volume leads to lower quoted and effective half-spreads, but only when returns are low and not during other periods of market stress. The 2020 market stress evidence thus shows that dark pool trading leads to wider spreads and higher short-term volatility, not just on average as shown in Table 5 and for Large caps as shown in Table 6, but also for periods of market stress on average.

6 | DISCUSSION

Our results show that dark trading is generally beneficial for market quality in 2009, but we find evidence suggesting a detrimental effect of dark pool trading, particularly for Large caps, in 2020. There are several potential explanations for the discrepancy between the results, and we discuss these in turn in this section.

One possible explanation is that the FINRA data are weekly, and the lower frequency may contribute to the lack of power to detect significant effects of dark trading on market quality for the 2020 samples, both on average when examining aggregate dark trading and for trades internalized by OTC market makers. The FINRA *Non-ATS* data are particularly noisy, and this may explain why we find no significant effects of trades reported by OTC market makers in any of our tests.[24] It is also possible that the noisy *Non-ATS* data cause problems for estimating the effect of dark pool trading on market quality. However, dropping the *Non-ATS* category from the analyses in Tables 5 and 6 does not change the conclusion, so this does not appear to be the explanation (Tables SA5.2 and SA6.2). For dark pool trading, the lack of power appears to be mostly an issue with Small and Medium caps. We do find consistently harmful effects of dark pool trading for the subsample of Large caps suggesting that at least for these stocks, we are able to pin down the effect of dark pool trading on market quality.

Another possibility is that the 2009 SIFMA data only cover the dark pools that voluntarily reported their trading volume, and that the "good" dark pools selected to report their trading activity and those that had a negative impact on market quality did not participate in the survey. To address this concern, we create a subset of the dark pools that existed in 2009, and still are operating in 2020. While we do not know if these dark pool survivors were in the original SIFMA sample, they should be more similar to the dark pools that reported to SIFMA in 2009. We add the dark pools that started operating after 2009 to the *Non-ATS* subsample, creating a similar aggregate dark trading measure that we have for the 2009 sample. If the self-reporting in 2009 created a selection on "good" dark pools, we hope to detect this using the new decomposition of dark trading for 2020. We repeat our analysis to study the effect of these two "pseudo" forms of dark trading on market quality for our 2020 sample to investigate if the dark pool survivors have

[23] These results are robust to defining low returns as the lowest decile of returns for each stock.

[24] FINRA *Non-ATS* includes an aggregate of smaller OTC market makers under the heading "De Minimis Firms." Their reporting appears to be much less consistent than that of the individually reported OTC market makers. For example, their reported volume often exceeds consolidated volume.

TABLE 8 Market shares and market quality in periods of market stress

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = *log(Quoted spread)*				$Y_{i,t}$ = *log(Effective half-spread)*				$Y_{i,t}$ = *log(StD returns)*			
A. 2009	*H1*	*ret_low*	*bs_low*	*vol_extr*	*H1*	*ret_low*	*bs_low*	*vol_extr*	*H1*	*ret_low*	*bs_low*	*vol_extr*
DP	−1.4871***	−1.0091***	−1.1675***	−0.7007**	−1.2870***	−1.0044***	−1.2241***	−0.4878	−1.5007***	−0.9676***	−0.8524***	−0.6642
	(0.2311)	(0.2646)	(0.2577)	(0.2888)	(0.2999)	(0.3158)	(0.2923)	(0.4512)	(0.2687)	(0.2277)	(0.1780)	(0.5399)
INT	−0.2866**	−1.0257***	−1.3631***	0.0104	−0.1881	−0.8969***	−0.8783***	0.1595	−0.7134**	−0.7283***	−1.0528***	0.0399
	(0.1253)	(0.1679)	(0.1889)	(0.3253)	(0.1842)	(0.1888)	(0.1998)	(0.2576)	(0.3578)	(0.2225)	(0.2187)	(0.3611)
COMP	−0.4059***	0.0331	−0.4480**	−0.0560	−1.0376***	−0.4521**	−0.5522***	−0.9762***	−0.6728***	−0.6674***	−0.7427***	−0.5349**
	(0.1364)	(0.1843)	(0.1850)	(0.2205)	(0.2076)	(0.2253)	(0.2096)	(0.3111)	(0.1869)	(0.1968)	(0.1849)	(0.2686)
log(Volume)	−0.0998***	−0.1122***	−0.1003***	−0.1232***	0.0261***	0.0232***	0.0217***	0.0374***	0.2089***	0.1947***	0.1733***	0.1999***
	(0.0062)	(0.0057)	(0.0066)	(0.0099)	(0.0067)	(0.0062)	(0.0067)	(0.0084)	(0.0111)	(0.0066)	(0.0071)	(0.0108)
Y_{Noti}	0.8936***	0.9200***	0.9277***	0.7820***	0.8072***	0.8164***	0.8431***	0.6755***	0.8823***	0.8463***	0.8923***	0.4679***
	(0.0147)	(0.0182)	(0.0201)	(0.0252)	(0.0219)	(0.0226)	(0.0226)	(0.0290)	(0.0216)	(0.0200)	(0.0167)	(0.0283)
Observations	341,859	233,943	233,976	69,231	341,412	233,767	233,807	69,158	341,859	233,999	233,997	69,237
R-squared	0.3056	0.3101	0.2454	0.3521	0.0803	0.1184	0.1145	0.0470	0.2472	0.3821	0.3742	0.1792
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = *log(Quoted spread)*				$Y_{i,t}$ = *log(Effective half-spread)*				$Y_{i,t}$ = *log(StD returns)*			
B. 2020	*H1*	*ret_low*	*bs_low*	*vol_high*	*H1*	*ret_low*	*bs_low*	*vol_high*	*H1*	*ret_low*	*bs_low*	*vol_high*
ATS	−0.0232	4.0580**	4.5235	10.1089***	−1.4955*	1.9940	2.1741	6.2900***	−2.4471**	1.6395	0.9477	7.3868***
	(1.0925)	(1.6072)	(2.9989)	(3.3351)	(0.7556)	(1.5049)	(1.5690)	(1.4479)	(1.1258)	(1.7083)	(1.7871)	(2.2719)
Non-ATS	−2.9879*	−2.9225	−7.9555	−8.8091*	0.0340	0.4225	−1.6477	−4.6493**	−2.7477	−2.5195	−4.1004*	−7.9909***
	(1.6606)	(3.7037)	(9.1517)	(5.0390)	(1.1235)	(3.2030)	(1.9815)	(2.2460)	(1.6560)	(2.9541)	(2.3987)	(2.5421)

(Continues)

TABLE 8 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$Y_{i,t}$ = log(Quoted spread)				$Y_{i,t}$ = log(Effective half-spread)				$Y_{i,t}$ = log(StD returns)			
B. 2020	*H1*	*ret_low*	*bs_low*	*vol_high*	*H1*	*ret_low*	*bs_low*	*vol_high*	*H1*	*ret_low*	*bs_low*	*vol_high*
COMP	0.7615	−3.0749***	−1.2737	−5.8588	−0.1149	−3.1172***	−1.8256*	−2.9970*	1.7482	−1.4933	−0.3363	−0.9788
	(1.2378)	(1.0691)	(2.3984)	(3.6811)	(0.8490)	(0.8585)	(1.0186)	(1.5648)	(1.2140)	(1.0868)	(1.2590)	(2.7098)
log(Volume)	−0.0359	−0.1569***	−0.0059	−0.1384	−0.0644**	−0.1708***	−0.0699	−0.0587	0.0548	−0.0630	0.0354	0.0381
	(0.0461)	(0.0557)	(0.1694)	(0.0972)	(0.0309)	(0.0537)	(0.0449)	(0.0493)	(0.0415)	(0.0464)	(0.0508)	(0.0619)
Y_{Noti}	1.1088***	1.0581***	0.9288**	0.6740**	1.2139***	1.1370***	1.0453***	0.5556***	1.1076***	0.9559***	0.8738***	0.1827
	(0.0974)	(0.1986)	(0.4001)	(0.2925)	(0.0805)	(0.2156)	(0.1292)	(0.1532)	(0.1041)	(0.1663)	(0.1190)	(0.1768)
Observations	74,139	50,792	50,641	47,957	74,137	50,792	50,639	47,956	74,123	50,789	50,631	47,950
R-squared	0.3573	−0.0753	−1.8301	−4.1119	0.5431	0.2221	0.1699	−1.9246	0.2402	0.2073	−0.1987	−5.5483
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
The table reports the results of regressions of Quoted and Effective spreads on indicators of market stress for 2009 in Panel A and for 2020 in Panel B. We estimate the following panel regression of market shares ($MS_{i,t}$) using IV/2SLS:

(1) $DP_{i,t} = a_i\, d_q + b_1\, W_{i,t} + c_1\, X_{i,t} + e_{1,i,t}$,
(2) $INT_{i,t} = a_i\, d_q + b_2\, W_{i,t} + c_2\, X_{i,t} + e_{2,i,t}$,
(3) $COMP_{i,t} = a_i\, d_q + b_3\, W_{i,t} + c_3\, X_{i,t} + e_{3,i,t}$,
(4) $Y_{i,t} = \alpha_i\, d_q + \beta_1\, DP_{i,t} + \beta_2\, INT_{i,t} + \beta_3\, COMP_{i,t} + \gamma\, X_{i,t} + e_{4,i,t}$,

where $X_{i,t}$ is a vector of control variables that includes $Y_{Noti,t}$ and *log(Volume)*$_{i,t}$. $W_{i,t}$ is a vector that includes $DP_{Noti,t}$, $INT_{Noti,t}$, and $COMP_{Noti,t}$, where *Noti,t* stands for the day *t* average across stocks in the same size group, but that are Not in same four-digit SIC code, and Not in same index (S&P 500 and/or Nasdaq 100). For 2020, *DP* = *ATS* and *INT* = *Non-ATS*. The second-stage regressions (4) use the fitted value from the first-stage regressions (1)–(3) of the endogenous variables. Market stress periods are the first two quarters (*H1*), the lowest tercile of individual stock returns (*ret_low*), and the lowest tercile of stock-specific buy-order imbalances (*bs_low*). For 2009 (2020), we sample the highest decile (tercile) of individual stock volatility, *vol_extr* (*vol_high*).
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

a positive impact on market quality in 2020. The results show that neither the dark pool survivors nor the combined *Non-ATS* plus new dark pools aggregate has any effect on market quality for any of the subsamples by size (Table SA6.3). Hence, while the trading that takes place through dark pool survivors appears to be less harmful than that of dark pool trading on average, we do not find that the dark pool survivors are beneficial for market quality. We conclude that the selection induced by self-reporting is unlikely to explain the differences we observe between 2009 and 2020.

Other possible explanations relate to the fact that the mix of orders executed in dark pools has changed significantly between 2009 and 2020. Three particular trends deserve mention in this regard: proprietary trading, HFTs, and retail trading activity. We discuss each in turn. Dark pools' main selling point was originally that they gave institutional traders the opportunity to execute large trades without causing markets to move against them. They explicitly screened out order flow that could be toxic such as proprietary traders and HFTs. Quoting former SEC Commissioner Luis A. Aguilar:[25] "Dark pools initially portrayed themselves as havens from predatory traders." He continued: "They achieved this, in part, by excluding high-frequency traders, who supposedly use brute speed to front run institutional investors' large orders." As competition for order flow increased, dark pool operators started welcoming (albeit not always openly) both proprietary traders and HFTs, recognizing that they provide additional liquidity and increase the probability of execution.[26] However, since some HFTs use sophisticated pinging strategies to detect hidden orders, and proprietary traders may front-run large orders, allowing them access to dark pools may reduce the benefit of these venues for institutional traders (Korajczyk & Murphy, 2018; Van Kervel & Menkveld, 2019).

The second trend is the dramatic rise in retail trading activity that took place in 2020 as the United States went into lockdown sequestering most everyone at home. Retail brokerages experienced tremendous growth in accounts, as well as in trading activity, as individuals turned to the stock market for entertainment and distraction. For example, retail broker Robinhood had 13 million users at the end of 2020, up 30% from 2019.[27] Virtu Financial, one of the largest OTC market makers, estimates that retail represented over 30% of trading in late 2020, up from about 17% at the beginning of the year.[28] Brokers route retail orders to an OTC market maker, a dark pool, or to an exchange for execution. Retail orders typically receive price improvement of say 0.1 cents per share relative to the NBBO when internalized by an OTC market maker, and this fact means that a proxy for retail trading is the volume of sub-penny executions. While the traditional view was that retail traders are uninformed, evidence using this proxy suggests that their order imbalances predict future returns (Boehmer et al., 2021). In practice, OTC market makers absorb retail order imbalances only when they have access to offsetting institutional orders, either via their own Single Dealer Platforms (SDPs) such as Citadel Connect or in dark pools (Barardehi et al., 2022).[29] During 2020, retail traders had an adversely affected stock liquidity during periods of market stress, timed their trades well relative to future returns, and generated an alpha (e.g., Ozik et al., 2021; Pagano et al., 2021; Welch, 2021). We conjecture that an increasing volume of potentially market moving retail order imbalances likely reached dark pools in 2020, both directly when routed by retail brokers, and indirectly as OTC market makers laid off order imbalances.

In sum, the mix of order flow that reaches dark pools likely includes both more proprietary orders, orders from HFTs and more retail order imbalances in 2020 compared to 2009. Research suggests that these types of order flow move prices and potentially contribute to short-term volatility. Hence, dark pools have gone from being venues where institutional traders could find liquidity while avoiding broadcasting their trading intentions, to venues where they

[25] https://www.sec.gov/news/statement/shedding-light-on-dark-pools.html.

[26] In 2014, the New York Attorney General sued Barclays for its dark pool operations, specifically for misstating the level of HFT activity in its dark pool, thus defrauding investors. In January 2016, Barclays agreed to pay a fine of $35 million to the SEC and $70 million to the New York Attorney General for its misconduct related to the dark pool.

[27] Robinhood was not alone in experiencing strong growth in retail accounts. Fidelity had 26 million retail accounts at the end of 2020, up 17% from a year earlier, and Charles Schwab had 30 million active accounts, up 13% from a year earlier (net of acquisitions of TD Ameritrade and USAAs investment management company).

[28] Virtu Financial, Inc. (2020) and McCrank (2021).

[29] SDPs are not ATSs according to the current rules, and they therefore do not report to FINRA. FINRA recently proposed to expand OTC equity trading data published on FINRA's website to include SPD trading (Regulatory Notice 18-28).

face an increasing amount of pinging and front running from proprietary traders and HFTs, and market moving order flow from retail investors. We conjecture that dark pools are less attractive to institutional traders in 2020 compared to 2009. This is consistent with the aggregate data—the market share of dark pools has declined from 14.5% in 2016 to 10.1% in 2020, while trading internalized by OTC market makers has increased from 22.1% to 31.2%. This trend has continued, and dark pools represent 8.1%, while OTC market makers represent 37.2% of volume in 2021.[30] Our weekly data do not permit us to study the mix of traders in dark pools, but it is clearly a topic of interest for future research.

7 | CONCLUSIONS

We study dark trading based on 2009 and 2020 data. Each sample includes roughly 3000 stocks and covers a tumultuous period in U.S. stock markets, related to the Great Financial Crisis in 2009 and the COVID pandemic in 2020, respectively. This permits us to study dark trading for different stocks (Small and Large caps), of different types (dark pools and internalized trades), and during periods of market stress. It also allows us to examine whether dark trading plays a different role in 2020 compared to 2009.

The picture that emerges is that dark trading activity differs systematically across stocks and between dark pools and internalization by OTC market makers. For example, dark pools are more active for large caps, while OTC market makers internalize more for small caps. OTC market makers internalize more when spreads are wide and depth is high so they can more easily lay off order imbalances in the lit market. By contrast, how the dark pool market share depends on the order book changes between samples for large caps. In the earlier sample, traders in large caps use dark pools to jump the queue when the order book is competitive; in the later sample, they instead use dark pools to avoid crossing a wide spread. The effect of dark trading on market quality has also changed. A higher dark pool market share leads to lower quoted and traded spreads and does not affect short-term volatility in the early sample. By contrast, we find that a higher dark pool market share leads to wider spreads and higher short-term volatility for large caps in the later sample. Similarly, more internalization by OTC market makers leads to improved spreads in the early sample, but does not affect market quality in the later sample.

While a full explanation for these changes is beyond the scope of the current paper, we speculate that the difference we observe between the two samples arises because proprietary order flow, HFTs, and retail order imbalances reach dark pools in recent years, and that these venues have become less attractive for institutional traders as a result. This conjecture is consistent with statements from the former SEC Commissioner Luis A. Aguilar[31] and with the recent decline in dark pool trading. Taken together, the evidence we present shows that dark trading is evolving over time, and we believe that research based on more granular recent data is needed to better understand what role it will play going forward.

ACKNOWLEDGMENTS

We are grateful to SIFMA for assisting us in collecting the dark pool data that form the basis for this study. We also thank Jamie Selway for making this study possible, Kewei Hou for helpful comments on the empirical design, Minki Kim and Ming Ju for research assistance, participants at the Hong Kong University of Science and Technology Symposium, the Notre Dame Conference on Current Topics in Financial Regulation, and the University of Cyprus academic conference, and seminar participants at Columbia University for comments.

[30] Hadiaris (2021).

[31] Op. cit.

REFERENCES

Anand, A., Samadi, M., Sokobin, J., & Venkataraman, J. (2021). Institutional order handling and broker-affiliated trading venues. *Review of Financial Studies*, *34*, 3364–3402.

Aquilina, M., Foley, S., O'Neill, P., & Ruf, T. (2021). *Sharks in the dark: Quantifying latency arbitrage*. Working paper. https://ssrn.com/abstract=2848120

Bacidore, J. M. (2020). *Algorithmic trading: A practitioner's guide*. TBG Press.

Barardehi, Y. H., Bernhardt, D., Da, Z., & Warachka, M. (2022). *Institutional liquidity demand and the internalization of retail order flow: The tail does not wag the dog*. https://ssrn.com/abstract=3966059

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Trading retail investor activity. *Journal of Finance*, *76*, 2249–2305.

Buti, S., Rindi, B., & Werner, I. M. (2017). Dark pool trading strategies, market quality and welfare. *Journal of Financial Economics*, *124*, 244–265.

Chakravarty, S., Jain, P., Upson, J., & Wood, R. (2012). Clean sweep: Informed trading through intermarket sweep orders. *Journal of Financial and Quantitative Analysis*, *47*, 415–35.

Chiyachantana, C. N., Jain, P. K., Jiang, C., & Wood, R. A. (2004). International evidence on institutional trading behavior and price impact. *Journal of Finance*, *59*, 869–898.

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

Comerton-Forde, C., Malinova, K., & Park, A. (2018). Regulating dark trading: Order flow segmentation and market quality. *Journal of Financial Economics*, *130*, 347–366.

McCrank, J. (2021). *Factbox: The U.S. retail trading frenzy in numbers*. Reuters. https://www.reuters.com/article/us-retail-trading-numbers/factbox-the-u-s-retail-trading-frenzy-in-numbers-idUSKBN29Y2PW

Degryse, H., de Jong, F., & van Kervel, V.. (2015). The impact of dark and visible fragmentation on market quality. *Review of Finance*, *19*, 1587–1622.

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., & Menkveld, A. J. (2008). Competition for order flow and smart order routing systems? *Journal of Finance*, *63*, 119–158.

Gawronski, J., & Schack, J. (2010). Let there be light, Rosenblatt's Monthly Dark Liquidity Tracker, Special Issue: 2009. Rosenblatt Securities Inc.

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, *16*, 646–679.

Hadiaris, J. (2021). *Cowen market structure: Retail trading – What's going on, what may change, and what can you do about it?* https://www.cowen.com/insights/retail-trading-whats-going-on-what-may-change-and-what-can-institutional-traders-do-about-it/

IIROC. (2015). *Study of the impact of the dark rule amendments, IIROC Notice*. https://www.iiroc.ca/Documents/2015/bcb13158-7b37-4780-9962-8c0a751e12b0_en.pdf

Korajczyk, R. A., & Murphy, D. (2018). High frequency market making to large institutional trades. *Review of Financial Studies*, *32*, 1034–1067.

Kwan, A., Masulis, R., & McInish, T. H. (2015). Trading rules, competition for order flow and market fragmentation. *Journal of Financial Economics*, *115*, 330–348.

Mittal, H. (2008). Are you playing in a toxic dark pool? A guide to preventing information leakage. *Journal of Trading*, *3*(3), 20–33.

O'Hara, M., & Ye, M. (2011). Is market fragmentation harming market quality? *Journal of Financial Economics*, *100*, 459–474.

Ozik, G., Sadka, R., & Shen, S. (2021). Flattening the liquidity curve: Retail trading during the COVID-19 lockdown. *Journal of Financial and Quantitative Analysis*, *56*, 2356–2388.

Pagano, M. S., Sedunov, J., & Velthuis, R. (2021). How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality. *Finance Research Letters*, *43*, 1–11.

Securities and Exchange Commission. (2005). *Regulation National Market System (NMS), 17 CFR PARTS 200, 201, 230, 240, 242, 249, and 270* (Release No. 34–51808; File No. S7-10-04). Author.

Securities and Exchange Commission. (2010). *Concept Release on Equity Market Structure, 17 CFR PART 242* (Release No. 34–61358; File No. S7-02-10). Author.

Securities and Exchange Commission, (2014a). In the Matter of LavaFlow Inc. (Securities Exchange Act Release No. 72673). http://www.sec.gov/litigation/admin/2014/34-72673.pdf

Securities and Exchange Commission, (2014b). In the Matter of Liquidnet, Inc. (Securities Exchange Act Release No. 72339). http://www.sec.gov/litigation/admin/2014/33-9596.pdf

Securities and Exchange Commission, (2015). In the Matter of UBS Securities LLC (Securities Exchange Act Release No. 74060). http://www.sec.gov/litigation/admin/2015/33-9697.pdf

Securities and Exchange Commission. (2021). *Staff Report on Equity and Options Market Structure Conditions in Early 2021*. https://www.sec.gov/files/staff-report-equity-options-market-struction-conditions-early-2021.pdf

Van Kervel, V., & Menkveld, A. J. (2019). High-frequency trading around large institutional orders. *Journal of Finance*, *74*, 1091–1137.
Virtu Financial, Inc. (2020). *Annual report*. https://2020annualreport.virtu.com/
Welch, I. (2021). The wisdom of the Robinhood crowd. *Journal of Finance*, forthcoming.
Yao, C., & Ye, M. (2018). Why trading speed matters: A tale of queue rationing under price controls. *Review of Financial Studies*, *31*, 2157–2183.
Zhu, H. (2014). Do dark pools harm price discovery. *Review of Financial Studies*, *27*, 747–789.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Buti, S., Rindi, B., & Werner, I. M. (2022). Diving into dark pools. *Financial Management*, e12395. https://doi.org/10.1111/fima.12395