Shades of Darkness: A Pecking Order of Trading Venues[1]

Albert J. Menkveld[2]   Bart Zhou Yueshen[3]   Haoxiang Zhu[4]

First version: October 16, 2014
This version: December 23, 2014

[2] VU University Amsterdam, Tinbergen Institute, and Duisenberg School of Finance; FEWEB, VU, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands; +31 20 598 6130; albertjmenkveld@gmail.com.

[3] INSEAD; Boulevard de Constance, 77300 Fontainebleau, France; +33 1 60 72 42 34; b@yueshen.me.

[4] MIT Sloan School of Management and NBER; 100 Main Street E62-623, Cambridge, MA 02142, USA; +1 617 253 2478; zhuh@mit.edu.

Shades of Darkness: A Pecking Order of Trading Venues

Abstract

Investors trade in various types of venues. When demanding immediacy, they trade off price impact and execution uncertainty. The "pecking order" hypothesis (POH) states that investors rank venues accordingly, with low-cost-low-immediacy venues on top and high-cost-high-immediacy venues at the bottom. Hence, midpoint dark pools on top, non-midpoint pools in the middle, and lit markets at the bottom. When urgency increases, investors tilt their flow from top to bottom. We document such pattern for U.S. data, confirming POH. A simple model obtains POH in equilibrium and suggests that the availability of dark pools reduces investor (utility) cost by \$1.43 billion annually.

Keywords: dark pool, pecking order, fragmentation, high-frequency trading

JEL Classifications: G12, G14, G18, D47

1   Introduction

A salient trend in global equity markets over the last decade is the explosion of off-exchange, or "dark" trading venues. In the United States, dark venues now account for about 30% of equity trading volume (see Figure 1(a) for an illustration of Dow-Jones stocks). In Europe, about 40% of volume trades off-exchange for leading equity indices (see Figure 1(b)).

Equally salient is the wide fragmentation of trading volume across dark venues. The United States has more than 30 "dark pools" and more than 200 broker-dealers that execute trades away from exchanges (see SEC, 2010). Dark pools, which are automated traded systems that allow investor-to-investor trades, have grown fast in market shares and now account for about 15% of equity trading volume in the U.S., according to industry estimates.[1] In Europe, dark venues also face a high degree of fragmentation, with at least 10 multilateral dark venues operating actively.[2]

The fragmentation of trading—between exchanges and dark venues, as well as across dark venues—is a double-edged sword. It creates a conflict between the efficient interaction among investors and investors' demand for a diverse set of trading mechanisms. The SEC (2010, p. 11-12) highlights this trade-off in its Concept Release on Equity Market Structure:

> "Fragmentation can inhibit the interaction of investor orders and thereby impair certain efficiencies and the best execution of investors' orders. . . . On the other hand, mandating the consolidation of order flow in a single venue would create a monopoly and thereby lose the important benefits of competition among markets. The benefits of such competition include incentives for trading centers to create new products, provide high quality trading services that meet the needs of investors, and keep trading fees low."

These two views of fragmentation neatly correspond to two complementary strands of theories on fragmented markets. Early theories involving multiple trading venues show that investors are

---

[1] Industry estimates are provided by Tabb Group, a consultancy firm, and Rosenblatt Securities, a broker. On June 2, 2014, FINRA started publishing weekly statistics of transaction volumes in alternative trading systems (ATS), with a two-week lag. Many dark pools are registered as ATS. For more details, see https://www.finra.org/Newsroom/NewsReleases/2014/P519139.

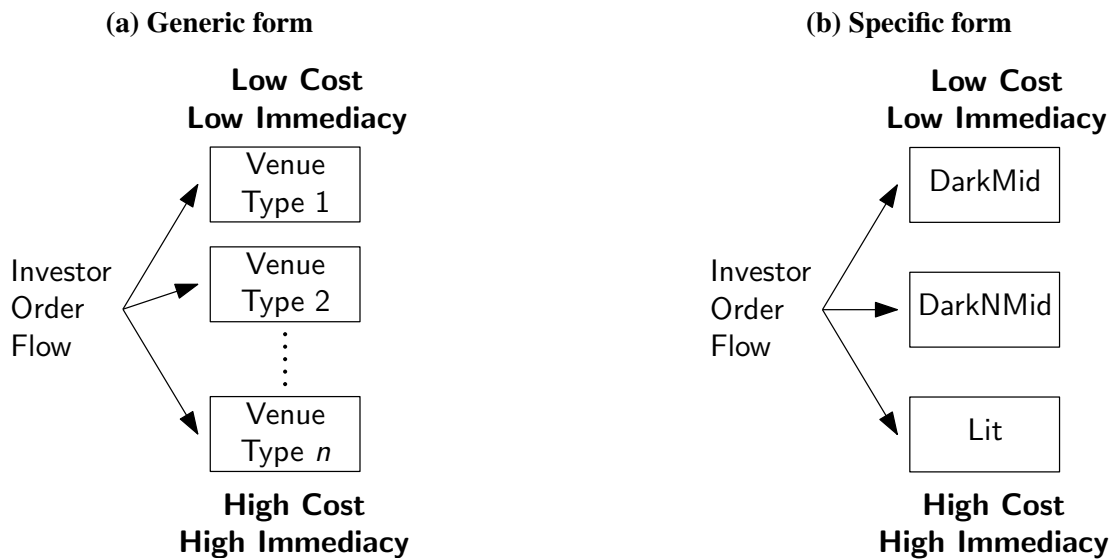[2] See "European Dark Trading Analysis," Fidessa, October 2013.

Figure 1: Dark market share in United States and Europe. This figure shows the market shares of dark trading in the U.S. and in Europe. Panel (a) plots the monthly average dark shares of the 30 stocks in the Dow Jones Industrial Average from 2006 to October 2014. We use the stocks that are in the Dow index as of November 2014. Volume data are obtained from Bloomberg and TAQ. Dark trades are defined by those reported to FINRA (code "D" in TAQ definition). From May 2006 to February 2007 the estimates are missing because (i) TAQ data mixes trades reported to FINRA with some NASDAQ trades, and (ii) Bloomberg data do not cover this period. Panel (b) plots the averages of dark shares of FTSE100, CAC40, and DAX30 index stocks. These estimates are directly obtained from Fidessa.

(a) U.S.: Dow 30 stocks **(b) Europe**
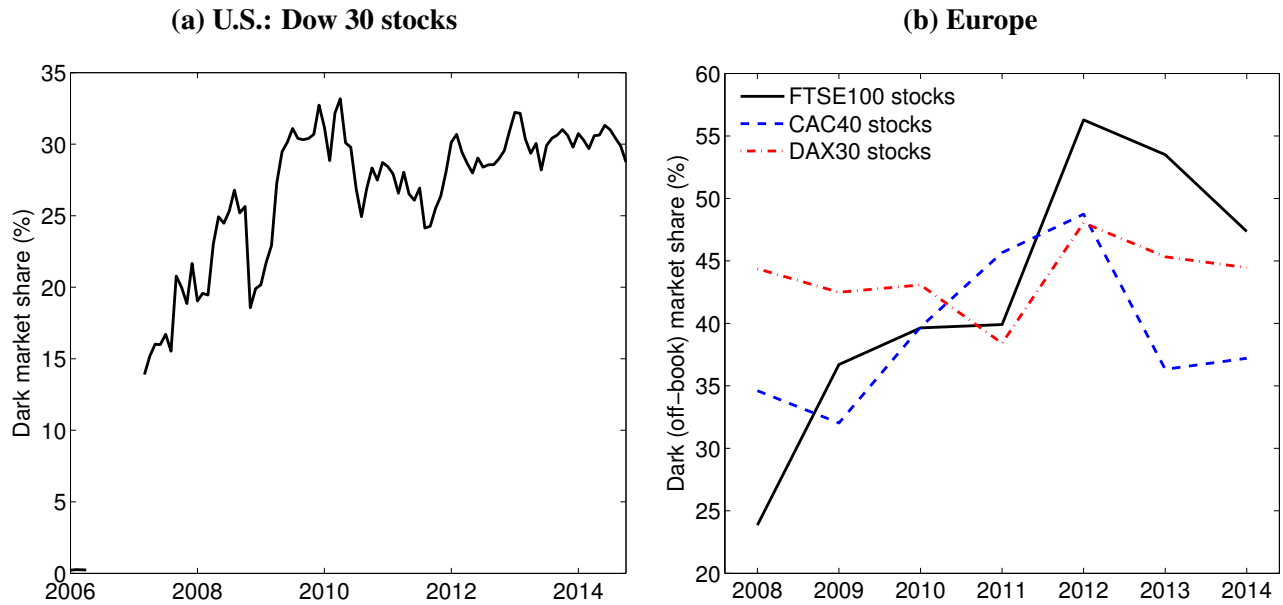
attracted to venues where other investors are trading (see, for example, Pagano, 1989; and Chowdhry and Nanda, 1991). This "liquidity-begets-liquidity" insight suggests that fragmentation reduces the efficiency of trading among investors. Applied to today's equity markets, this argument further implies that fragmentation of dark venues, which provide little or no pre-trade transparency, causes a particular concern because investors cannot observe the presence of counterparties *ex ante* and must engage in costly search in multiple dark venues.

In contrast, recent theories of dark pools show that precisely because of their pre-trade opacity and the associated execution uncertainty, dark venues attract a different type of investors from those on the exchanges (Hendershott and Mendelson, 2000; Degryse, Van Achter, and Wuyts, 2009; Buti, Rindi, and Werner, 2014, Ye, 2011; Zhu, 2014; Brolley, 2014). Under this "venue heterogeneity" or "separating equilibrium" view, fragmentation can be an equilibrium response to the heterogeneity of

2

Figure 2: Pecking order hypothesis. This figure depicts the pecking order hypothesis. Panel (a) shows the generic form. Panel (b) shows the specific form. Midpoint crossing dark pools (DarkMid) are on top. Non-midpoint dark pools (DarkNMid) are in the middle. Lit market (Lit) are at the bottom. Detailed descriptions of various dark pool types are collated in Section 3.1.

(a) Generic form

Low Cost
Low Immediacy

Investor Order Flow →	Venue Type 1
	Venue Type 2
	⋮
	Venue Type *n*

High Cost
High Immediacy

(b) Specific form

Low Cost
Low Immediacy

Investor Order Flow →	DarkMid
	DarkNMid
	Lit

High Cost
High Immediacy

investors and time-varying market conditions.

Pecking order hypothesis

This paper characterizes the dynamic fragmentation of U.S. equity markets. We propose and test a "pecking order hypothesis" for the fragmentation of trading volumes, in line with the second view in the SEC remark. We hypothesize that investors "sort" dark and lit venues by the associated costs (price impact, information leakage) and immediacy, in the form of a "pecking order."[3] The top of the pecking order consists of venues with the lowest cost and lowest immediacy, and the bottom of the pecking order consists of venues with the highest cost and highest immediacy. The pecking order hypothesis predicts that, on average, investors prefer searching for liquidity in low-cost, low-immediacy venues but will move towards higher-cost, higher-immediacy venues if their trading needs become more urgent. This intuitive sorting is illustrated in Figure 2(a).

---

[3] We choose "immediacy" here as we believe it is the appropriate interpretation of liquidity in the context of securities trading (Grossman and Miller, 1988).

More concretely, recent theories of dark pools mentioned above predict that dark venues are at the top of the pecking order, whereas lit venues are at the bottom, because of their difference in execution probabilities. In addition, through a simple stylized model, we rank two important categories of dark pools—midpoint dark pools ("DarkMid") and non-midpoint dark pools ("DarkNMid")—according to investors' cost and immediacy considerations. The specific ordering of the three venue types is illustrated in Figure 2(b). This sorting combines—in an intuitive description of the underlying theory—"exchanges are liquidity of last resort" and "not all dark pools are created equal."

We test the pecking order hypothesis by exploiting a unique dataset on dark trading in U.S. equity markets. Our dataset disaggregates dark transactions into five categories by trading mechanism, including the two types of dark pools shown in Figure 2(b). The other three categories of dark transactions are internalized trades with retail brokers, average-price trades, and other (mostly institutional) trades. (The detailed descriptions are provided in Section 3.) To the best of our knowledge, this dataset provides the most comprehensive and granular view of U.S. dark trading that is accessible by academics.

We estimate a panel vector-autoregressive model with exogenous variables (VARX) to characterize the dynamic interrelation among dark volumes, high-frequency trading activity in lit venues, and various market condition measures. The focus of our empirical strategy is on two exogenous variables, VIX and earnings surprise, which we use as proxies for shocks to investors' demand for immediacy, due to aggregate volatility and firm-specific news, respectively. The pecking order hypothesis predicts that, following an upward shock to VIX or earnings surprise, trading volumes should migrate from low-cost low-immediacy venues to high-cost high-immediacy venues; therefore, the change in volume share of a venue should become progressively larger the further out it is in the pecking order.

The data support the pecking order hypothesis. Following a 100% upward shock to VIX, dark pools that cross orders at the midpoint lose 52% of their market share (from 2.3% to 1.2%), dark pools that allow some price flexibility lose 30% of their market share (from 7.1% to 5.0%), and lit venues gain 16% of their market share (from 77.5% to 90.0%). The same qualitative pattern is

observed following an earnings-surprise shock.

We also develop a simple, parsimonious model to formally examine the pecking order hypothesis. The model serves two main purposes. First, it provides a "minimalism" framework under which investors strategically split their order flows across a variety of trading venues—including a lit exchange and multiple heterogeneous dark pools. The pecking order hypothesis arises from our model as an equilibrium outcome. Second, the model offers a natural setting to compare social welfare under different market structures: trading with or without dark venues. Guided by the model, we estimate that the existence of midpoint dark pools and non-midpoint dark pools achieves a cost saving of about $1.43 billion per year in the U.S. equity market. To put this dollar amount into perspective, $1.43 billion is about 7.4% of the annual value added by U.S. mutual funds (gross of fees) estimated by Berk and van Binsbergen (2014). In sum, market structure matters and has a meaningful impact on investor welfare.

At this point it is important to make a few cautionary remarks on the interpretation of our results. First, our simple model abstracts from information asymmetry. With information asymmetry, price discovery and exchange liquidity will be affected by the presence of dark venues, which has ambiguous additional welfare implications (see Zhu 2014). Second, while our results suggest that the fragmentation among dark venue types can be an equilibrium response to investor heterogeneity and time-varying market conditions, our analysis is silent on the fragmentation within each dark venue type. The latter question requires more detailed data on venue identities, not only venue types. Third, the pecking order hypothesis implicitly assumes that at least some investors make rational venue choices based on correct information of how these venues operate. This point is important in light of recent cases about alleged fraud by a couple of dark pools for misrepresenting information to investors.[4] Fourth, our analysis ignores the fact that investors can place "hidden" orders in Lit. It is, however, hard to completely hide orders on lit venues as these orders are frequently detected

---

[4] For example, on June 25, 2014, Eric Schneiderman, Attorney General of the State of New York, alleges that Barclays has falsified marketing material and misrepresented information to clients about the presence of high-frequency traders in its dark pool. In October 2011, SEC finds that Pipeline, a dark-pool operator that claimed to only allow buyside firms to participate, had filled the majority of customer orders through its own trading affiliate (see www.sec.gov/news/press/2011/2011-220.htm).

(de Winne and D'Hondt, 2007). These caveats are likely to garble the pecking order hypothesis and its welfare implications.

The primary contribution of this paper to the literature consists of three parts: (i) we characterize the *dynamic* fragmentation of dark and lit venues, (ii) we offer a "pecking order" explanation of the observed market-share pattern as an equilibrium outcome of a simple and parsimonious model, and (iii) based on the model, we quantify the welfare implications of market fragmentation.

Our result that adding dark venues improves welfare for investors is consistent with Degryse, Van Achter, and Wuyts (2009), who model the competition between a dealers market and a crossing network (equivalent to the midpoint dark pool in our model). They find that adding a crossing network improves welfare for investors; but if dealers' profits are included, then the welfare effect is ambiguous. Buti, Rindi, and Werner (2014) model the competition between a limit order book and a midpoint dark pool. They find numerically that adding the dark pool harms small traders' welfare but may harm or improve large traders' welfare. A main feature of their model is that there are no specialized liquidity providers who post sufficiently good quotes on the limit order book, so attractive quotes must be provided by the traders. This modeling feature helps explain why the welfare result of Buti, Rindi, and Werner (2014) differ from those of Degryse, Van Achter, and Wuyts (2009) and ours. Neither of these two theory papers calibrate to the data.

Prior empirical studies of fragmentation and dark venues typically focus on standard measures of "market quality," such as spread, depth, and volatility. They do not make welfare statements or estimates. In addition, they are based on either aggregate off-exchange trades or trades in a few selected dark pools. Studies based on aggregate trades include O'Hara and Ye (2011) and Hatheway, Kwan, and Zheng (2013) (U.S. data), Degryse, de Jong, and van Kervel (2014) (Dutch data), and Comerton-Forde and Putniņš (2013) (Australian data), among others. Studies based on trades in a few dark pools include Hendershott and Jones (2005) (Island ECN), Ready (2014) (Liquidnet and POSIT), Buti, Rindi, and Werner (2011) (11 anonymous dark pools), Boni, Brown, and Leach (2012) (Liquidnet), Nimalendran and Ray (2014) (one anonymous dark pool), Foley, Malinova, and Park

(2013) (dark order types on Toronto Stock Exchange), and Foley and Putniņš (2014) (four dark venues in Canada), among others.

Kwan, Masulis, and McInish (2014) use a similar dataset from NASDAQ. They study how the minimum tick size affects the competitiveness of exchanges relative to dark venues, which is a very different research question from ours. SEC (2013) provides a comprehensive review of the recent empirical literature on fragmentation.

Venue pecking order and high-frequency trading

One commonly cited reason for the increase in dark volume is that investors "fear" high-frequency traders (HFTs) in lit markets.[5] HFTs are proprietary traders who use "extraordinarily high-speed and sophisticated computer programs for generating, routing, and executing orders" (SEC 2010). Some investors and regulators are concerned that HFT may engage in "front-running" or "predatory trading." In the context of the pecking order hypothesis, the perceived risk of interacting with HFTs is one form of cost for trading in lit venues.

Motivated by these concerns, we are interested in the extent to which HFT activity affects investors' order flow along the venue pecking order. Without an exogenous shock to HFT activity, we attempt a preliminary, partial answer by studying the responses of HFT activity to VIX and earnings shocks. Specifically, if HFT activity declines after the VIX and earnings shocks, this decline would be at least "consistent with" the conjecture that investors trade more in lit venues because they face a lower risk of interacting with HFTs. If the data show the opposite, this conjecture is a lot less likely.

Empirically, we find that HFT quoting and trading activities both increase significantly in large stocks following an upward shock to VIX. The changes in HFT activity after earnings shocks are statistically insignificant. This result suggests that investor order flow migrates to lit venues under a higher aggregate volatility *not because* HFTs scale back in activity. Instead, it suggests that investors' fear for interacting with HFT, if any, is dominated by their demand for immediacy in high volatility situations.

---

[5] See, for example, "As market heats up, trading slips into shadows," New York Times, March 31, 2013.

Our finding that a high HFT activity is associated with a high *exogenous* volatility contributes to an emerging literature on HFT and volatility. Brogaard, Hendershott, and Riordan (2014) document that HFTs participate more on high volatility days and generally trade in the direction of public news (e.g., macro news, market returns, or order book imbalances). Jovanovic and Menkveld (2011) show that HFTs participate more on days when a stock's volatility largely reflects market-wide volatility (i.e., low idiosyncratic volatility). Our evidence is consistent with theirs. Existing evidence on the relation between HFT activity and the size of endogenous (short-term) pricing errors is mixed. For example, Boehmer, Fong, and Wu (2014), Egginton, Ness, and Ness (2013), and Hasbrouck (2013) suggest a positive relation, whereas Brogaard, Hendershott, and Riordan (2014) and Hagströmer and Nordén (2013) suggest the opposite.

2   A Pecking Order Hypothesis of Trading Venues

In this section we further motivate the pecking order hypothesis in its specific form (Figure 2(b)). It is this specific-form hypothesis that will be the main hypothesis discussed and tested throughout the remainder of the paper. First, we discuss how various dark pool papers could be interpreted as suggestive of the predicted ordering in venues. Second, we develop an empirical strategy to test the hypothesis.

2.1   The pecking order hypothesis in its specific form

Panel (b) of Figure 2 illustrates the specific form of the pecking order hypothesis: dark pools that cross orders at the midpoint of NBBO (labeled "DarkMid") are on the top of the pecking order, dark pools that allow some price flexibility (labeled "DarkNMid") are in the middle, and transparent venues (labeled "Lit") are at the bottom.

DarkNMid is a "lighter shade of dark," sitting between DarkMid and Lit. In DarkNMid an order imbalance can move the price so that the dark pool operator can create more matched volume and

the trade-through restriction guarantees that investors still get a price improvement relative to Lit. In other words, in DarkNMid execution is more likely than in DarkMid, but earning the full half-spread is less likely. If an investor moves from DarkNMid to Lit, execution becomes guaranteed but he loses the ability to earn any of the spread at all. In Section 6 we propose a simple candidate model that characterizes the competition among the three types of venues: DarkMid, DarkNMid and Lit. The model and its analysis formalize the intuition underpinning the pecking order hypothesis of trading venues.

The specific form of the pecking order hypothesis is richer than existing theories of dark pools. For example, models of Hendershott and Mendelson (2000), Degryse, Van Achter, and Wuyts (2009), Buti, Rindi, and Werner (2014), Ye (2011), and Zhu (2014) all have the feature that midpoint dark pools provides potential price improvement relative to exchange prices, but they do not guarantee the execution of investors' orders. In all these models, investors have access to all venues and actively choose the best type of venue that serves their speculative or hedging needs.[6] While these models predict that dark venues are near the top of the pecking order and lit venues near the bottom, they do not distinguish different types of dark pools.

There is an alternative motivation for the pecking order hypothesis, based on an agency conflict between investors and their brokers. Its starting point is that brokers decide where to route investors' orders, and investors monitor brokers insufficiently. If brokers earn more profits by routing investors' orders first to their own dark pools, then orders would first flow into broker-operated dark pools and then to other dark and lit venues. Only when investors tell brokers to execute quickly will they have no choice but to send it to lit venues. This alternative motivation is less rich as it does not suggest a particular ranking of dark venue types. It simply states that dark venues take priority, whatever type of dark venue the broker is running, except when investors emphasize quick execution.

---

[6] Yet another motivation for trading in dark venues is suggested by Boulatov and George (2013). They show that if liquidity providers are informed, displaying their orders reduces their information rents and discourages informed traders from providing liquidity. As a result, an opaque, or dark, market can have a narrower bid-ask spread and more efficient prices than a transparent one. Different from the dark pool studies, however, the model of Boulatov and George (2013) has only one market at a time and does not make predictions regarding the fragmentation of dark and lit venues.

2.2 Empirical strategy for testing the pecking order hypothesis

Because venues are sorted by the trade-off between immediacy and trading cost, an upward shock to investors' urgency to trade should predict a loss of market share by venues at the top of the pecking order and a gain of market share by venues at the bottom. Testing the pecking order hypothesis requires (i) disaggregated data on dark volumes and (ii) an exogenous shock to demands for immediacy and an econometric model.

Section 3 discusses the high-frequency disaggregated dark volume data available to us. Our empirical strategy is to use the high-frequency nature of the data to identify the dynamic interrelation among dark volumes in the various categories. We use two types of exogenous shocks as instruments for urgency: shocks to VIX and shocks to earnings.

We first study trade response to a market-wide volatility shock implemented through a doubling of VIX. Because VIX is calculated from prices of short-dated (within one month maturity) S&P 500 index options, we are reasonably confident that fragmentation of order flows is unlikely to change VIX. Rather, VIX is driven mostly by systematic risks and the time-varying risk premium in the aggregate U.S. equity market, both of which can increase the effective cost of holding an undesirable inventory.

Next, we study trade response to an idiosyncratic shock to firms' fundamental values, measured by earnings surprises on announcement dates relative to pre-announcement expectations. Since earnings announcements are scheduled in advance and are made outside trading hours, they are not affected by trading activities. Trades naturally happen given the new information from the announcements. For example, investors may wish to liquidate speculative or hedging positions established before earnings announcements. The larger is the magnitude of an earnings surprise (positive or negative), the more quickly losses accumulate for investors who "made the wrong bet," and the more urgently they want to close out these losing positions (e.g., due to margin constrain). Trades can also be generated when investors interpret the same earnings news differently (Kim and Verrecchia, 1994). The larger the earnings surprise, the more immediacy we expect investors to demand to adjust their positions in the

10

stock.

The next two sections discuss the data and the econometric approach in full detail. Section 5 discusses the test results.

3   Data

Our data sample covers 117 stocks in October 2010 (21 business days). It consists of large-cap, medium-cap, and small-cap stocks in almost equal proportions.[7] In addition to TAQ trading volumes and National Best Bid and Offer, we use two main proprietary data sources: (1) transactions in five types of dark venues, (2) HFT activity in the NASDAQ main market. We complement these trading data with intraday VIX and earnings announcements by firms in our sample. These data sources are described below in detail.

3.1   Dark volumes

In the United States, off-exchange transactions in all dark venues are reported to trade-reporting facilities (TRFs). The exact venue in which the dark trade takes place is not reported in public data. (Recently, FINRA started to publish weekly transaction volumes in alternative trading systems, but these volumes are aggregated and not on a trade-by-trade basis.[8]) The NASDAQ TRF is the largest TRF, accounting for about 92% of all off-exchange volumes in our sample. Our first data source is the dark transactions reported to NASDAQ TRF. These trades are executed with limited pre-trade transparency.[9] A salient feature of our data is that the dark transactions are disaggregated into five cat-

---

[7] This sample of stocks is the same as that in the NASDAQ HFT dataset used by many studies of HFT. The origin sample compiled by NASDAQ contains 120 stocks, with 40 stocks in each size category. But only 117 of the 120 stocks are present in our sample period. Brogaard, Hendershott, and Riordan (2014) state that "The top 40 stocks are composed of 40 of the largest market capitalization stocks... The medium-size category consists of stocks around the 1000th largest stocks in the Russell 3000..., and the small-size sample category contains stocks around the 2000th largest stock in the Russell 3000."

[8] See https://www.finra.org/Newsroom/NewsReleases/2014/P519139.

[9] Our dark transaction data do not include trades on electronic communication networks (ECNs). ECNs are transparent venues that register as alternative trading systems (ATS), but they are not exchanges for regulatory purposes. In our sample, ECNs account for a very small fraction of total transaction volume. BATS and DirectEdge, two major exchanges

egories by trading mechanism. The exact method of such disaggregation is proprietary to NASDAQ, but we know their generic features. The five categories include:

1. **DarkMid.** These trades are done in dark pools that use midpoint crossing as much as possible. Midpoint crossing means the buyer and the seller in the dark venue transact at the midpoint of the National Best Bid and Offer (NBBO). "Agency-only" dark pools (i.e. those with no proprietary order flows) typically operate this way.[10]

2. **DarkNMid.** These trades are done in dark pools that allow flexibility in execution prices (not necessarily midpoint). By this features, we infer that dark pools operated by major investment banks belong to this category.

3. **DarkRetail.** These trades are internalized volume by broker-dealers. Retail brokers often route order flows submitted by retail investors to major broker-dealers, who then fill these orders as principal or agent. These transactions would be classified as DarkRetail.

4. **DarkPrintB.** These trades are "average-price" trades. A typical example is that an institutional investor agrees to buy 20,000 shares from a broker, at a volume-weighted average price plus a spread. This trade of 20,000 shares between the investor and the broker would be classified as a "print back" trade, abbreviated as "PrintB."

5. **DarkOther.** These are other dark trades not covered by the categories above. A typical example in this category would be a negotiated trade between two institutions on the phone (i.e., not done on any electronic platform).

We emphasize that each category is not a single trading venue, but a collection of venues that are qualitatively similar in terms of their trading mechanisms. In the interest of brevity, however, we will use the terms "venue" and "type of venue" interchangeably.

---

that recently merged, used to be ECNs, but they have converted to full exchange status in November 2008 and July 2010, respectively.

[10] The NASDAQ classification puts a couple of midpoint-crossing dark pools that have large transaction sizes into "DarkOther" rather than "DarkMid," in order to guarantee that one cannot reverse-engineer the identities of these venues.

Figure 3: Volume shares of dark venues and aggregate lit venue. This chart illustrates the average volume share of the various types of dark pools in the data sample.
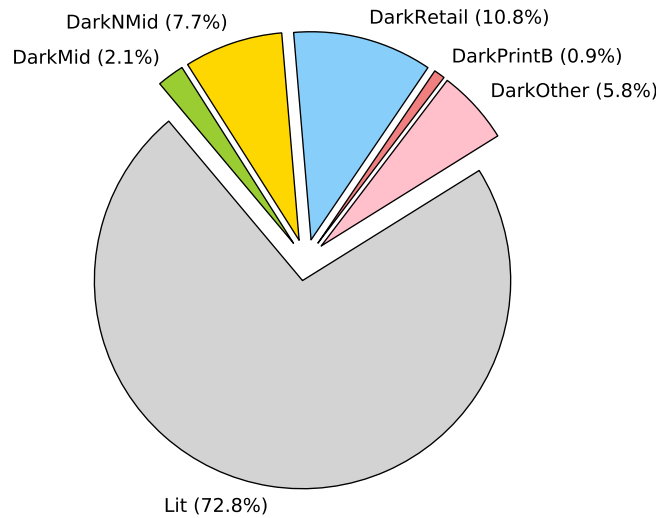

Figure 3 shows the market shares of the five types of dark venues as a fraction of total trading volume (obtained from TAQ) in our sample. We label the complement of these five dark venues as the "lit" venues. The "lit" label is an approximation, but it is a good one.[11] We observe that dark venues account for 27.2% of total transaction volume in the 117 stocks in October 2010. Ranked by market shares, the five dark categories are DarkRetail (10.8%), DarkNMid (7.7%), DarkOther (5.8%), DarkMid (2.1%), and DarkPrintB (0.9%).

It is informative to compare our five-way categorization of dark trading venues to that of the SEC. SEC (2010) classifies opaque trading centers into dark pools and broker-dealer internalization. If an approximate correspondence is to be made, our DarkMid and DarkNMid types roughly fall into the SEC's dark pool category, and our DarkRetail, DarkPrintB, and DarkOther types roughly fall into the SEC's broker-dealer internalization category.

Using one week of FINRA audit trail data in 2012, Tuttle (2014) reports that about 12.0% of trading volumes in U.S. equities are executed in off-exchange alternative trading systems (ATS), the

---

[11] Since NASDAQ TRF accounts for 92% of all off-exchange trading volumes in our sample, the "lit" category also contains the remaining 8% of off-exchange volumes, or about 2.4% of total volumes. For our purchase of characterizing the qualitative nature of venue types, this measurement noise is likely inconsequential.

majority of which are dark pools; and that about 18.8% of U.S. equity volume are executed off exchanges without involving ATS, which can be viewed as a proxy for trades intermediated by broker-dealers. The dark-pool volumes and internalized volumes implied by our dataset are comparable to those reported by Tuttle (2014).

3.2 High-frequency trading and quoting activity in NASDAQ market

Our second data source contains the trading and quoting activities of high-frequency trading firms in the NASDAQ market. These firms are known to employ high-speed computerized trading, but their identities and their strategies are unknown to us. This dataset has two parts.

First, for each transaction on NASDAQ, we observe the stock ticker, transaction price, number of shares traded, an indicator of whether the buyer or the seller is the active side, and an indicator on whether either side of the trade is an HFT firm. We refer to a trade as an HFT trade if at least one side of the trade is an HFT firm. All transactions are time-stamped to milliseconds. Second, we observe the minute-by-minute snapshot of the NASDAQ limit order book. For each limit order, we observe the ticker, quantity, price, direction (buy or sell), a flag on whether the order is displayed or hidden, and a flag on whether or not the order is submitted by a high-frequency trading firm.

An important and well-recognized caveat of the NASDAQ HFT dataset is that it excludes trades by HFT firms that are routed through large, integrated brokers. Nor does the dataset distinguish various HFT strategies, be it market making or "front-running." Given this caveat, a prudent way to interpret the HFT measures is that they are "control variables," needed to capture market conditions. For additional details about the NASDAQ HFT dataset, see Brogaard, Hendershott, and Riordan (2014).

3.3 VIX and earnings announcements

The last components of our data include two measures of shocks: VIX and earnings announcements. The Chicago Board of Exchange (CBOE) disseminates VIX every 15 seconds. Minute-by-minute VIX data are obtained from pitrading.com. Among the 117 firms in our sample, 68 firms announced

Figure 4: Time series of dark volume share and VIX. This chart illustrates the average volume share of the five types of dark venues combined (shaded area, left axis) and VIX (solid line, right axis). All series are aggregated at the daily level and averaged across all stocks.
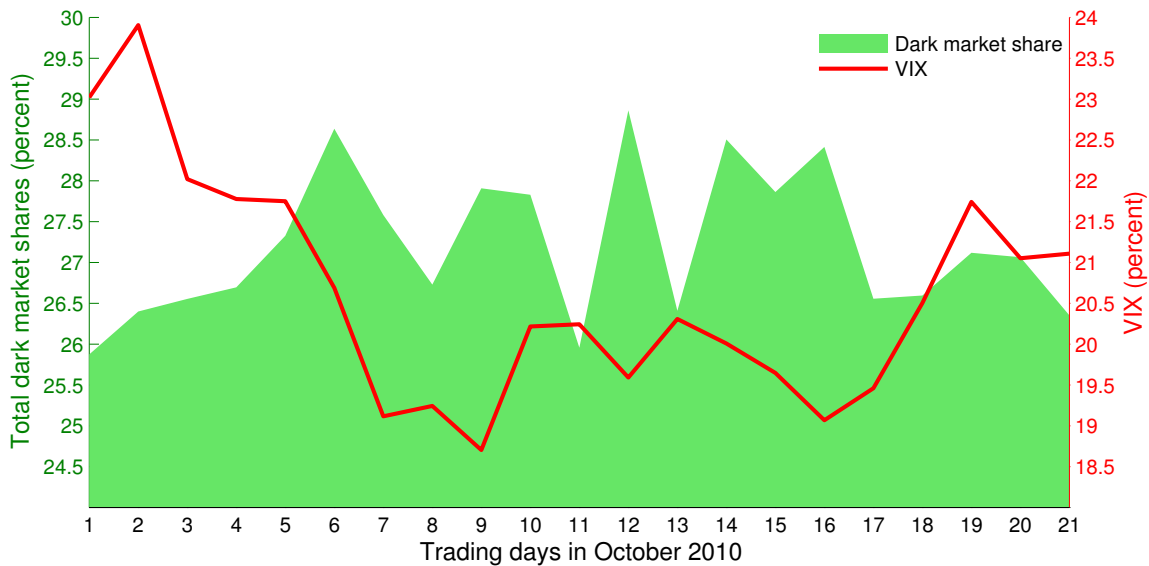
earnings in October 2010. For each of these earnings announcements, we download the announcement dates, time stamps, announced earnings per share, and expected earnings per share from Bloomberg. (We are able to collect EPS forecast from Bloomberg for 67 of the firms, and hence can construct the *EpsSurprise* variable defined below.)

Figure 4 shows the time series of the volume shares of the five types of dark venues and VIX. They are aggregated at the daily level and averaged across the 117 stocks. One interesting observation is that in the time series, a higher VIX is associated with a lower dark market share. In the next few sections we will further explore VIX shocks at a higher frequency to test the pecking order hypothesis.

4 A VARX Model of Dark Volumes

In this section, we characterize the dynamic interrelation among dark volumes, HFT participation, and various standard trade variables through a panel vector autoregressive model with exogenous variables (a panel VARX).

15

4.1 Model components

From the raw data we calculate three broad types of variables for our analysis: (1) transaction volumes in five types of dark venues, (2) characteristics of the NASDAQ lit market, and (3) overall market conditions. We describe these components below. For ease of reference, these variables and their descriptions are tabulated in Table 1. All variables are constructed from the raw data at the minute frequency during the normal trading hours (9:30am to 4:00pm). This gives us a stock-minute panel data with $117 \times 21 \times 390 = 958,230$ observations.

Table 1: Variable descriptions. This table lists and describes all variables used in this study. All variables are generated for one-minute intervals. Variables that enter the econometric model (Section 4) are underscored. The subscript j indexes stocks; t indexes minutes. Type "Y" and "Z" are described in the panel VARX model.

Type	Variable Name	Description
Panel A: Dark venue trading volumes		
Y	$VDarkMid_{jt}$	Volume of midpoint-cross dark pools
	$VDarkNMid_{jt}$	Volume of non-midpoint dark pools
	$VDarkRetail_{jt}$	Volume of retail flow internalization
	$VDarkPrintB_{jt}$	Volume of average-price trades ("print back")
	$VDarkOther_{jt}$	Volume of other dark venues
	$VLit_{jt}$	Total volume minus all dark volume
Panel B: NASDAQ lit market characterization		
	$BASpread_{jt}$	NASDAQ lit market bid-ask spread divided by the NBBO midpoint
	$TopDepth_{jt}$	Sum of NASDAQ visible best bid depth and best ask depth
	$HFTinTopDepth_{jt}$	$Depth_{jt}$ based on only HFT limit orders divided by $Depth_{jt}$
	$HFTinVolume_{jt}$	NASDAQ lit volume in which HFT participates divided by total NASDAQ lit volume
Panel C: Overall market conditions		
	$TAQVolume_{jt}$	TAQ volume
	$RealVar_{jt}$	Realized variance, i.e., sum of one-second squared NBBO midquote returns
	$VarRat10S_{jt}$	Variance ratio, i.e., ratio of realized variance based on ten-second returns relative to realized variance based on one-second returns (defined to be one for a minute with only one-second returns that equal zero)
Z	VIX_t	One-month volatility of S&P500 index (in annualized percentage points)
	$EpsSurprise_{jt}$	Surprises in announced EPS, calculated as the absolute difference in announced EPS and the forecast EPS, scaled in share price: |announced EPS − forecast EPS|/(closing price on the day before). This variable is zero if a stock-day pair does not have a earnings announcement.

Dark volumes. We separately calculate trading volumes across the five types of dark venues: *VDark-Mid*, *VDarkNMid*, *VDarkRetail*, *VDarkPrintB*, and *VDarkOther*. Volume is measured in thousands of shares, which we prefer over dollar value as it ensures that quantity variables are not contaminated by price effects. For example, two consecutive 1000 share transactions are considered to be of equal size, even though they may be executed at slightly different prices.

Characteristics of the NASDAQ lit market. We use two commonly-used measures of liquidity: spread and depth. *BASpread* is the relative bid-ask spread of the NASDAQ lit market, measured in basis points. *TopDepth* is the sum of visible liquidity supply on the bid side and the ask side of the NASDAQ limit order book, measured in thousands of shares. The motivation for using the number of shares (rather than market value) to measure depth is the same as that for transaction volumes, as discussed earlier.

In addition, we use two measures of HFT activity: *HFTinVolume* and *HFTinTopDepth*. *HFTinVolume* is the number of shares traded by HFT (on at least one side) on the NASDAQ market divided by the number of shares traded on the NASDAQ market.[12] *HFTinTopDepth* is the number of shares posted (quoted) by HFT on the top of the NASDAQ limit order book divided by the total depth on the top of the NASDAQ limit order book.

Overall market conditions. We add endogenous and exogenous variables describing the overall market conditions. For each stock and each minute, the endogenous variables are total transaction volume *TAQVolume*, realized return variance within the minute *RealVar*, and the variance ratio within the minute *VarRat10S*. The first variable is based on TAQ data, whereas the other two are based on millisecond-level NBBO data provided by NASDAQ.

Two exogenous variables are included: the one-month volatility on the S&P 500 index, *VIX*, and earnings surprise, *EpsSurprise*. VIX is based on option market prices and therefore is a forward-

---

[12] The stock-day-minute average of *HFTinVolume* is around 40% in our sample period (see table 2). If HFT volumes are first aggregated across days and stocks, then the total HFT volume accounts for about 45% of the total NASDAQ volume in our sample.

looking volatility measure.

There are 68 earnings announcements, 67 of which can be matched with a forecast, in our sample. Consistent with the accounting literature (e.g. Kinney, Burgsthler, and Martin, 2002), the earnings surprise is calculated as the absolute difference between announced EPS and pre-announcement expected EPS, then divided by the closing price on the business day immediately before the announcement. If stock j announced its earnings on a particular day, we set the variable $EpsSurprise_{jt}$ to this surprise measure for stock j and all minutes t in the business day immediately following the earnings announcement.[13] Otherwise, $EpsSurprise_{jt}$ is set to zero. This variable captures the presence of material news for the firm's fundamental as well as the magnitude of the surprise. Earnings announcements are typically scheduled months or years ahead by company management; therefore, they are exogenous.

4.2 Data preparation and summary statistics

We convert all variables into logs, except the earnings surprise. A log-linear model has a couple of advantages over a linear model. First, a log-linear model comes with a natural interpretation that estimated coefficients are elasticities. Second, all endogenous variables (e.g., volume, realized variance, and depth) are guaranteed to remain nonnegative. In other words, the error term does not need to be bounded from below, which would be the case for a linear model.

In order to take the log, data need to be winsorized to eliminate the zeros. We use the following procedure. If a particular dark venue has a zero transaction volume for stock j and minute t, its volume for that stock-minute is reset to one share. If a particular stock j does not trade in period t on the NASDAQ market, the *HFTinVolume* variable is undefined. In this case, to not lose data, *HFTinVolume* is forward filled from the start of the day. The motivation is that market participants may learn HFT activity on NASDAQ by carefully parsing market conditions. If there is no update in a particular time interval, they might rely on the last observed value they learned about. Zero entries

---

[13] That is, if the earnings announcement is made before the market opens (9:30am), the immediate following business day is the same business day. If the earnings announcement is made after the market closes (4pm), the immediate following business day is the next business day.

for all other variables are left-winsorized at the 0.01% level.

Table 2 reports the summary statistics of the model variables, before taking the log. One important observation is that the data preparation procedure discussed above has a minimum effect on the raw data. Total trading volume per stock-minute is about 12,300 shares on average. The market shares of the five types of dark venues are shown in Figure 3. (For completeness, we also include *VLit*, defined as *TAQVolume* less the sum of the five dark volumes, although *VLit* will not be part of the panel VARX model.) On average, HFT participation in quoting at the best NASDAQ prices, *HFT-inTopDepth*, is about 35%, whereas HFT participation in transactions on NASDAQ, *HFTinVolume*, is about 40%. The average depth and relative spread on NASDAQ limit order book are about 4440 shares and 16.8 basis points, respectively. VIX in our sample has an average of 20.6%. The earnings surprises measure, *EpsSurprise*, has an average of about 47 basis points across the 68 firms that made earnings announcements in our sample.[14]

4.3 Panel VARX model

The panel VARX model used for the main empirical analysis has the following form:

$$y_{jt} = \alpha_j + \Phi_1 y_{j,t-1} + \cdots + \Phi_p y_{j,t-p} + \Psi_1 z_{j,t-1} + \cdots + \Psi_r z_{j,t-r} + \varepsilon_{jt}, \tag{1}$$

where, for stock j and period t:

y_{jt} is a vector of log-transformed endogenous variables. Our baseline model includes the following variables, with detailed explained in Table 1:

- Volumes in the five types of dark venues;

- Other endogenous variables that characterize trade conditions (e.g., volatility, the bid-ask spread, etc.).

---

[14] The statistics of *EpsSurprise* reported in Table 2 are across all $117 \times 21 \times 360$ stock-day-minutes. It should be re-minded that there are only 67 *EpsSurprises* during our sample period. The sample average based on these 67 observations is about 0.15% and the sample standard deviation is 0.89%. Based on these numbers, we give a 1% *EpsSurprise* shock in the pecking order analysis below (see Figure 6).

Table 2: Summary statistics. This table reports summary statistics of the variables used throughout the paper. These statistics are calculated for the raw data as well as the prepared data, before taking the logarithm. The sample frequency is minute. The units of each series is in the square brackets.

	mean		std. dev.		skewness		min		max		zeros		count	
	raw	prep/d	raw	prep/d	raw	prep/d	raw	prep/d	raw	prep/d	raw	prep/d	raw	prep/d
VDarkMid [1k shares]	0.259	0.259	3.161	3.161	377.994	378.016	0	0.001	2032.340	2032.340	737273	0	954892	954892
VDarkNMid [1k shares]	0.941	0.941	5.437	5.437	54.600	54.602	0	0.001	1040.352	1040.352	574616	0	954892	954892
VDarkRetail [1k shares]	1.329	1.330	6.197	6.196	34.609	34.611	0	0.001	1641.229	1641.229	592060	0	954892	954892
VDarkPrintB [1k shares]	0.106	0.107	8.040	8.040	269.541	269.542	0	0.001	3524.920	3524.920	924829	0	954892	954892
VDarkOther [1k shares]	0.709	0.709	10.798	10.798	104.222	104.223	0	0.001	3018.263	3018.263	690408	0	954892	954892
VDark [1k shares]	3.437	3.438	41.120	41.120	414.272	414.273	0	0.001	23616.586	23616.586	440491	0	958230	958230
VLit [1k shares]	8.932	8.932	45.397	45.397	60.567	60.568	0	0.001	11892.866	11892.866	233599	0	958230	958230
RelSpread [bps]	16.776	16.776	28.466	28.466	37.114	37.114	0.162	0.334	6681.449	6681.449	0	0	957769	957769
TopDepth [1k shares]	4.443	4.443	13.909	13.909	6.058	6.058	0	0.004	502.515	502.515	9	0	958230	958230
HFTinTopDepth [percent]	35.199	35.225	29.512	29.480	0.384	0.386	0	0.105	100	100	244386	0	958221	958230
TAQVolume [1k shares]	12.264	12.268	55.722	55.722	40.816	40.817	0	0.001	12196.768	12196.769	208687	0	958230	958230
HFTinVlm[percent]	39.985	37.157	23.939	26.038	0.053	0.166	0	0.125	100	100	72830	0	583561	955059
RealVar [bps]	6.139	6.188	9.112	9.079	14.518	14.668	0	0.172	1367.197	1367.197	274466	0	958218	958218
VarRat10S [percent]	99.772	99.778	44.438	44.426	1.142	1.144	0	0.317	853.407	853.407	16561	0	958218	958218
VIX [percent]	20.626	20.626	1.369	1.369	0.656	0.656	17.940	17.940	24.330	24.330	0	0	954603	954603
EpsSurprise [percent]	0.013	0.013	0.150	0.150	18.871	18.871	0	0	4.190	4.190	931710	931710	957840	957840

z_{jt} is a vector of arguably exogenous variables, including the log transformations of VIX_{jt} and the raw (not logged) earnings surprise measure $EpsSurprise_{jt}$.

A stock fixed effect α_j ensures that only time variation is captured, not cross-sectional variation, as our focus is on dynamic interrelations among variables. The number of lags (in minutes) is determined based on the Bayesian Information Criterion (BIC): $p = 2$ and $r = 1$. Further details on the implementation of the estimation are provided in Appendix A.

4.4   Estimation results

Table 3 reports the estimation results of the VARX model. Panel (a) reports the estimated coefficients $\{\Phi_1, \Phi_2, \Psi_1\}$, which should be interpreted as elasticities, and Panel (b) reports the correlations of the residuals. We make a few observations. First, a higher VIX or a higher earnings surprise forecasts higher volumes in dark venues and total TAQ volume, but the elasticity of TAQ volume is higher than that of dark volumes. This suggests that the market shares of various venue types are likely to have heterogeneous reactions to VIX and earnings shocks. Second, a higher HFT participation in this minute tends to forecast a mildly lower volumes in all venues in the next minute. Third, dark volumes in different types of venues have heterogeneous dynamic correlations with spreads, depths, and variance ratios, sometimes with opposite signs (see the first five lines of Panel (a)).

In addition to the parameter estimates, we also calculate and plot standard impulse-response functions (IRFs) of endogenous variables to shocks. With few exceptions, we set the magnitudes of the initial shocks such that the shocked variable doubles in its level (i.e. shock the logged variable by $\log(2)$). For a generic upward shock to some variable in y or z, the IRF essentially identifies two types of elasticities:

- an immediate one-period response of some variable in y, and

- a cumulative long-term response of some variable in y.

IRFs make the dynamic interrelations among variables transparent. Besides its intuitive appeal,

21

another advantage of IRFs is that they show the duration of the responses. Thus, we will mostly rely on IRFs for exposition in the remaining of the paper. As the IRF is a non-linear function of parameter estimates, we calculate the 95% confidence bounds of the IRF through simulations. In each iteration a parameter value is drawn from a normal distribution with a mean equal to the point estimate and a covariance matrix equal to the estimated parameter covariance matrix. Details of this simulation method are provided in Appendix A. In the next two sections, we present the IRFs that are most directly related to our main findings.

5   Results: Pecking Order

In this section we test the pecking order hypothesis laid out in Section 2.

5.1   Testing the pecking order hypothesis

We focus on testing the specific form of the pecking order hypothesis (Figure 2(b)) because it makes stronger (i.e. more specific) predictions. Evidence supporting the specific form would also support the generic form. For completeness, after presenting results for DarkMid, DarkNMid, and Lit, we also discuss the other three dark venue types (DarkRetail, DarkPrintB, and DarkOther).

A VIX shock

Starting with our estimated VARX model and the steady state, we shock VIX by 100% upward (i.e. shock $\ln(VIX)$ by $\ln(2)$) and examine the resulting changes in volume shares of DarkMid, DarkNMid, and Lit in the following minutes.[15] Their volume shares are denoted as *SDarkMid*, *SDarkNMid*, and

---

[15] While the VARX model is written in terms of volumes, the calculation of market shares is straightforward from the estimated coefficients. Specifically, following the VIX shock at the steady state, the VARX model we already estimated spells out the future paths of all transaction volumes in the next 1, 2, 3, … minutes. From these volumes we calculate the market shares and associated confidence bound by simulation. The steady state levels of the market shares are defined as the average of all the minute-day-stock observations. The steady-state market shares calculated this way are slightly different from those shown in Figure 3, in which the market share of each type is computed as the sum of volume in that venue type divided by the sum of TAQ volume during the entire sample period.

Table 3: VARX Estimation. This table reports the estimation result of the VARX model at the minute frequency. All variables are log-transformed except the earnings announcement surprise *EpsSurprise*. Panel (a) reports the estimated coefficients. Panel (b) reports the correlations of the estimated residuals. In Panel (b) we also included *VDark*, defined as the sum of the volumes in five types of dark venues, and *VLit*, defined as *TAQVolume* less *VDark*.

(a) Estimated coefficients

	VDarkMid	VDarkNMid	VDarkRetail	VDarkPrintB	VDarkOther	RelSpread	TopDepth	HFTinTopDepth	TAQVolume	HFTinVlm	RealVar	VarRat10S
Endogenous variables: 1 minute lag												
VDarkMid (-1)	0.213**	0.037**	0.030**	0.005**	0.041**	-0.001*	0.002**	-0.000	0.019**	0.001	0.005**	-0.001*
VDarkNMid (-1)	0.033**	0.196**	0.053**	0.004**	0.049**	-0.000	0.005**	0.001	0.042**	0.000	0.003*	-0.002**
VDarkRetail (-1)	0.020**	0.035**	0.112**	0.002**	0.040**	0.001*	-0.001	-0.001	0.020**	0.003**	0.007**	0.002**
VDarkPrintB (-1)	0.009*	0.007	0.009*	0.054**	0.018**	-0.000	0.001	0.002	0.003	0.000	0.003	0.000
VDarkOther (-1)	0.021**	0.036**	0.038**	0.002**	0.146**	0.000	-0.001**	-0.000	0.019**	-0.002*	0.006**	-0.001
RelSpread (-1)	-0.072**	-0.089**	-0.038**	-0.003	-0.061**	0.409**	-0.066**	-0.100**	-0.368**	-0.116**	0.233**	-0.039**
TopDepth (-1)	0.104**	0.170**	0.077**	0.006	0.082**	-0.040**	0.391**	-0.028*	0.158**	-0.011*	-0.134**	0.014**
HFTinTopDepth (-1)	-0.001	-0.001	-0.003*	-0.000	-0.004**	-0.004**	-0.002*	0.336**	-0.002	0.044**	0.001	0.002**
TAQVolume (-1)	0.015**	0.051**	0.047**	0.001*	0.032**	-0.003**	0.004**	-0.005	0.213**	-0.006*	0.033**	0.002
HFTinVlm (-1)	-0.003*	-0.005**	-0.002	-0.000	-0.007**	0.000	-0.000	0.004*	-0.004**	0.546**	0.001	0.001
RealVar (-1)	-0.025**	-0.046**	-0.013**	-0.000	-0.033**	0.010**	-0.016**	0.005	-0.037**	0.010**	0.234**	-0.022**
VarRat10S (-1)	0.004	0.001	0.018**	0.002	0.005	0.004**	-0.002**	0.001	0.013**	0.007**	0.023**	0.048**
Endogenous variables: 2 minutes lag												
VDarkMid (-2)	0.179**	0.024**	0.018**	0.003**	0.029**	-0.000	0.003**	0.001	0.011**	-0.001	0.004**	0.000
VDarkNMid (-2)	0.021**	0.143**	0.037**	0.002*	0.033**	0.001**	0.005**	0.001	0.026**	-0.001	0.002	-0.000
VDarkRetail (-2)	0.015**	0.026**	0.092**	0.001	0.031**	0.002**	-0.001*	-0.001	0.018**	0.002*	0.012**	0.002**
VDarkPrintB (-2)	0.009*	0.005	0.010*	0.031**	0.006	-0.000	0.001	-0.000	-0.001	-0.001	-0.001	-0.000
VDarkOther (-2)	0.010**	0.022**	0.031**	0.000	0.113**	0.000	-0.001**	-0.002	0.010**	-0.000	0.003**	0.000
RelSpread (-2)	0.021**	0.065**	0.066**	-0.006*	0.075**	0.223**	-0.035**	-0.101**	0.073**	0.010	0.113**	0.014**
TopDepth (-2)	0.009	0.039**	-0.011	0.005	-0.043**	-0.018**	0.235**	-0.001	-0.010	-0.014**	-0.102**	0.003
HFTinTopDepth (-2)	-0.001	-0.002	-0.004**	0.000	-0.004**	-0.005**	-0.000	0.171**	-0.008**	0.007**	-0.004**	-0.000
TAQVolume (-2)	0.001	0.030**	0.029**	0.003**	0.004*	-0.010**	0.010**	-0.007**	0.144**	-0.003	-0.009**	0.002*
HFTinVlm (-2)	0.002	0.002	-0.001	-0.000	0.002	-0.001**	0.000	0.018**	0.001	0.116**	0.005**	0.001
RealVar (-2)	-0.003	-0.014**	0.005*	-0.002*	-0.001	0.009**	-0.011**	0.000	0.002	0.005*	0.151**	-0.010**
VarRat10S (-2)	-0.001	-0.007*	0.004	0.001	-0.003	0.002**	0.001	0.002	-0.006*	0.005*	-0.001	0.030**

Table 3 continued...

(a) Estimated coefficients (continued)

	VDarkMid	VDarkNMid	VDarkRetail	VDarkPrintB	VDarkOther	RelSpread	TopDepth	HFTinTopDepth	TAQVolume	HFTinVlm	RealVar	VarRat10S
					Exogenous variables							
VIX (-1)	-0.142	0.268*	-0.401**	0.155**	0.353**	0.046**	0.009	-0.107	0.785**	0.176**	0.813**	-0.001
EpsSurprise	0.077*	0.184*	0.221**	0.086*	0.115*	0.005	-0.000	0.071	0.293**	0.035	0.079**	0.006
R^2	0.130	0.138	0.070	0.005	0.087	0.366	0.353	0.190	0.143	0.398	0.162	0.008
# obs.	937674	937674	937674	937674	937674	937674	937674	937674	937674	937674	937674	937674

*, ** Significant, respectively, at 5%, and 1%. All tests are two sided.

(b) Residual correlations

	VDarkNMid	VDarkRetail	VDarkPrintB	VDarkOther	VDark	VLit	RelSpread	TopDepth	HFTinTopDepth	TAQVolume	HFTinVlm	RealVar	VarRat10S
VDarkMid	0.185**	0.123**	0.010**	0.253**	0.361**	0.143**	0.004**	0.009**	0.010**	0.199**	0.022**	0.124**	-0.015**
VDarkNMid		0.212**	0.011**	0.226**	0.675**	0.227**	0.016**	0.014**	0.014**	0.335**	0.029**	0.171**	-0.014**
VDarkRetail			0.006**	0.161**	0.813**	0.182**	0.015**	0.004**	0.004**	0.318**	0.013**	0.129**	-0.002*
VDarkPrintB				0.009**	0.027**	0.033**	0.001	0.003*	-0.001	0.035**	0.000	0.006**	0.001
VDarkOther					0.460**	0.199**	0.019**	-0.008**	0.007**	0.270**	0.028**	0.176**	-0.011**
VDark						0.288**	0.022**	0.009**	0.011**	0.451**	0.031**	0.220**	-0.012**
VLit							0.059**	-0.027**	0.015**	0.985**	0.065**	0.495**	-0.005**
RelSpread								0.104**	0.150**	0.059**	0.040**	0.125**	-0.004**
TopDepth									0.067**	-0.024**	-0.018**	-0.095**	-0.003**
HFTinTopDepth										0.016**	0.010**	0.025**	0.001
TAQVolume											0.066**	0.501**	-0.007**
HFTinVlm												0.108**	0.001
RealVar													-0.106**

*, ** Significant, respectively, at 5%, and 1%. All tests are two sided.

SLit, respectively. The pecking order hypothesis stated in Section 2 predicts that the elasticities of Lit, DarkNMid, and DarkMid market shares to VIX are positive, mildly negative, and most negative, respectively.

Figure 5 depicts the findings. The ordering of these three venue types conforms to the pecking order hypothesis. In the minute immediately following the VIX shock, *SDarkMid* shows the most negative reaction, falling from the steady state of 2.3% to 1.2%, a 52% reduction. *SDarkNMid* also falls from the steady state of 7.1% to 5.0%, but its loss of market share is a smaller fraction of its steady state counterpart, plus 30% . By sharp contrast, *SLit* increases from 77.5% to 90.0%, a gain of 16%. In all three venue types, the effects on market shares last for about 5 minutes before dying out.

The intuitive sorting of the three venue types shown in Figure 5 is supported by formal econometric tests (a theory that motivates these tests is found in Section 6):

Null 1: The elasticities of *SDarkMid* and *SDarkNMid* to VIX shocks are the same;

Null 2: The elasticities of *SDarkNMid* and *SLit* to VIX shocks are the same.

The test results are reported in Table 4. Both nulls are rejected by the data. The tests show that after the upward VIX shock, the percentage change of *SDarkMid* is more negative than that of *SDarkNMid* at the 1% significance level over the next 5 minutes. The percentage change of *SDarkNMid* is more negative than that of *SLit* at the 1% significance level over the next 4 minutes.

Overall, the evidence from VIX shocks supports the pecking order hypothesis.

An earnings announcement shock

The second shock we exploit is the earnings surprises of individual firms. Starting with the estimated VARX model and the steady state, we impose a 1% earnings surprise and calculate the new steady-state market shares of DarkMid, DarkNMid, and Lit. For each venue type, this level is compared with the steady-state market share on days with no earnings announcements. Again, the pecking order hypothesis predicts that the percentage changes in the market shares should have the following ordering: DarkMid, DarkNMid, and Lit, from the most negative to the most positive.

Figure 5: Pecking order following a shock in VIX. This figure plots the impulse response functions of the market shares of DarkMid, DarkNMid, and Lit following a 100% upward shock to VIX. The impulse responses are plotted for a 15-minute horizon, toward the end of which the market shares of the venues revert to their steady state levels. At each minute, the two-standard-deviation confidence bounds are constructed by simulation. The sequence of the venues is motivated by the specific form of the pecking order hypothesis (see Figure 2(b)).
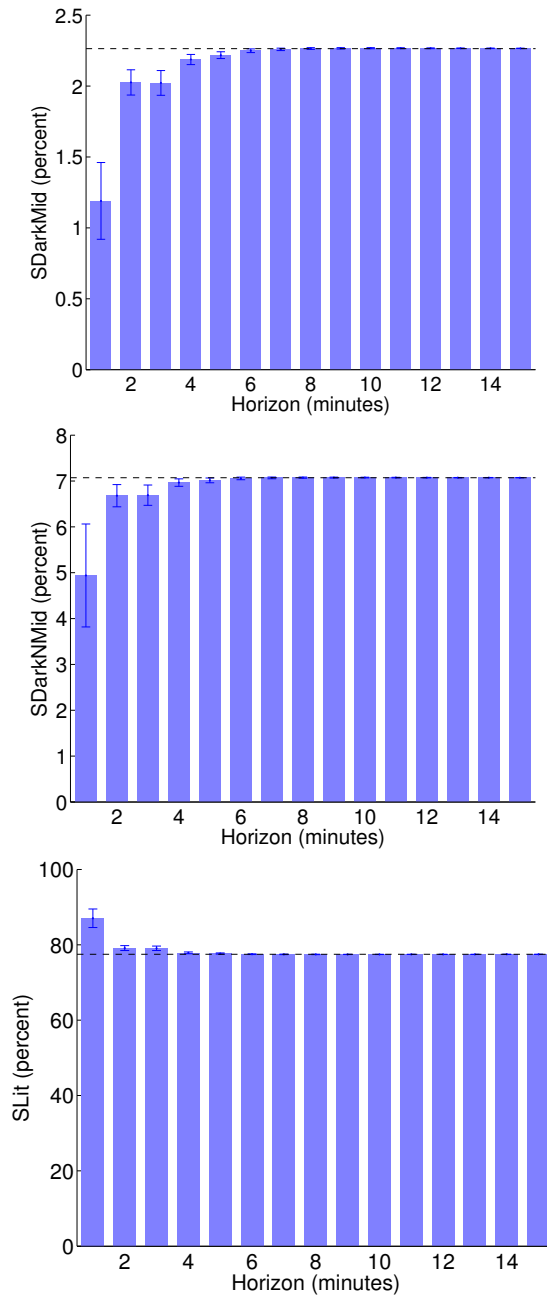
Table 4: Testing the pecking order of trade venues. This table shows the statistical significance of the difference in volume share elasticities (with respect to a VIX shock). Confidence bounds at 99% are constructed, based on 10,000 simulations, for the differences in the volume share changes for the 10 lags (10 minutes) after the VIX shock. If both the upper and the lower bounds are of the same sign, the difference of the elasticities is significantly different from zero. These significant numbers are shaded.

	1 min.	2 min.	3 min.	4 min.	5 min.	6 min.	7 min.	8 min.	9 min.	10 min.
$(\Delta SDMid/SDMid) - (\Delta SDNMid/SDNMid)$										
upper bound (99.5%)	-1.1	-0.7	-1.2	-0.2	-0.1	0.3	0.4	0.4	0.4	0.4
lower bound (0.5%)	-45.7	-13.2	-13.5	-5.2	-3.5	-1.5	-0.9	-0.4	-0.2	-0.1
$(\Delta SDNMid/SDNMid) - (\Delta SLit/SLit)$										
upper bound (99.5%)	-18.7	-2.8	-3.3	-0.2	0.2	0.6	0.7	0.7	0.6	0.5
lower bound (0.5%)	-79.8	-18.1	-17.2	-5.5	-3.2	-1.2	-0.6	-0.3	-0.1	-0.1

Figure 6 plots the results, where the venue types are arranged the same way as in Figure 5. A 1% higher earnings surprise significantly reduces *SDarkMid* by about 50%. While *SDarkNMid* and *SLit* do not show statistically significant changes, the point estimates point to the same direction as in Figure 5. The limited statistical significance here is likely due to the small number of firm-day observations that have earnings announcements (67 out of 117×21). Overall, the results for earnings announcement still support the pecking order hypothesis, although the statistical significance only shows at the DarkMid-DarkNMid step.

5.2 Reactions of DarkRetail, DarkPrintB, and DarkOther to VIX and earnings shocks

The specific form of pecking order hypothesis does not "rank" the remaining three types of dark venues: DarkRetail, DarkPrintB, and DarkOther. This is a conservative choice, for the following reasons. First, because retail investors have little control of their order flows in reality, volumes in the DarkRetail venue type are much more likely driven by the choice of broker-dealers, rather than

Figure 6: Pecking order following a shock in earnings surprises. This figure plots the steady-state values of venue shares for days with no earnings announcements (EA) and the corresponding venue shares following a 1% EPS surprise on an earnings day. The two-standard-deviation confidence bounds are constructed on the shocked market shares by simulation.
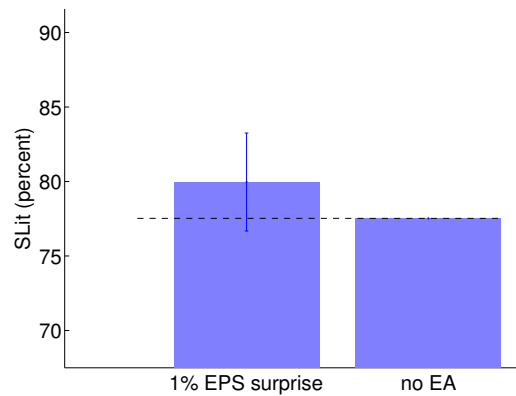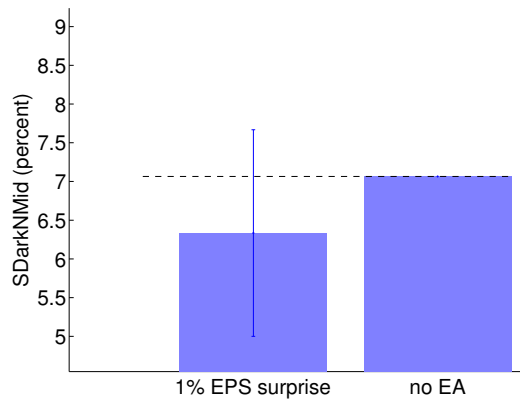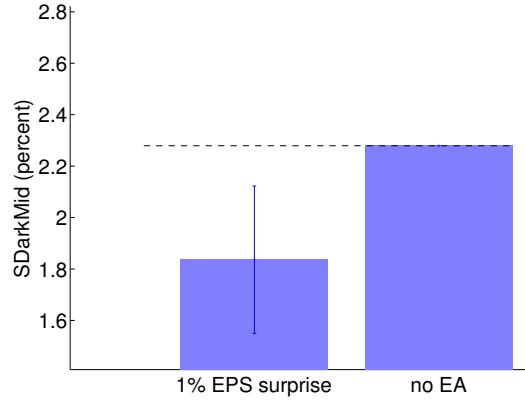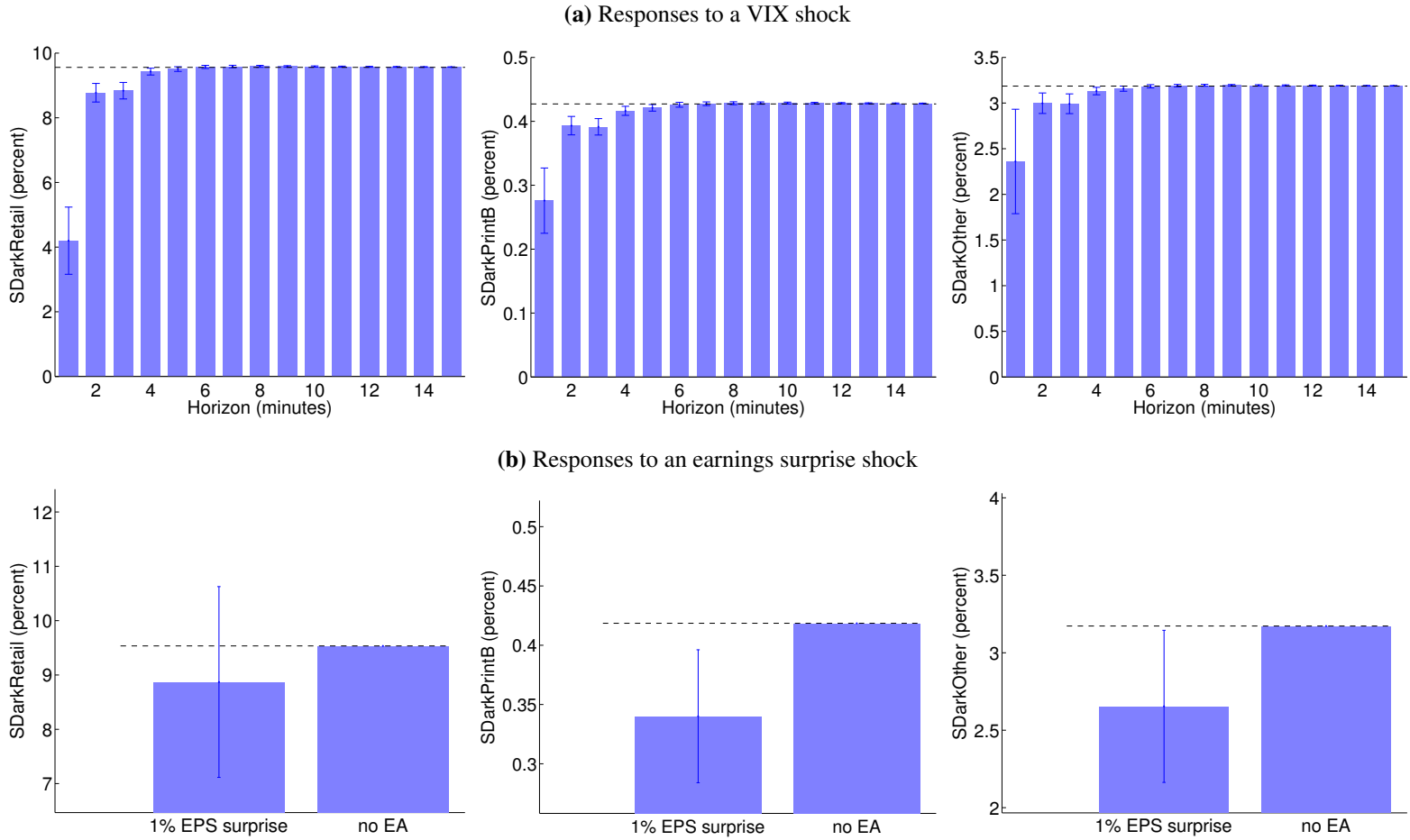
Figure 7: Pecking order of remaining venue types. This figure plots the responses of the market shares of the remaining three dark venue types—DarkRetail, DarkPrintB, and DarkOther—to a VIX shock (Panel (a)) and a earnings surprise shock (Panel (b)).

(a) Responses to a VIX shock

(b) Responses to an earnings surprise shock

investors themselves, to route and to execute these orders off exchange. Second, trades in DarkPrintB may be arranged way before execution takes place, so volumes in the DarkPrintB category may not reflect the current market condition as sensitively as other dark venues. Lastly, the mechanism of DarkOther is not fully transparent to us, as it is the residual category. To be prudent, we have excluded all three venue types from the specific form of pecking order hypothesis.

That said, it is still interesting study the responses of *SDarkRetail*, *SDarkPrintB*, and *SDarkOther* to shocks in VIX and earnings surprises. Figure 7 shows the results. Two salient patterns stand out. On the one hand, market shares in all three types of venues decrease following a positive shock to VIX. This drop is qualitatively similar to that of DarkMid and DarkNMid. In addition, it can also reflect broker-dealers' reluctance in providing liquidity to retail and institutional investors (at attractive enough prices) under a high aggregate volatility. This interpretation is consistent with the finding by Nagel (2012) that the return from liquidity provision in equity markets is highly predicted by VIX. On the other hand, following an earnings shock, the two institution-oriented dark venue types, DarkPrintB and DarkOther, significantly lose market shares, but the retail-oriented venue type, DarkRetail, does not. This difference may reflect the concern of adverse selection by broker-dealers when providing liquidity to institutional investors after firm-specific news, but this concern seems mild or nonexistent when broker-dealers trade against retail investors.

5.3   Venue pecking order and high-frequency trading

We close our empirical analysis by zooming in on high-frequency trading (HFT) and studying how it relates to the pecking order hypothesis. The prevalence of HFTs is commonly cited as one reason why investors are increasingly using dark venues. In the context of the pecking order hypothesis, investors' concern for interacting with HFTs is one form of cost for trading in lit venues.

If concerns regarding HFTs encourage investors to execute orders in the dark, a natural question arises: How does HFT activity affect investors' order flow down the pecking order? A definitive answer requires some exogenous shock that affects HFT activity but not volumes in dark venues.

We do not (yet) have such exogenous variation. Nonetheless, we attempt to provide a partial answer by examining the responses of HFT participation in lit venues to VIX and earnings shocks. More specifically, if HFT participation declines after the VIX and earnings shocks, this pattern would be at least "consistent with" the conjecture that investors trade more in lit venues because they face a lower risk of interacting with HFTs. If the data show the opposite, then this conjecture is less likely. In this exercise we only focus on large-cap stocks (top tercile of our sample), as HFTs are generally most active for actively traded stocks.

Panel (a) of Figure 8 depicts how HFT participation responds to a 100% upward shock to VIX. In the Nasdaq market, HFTs participation in both trading and quoting increase by about 20% after a 100% VIX shock. *HFTinTopDepth* reverts to the steady states rather quickly, in about two minutes, but *HFTinVolume* reverts back to the steady state after about 10 minutes. To the best of our knowledge, the result that HFTs participation increases in VIX is new to the HFT literature.

Panel (b) of Figure 8 depicts the steady-state HFT participation measures, *HFTinTopDepth* and *HFTinVolume*, on days with no earnings announcement and after a one-percentage-point shock to *EpsSurprise*. Again, probably because of a small sample of firm-days with earnings announcements, none of these tests show statistical significance.

The evidence reveals that HFTs become more active in trading and quoting under a higher aggregate volatility. If investors fear HFT for the risk of "front-running" or "predatory trading," this evidence suggests that they still prefer immediate transactions to avoiding HFT when volatility is high. In the context of the pecking order hypothesis, the evidence here is inconsistent with the conjecture that investors use lit venues more under a higher volatility *because* HFTs become less active in these situations.

Figure 8: HFT responses to shocks in VIX and earnings This figure shows the responses of the two HFT activity measures after a shock to VIX in Panel (a), and earnings, in Panel (b). In each panel, the left-hand graph shows the IRF of the participation fraction of HFT in the top depth, and the right-hand graph shows the IRF of the HFT trading volume fraction. The two-standard-deviation confidence bounds are constructed by simulation.

(a) Response to a VIX shock

(b) Response to an earnings announcement shock
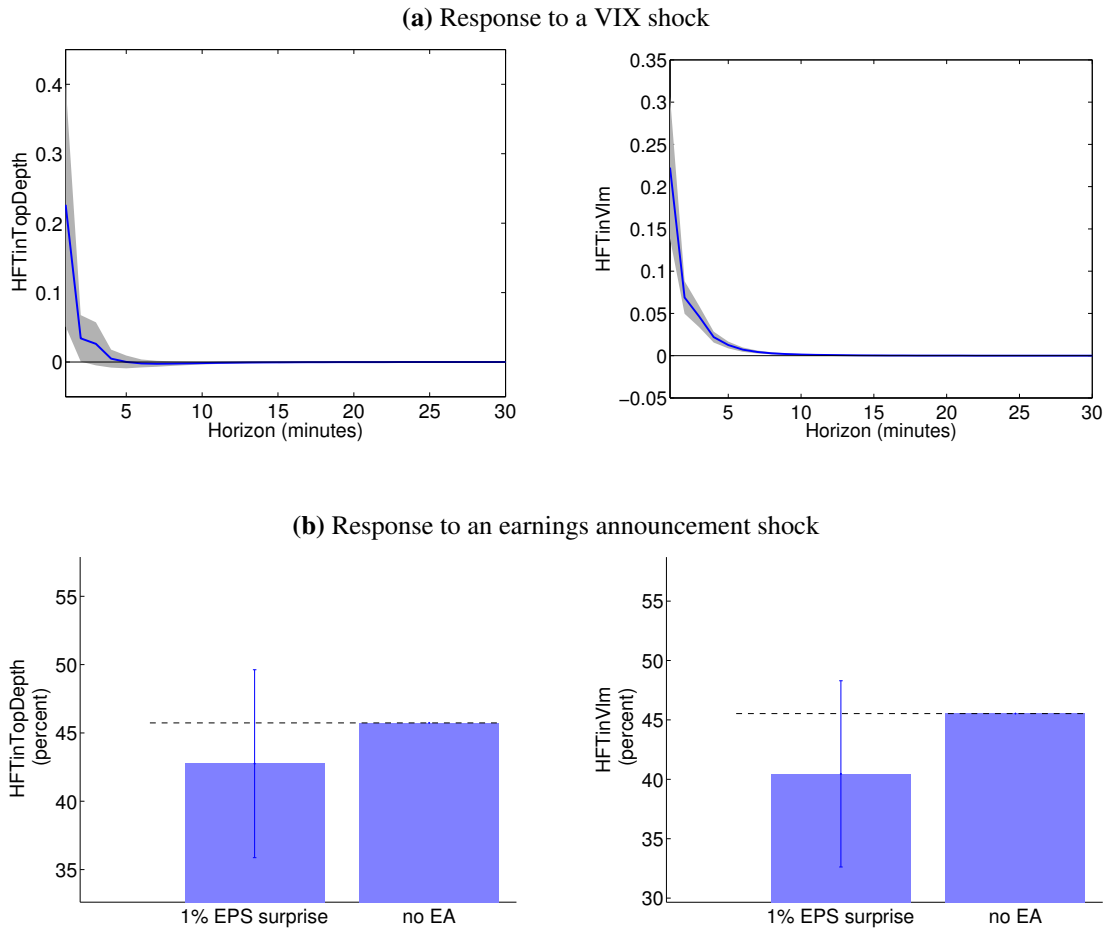
6 Simple Model to Formalize the Pecking Order Hypothesis, and Make Welfare Comparisons

This section proposes a simple model that characterizes investors' choices among three venue types: DarkMid, DarkNMid, and Lit. As discussed above, relative to existing theories of dark pools, our simple model distinguishes different types of dark venues. The model and its analysis formalize

the intuition that led to the pecking order hypothesis. The model further allows us to quantitatively measure what market fragmentation means for welfare.

6.1 Model setup

Asset. There is one traded asset. Its fundamental (common) value is normalized to zero. All players in this model have symmetric information about the asset and value it at zero. To formalize a pecking order hypothesis based on the urgency of trades, a symmetric-information setting suffices.

Venues. There are three trading venues: Lit, DarkMid, and DarkNMid.

- Lit is populated by infinitely many competitive and infinitesimal liquidity providers who have the same marginal cost β (> 0) for taking on one unit of the asset per capita, either long or short. The cost can be an operation cost or a margin cost. Together, these liquidity providers supply infinite depth at prices β and $-\beta$. This construct is similar to the "trading crowd" assumption in, for example, Seppi (1997) and Parlour (1998).

- DarkMid crosses buy and sell orders at the midpoint price, i.e., at 0. If unbalanced, only the matched part of the order flow gets executed. For example, if there are buy orders for 100 units in total and sell orders for 40 units then only 40 units are matched and executed.

- DarkNMid is run by a single competitive liquidity provider who starts with inventory zero, but incurs an inventory cost of $-\eta x^2/2$ for taking a (long or short) position of x, where $\eta > 0$. The liquidity provider posts an ask price schedule,

$$p^+ = \delta q^+,$$

where $q^+ > 0$ denotes the quantity to buy, and a bid price schedule,

$$p^- = -\delta q^-,$$

33

where $q^- > 0$ denotes the quantity to sell. The constant δ will be determined in equilibrium such that the liquidity provider breaks even in expectation. (We restrict to linear schedules for simplicity.) The average transaction price for sending in a buy order of size q is $\delta q/2$. For a sell order it is $-\delta q/2$. The liquidity provider can be viewed as the broker-dealer who operates this dark venue.

Investors. There are two large "representative" investors, a buyer and a seller. Having only two investors, rather than many, simplifies the exposition greatly at little cost of economics. The large buyer and seller receive independent and identically distributed liquidity shocks. The buyer starts with a short position $-Z^+ \leq 0$ and wishes to buy quantity Z^+. The seller starts with a long position $Z^- \geq 0$ and wishes to sell Z^-. Z^+ and Z^- are Bernoulli trials, $Q > 0$ with probability $\phi \in (0, 1)$ and zero with probability $(1 - \phi)$. This distributional assumption both keeps the analysis tractable and captures one salient feature of the data, i.e., for most minutes dark volume is zero.

If a large investor is left with a non-zero inventory after the trading round (i.e., the difference between the desired trade amount and the actual traded amount), say x, he incurs a quadratic cost of $(\gamma/2)x^2$, with $\gamma > 0$. Here, γ could be an inventory cost, a proxy for risk-aversion, or the cost of a missed opportunity to trade on a short-lived private signal. The γ is the key parameter of the model. We interpret it broadly as investors' urgency to trade: The higher is γ, the larger is the cost of holding a non-zero net position, and hence investors are more eager to trade.[16]

Timing. There is one trading round. Before trading, the large buyer and seller observe, privately, their shock sizes. They then decide simultaneously on the order sizes they send to the three venues. Venues execute trades simultaneously according to their specific trading protocols.

---

[16] For example, the interpretation of risk aversion is consistent with Campbell, Grossman, and Wang (1993), who argue that the market's aggregate risk aversion—which is correlated with VIX, for example—reflects the change of (a subset of) individual investors' risk aversion.

6.2 Equilibrium

We use "x" to denote the quantities sent to the trading venues. The superscripts "$+$" and "$-$" denote variables associated with the buyer and the seller, respectively. The subscripts "M", "N", and "L" denote DarkMid, DarkNMid, and Lit, respectively. We focus on a symmetric-strategy equilibrium, i.e., the buyer and the seller choose the same order flow sizes (but different signs): $x_i^+(Z) = x_i^-(Z)$ for all $(i, Z) \in \{M, N, L\} \times \{0, Q\}$.

Optimal order sizes. Fix the seller's strategy $x_i^-(Z^-)$ and consider the buyer's choice of $x_i^+(Z^+)$ for all $(i, Z^+, Z^-) \in \{M, N, L\} \times \{0, Q\} \times \{0, Q\}$. Suppose that the buyer receives a shock of size z. Note that from the buyer's perspective, the seller's order size Z^- is a Bernoulli random variable. To emphasize this randomness, we use capital letters to denote X_i^-, for $i \in \{M, N, L\}$, and let $V_M^+ := \min\{x_M^+, X_M^-\}$. Then, the buyer's expected profit is

$$
\pi^+(z) = \overbrace{-\mathbb{E}\left[0 \cdot V_M^+(z)\right]}^{\text{price to pay in DarkMid}} \overbrace{-\frac{\delta}{2} x_N^+(z)^2}^{\substack{\text{price to pay} \\ \text{in DarkNMid}}} \overbrace{-\beta \cdot x_L^+(z)}^{\text{price to pay in Lit}}
$$
$$
+ \underbrace{\mathbb{E}\left[0 \cdot \left(z - V_M^+(z) - x_N^+(z) - x_L^+(z)\right)\right]}_{\text{liquidation value of remaining position}} \underbrace{-\frac{\gamma}{2} \mathbb{E}\left(z - V_M^+(z) - x_N^+(z) - x_L^+(z)\right)^2}_{\text{quadratic cost for failing to trade}},
$$

which can be simplified to

$$
\pi^+(z) = -\frac{\gamma}{2} \mathbb{E}\left(z - V_M^+(z) - x_N^+(z) - x_L^+(z)\right)^2 - \frac{\delta}{2} x_N^+(z)^2 - \beta x_L^+(z). \tag{2}
$$

The buyer thus chooses six parameters to maximize his expected profit (2): $x_M^+(0)$, $x_N^+(0)$, $x_L^+(0)$, $x_M^+(Q)$, $x_N^+(Q)$, and $x_L^+(Q)$.

Because we look for a symmetric equilibrium, from this point on we suppress the superscript $\pm$ unless we need to explicitly distinguish a buyer from a seller.

Proposition 1 (Equilibrium order flows). *If*

$$Q \leq \Delta \equiv \beta \left(\frac{1}{(1 - \phi)\gamma} + \frac{1}{(1 - \phi)\eta} \right), \tag{3}$$

then there exists an equilibrium with the following strategies:

$$x_M(0) = 0; \quad x_M(Q) = \frac{\delta}{\delta + (1 - \phi)\gamma} Q, \tag{4}$$

$$x_N(0) = 0; \quad x_N(Q) = \frac{(1 - \phi)\gamma}{\delta + (1 - \phi)\gamma} Q, \tag{5}$$

$$x_L(0) = 0; \quad x_L(Q) = 0. \tag{6}$$

If $Q > \Delta$, then there exists an equilibrium with the following strategies:

$$x_M(0) = 0; \quad x_M(Q) = \frac{\beta}{(1 - \phi)\gamma}, \tag{7}$$

$$x_N(0) = 0; \quad x_N(Q) = \frac{\beta}{\delta}, \tag{8}$$

$$x_L(0) = 0; \quad x_L(Q) = Q - \Delta. \tag{9}$$

In both cases, the DarkNMid liquidity provider sets the slope of price schedules $\delta = (1 - \phi)\eta$.

The intuition of this equilibrium is as follows. Clearly, if the investor draws a zero liquidity shock then he does not send orders anywhere. If he draws the nonzero shock Q then he first considers Dark-Mid because it has the best transaction price (saving half the spread). Because the order size Q cannot be executed in DarkMid at zero (inventory) cost, he must strike a balance between nonexecution risk in DarkMid, the price impact in DarkNMid, and the spread in Lit.

The following heuristic argument illustrates this trade-off. A buyer's expected marginal cost of sending an additional infinitesimal quantity to DarkMid is $(1 - \phi)x_M(Q)\gamma$, where we use the fact that the DarkMid order size $x_M(Q)$ is not matched at all if the seller is small, and that the marginal inventory cost of holding unexecuted quantity q is γq. The marginal cost of sending an additional infinitesimal quantity to DarkNMid is $\delta x_N(Q)$. In equilibrium, these two marginal costs must be

equal, for otherwise the buyer would want to take a marginal unit from one dark venue to the other. Thus, we have

$$(1 - \phi)x_M(Q)\gamma = \delta x_N(Q). \tag{10}$$

The above equation says that $x_M(Q)$ and $x_N(Q)$ must increase simultaneously in a fixed proportion as Q increases.

If $Q \leq \Delta$, we can show that the marginal cost of using DarkNMid $\delta x_N(Q)$ is no larger than the Lit spread β. Thus, Lit is not used. The equilibrium order sizes can be solved by observing that in equilibrium $x_M(Q) = Q - x_N(Q)$, that is, the desired quantity in DarkMid is equal to the residual. If $Q > \Delta$, equation (10) hits a corner at

$$(1 - \phi)x_M(Q)\gamma = \delta x_N(Q) = \beta, \tag{11}$$

from which we can solve $x_M(Q)$ and $x_N(Q)$. Lit order size is given by $x_L(Q) = Q - x_M(Q) - x_N(Q)$.

Note that the equilibrium of Proposition 1 is not unique. There exists another, less interesting equilibrium in which neither the buyer nor the seller uses DarkMid, even if they happen to be large. This other equilibrium has the counterfactual implication that DarkMid has zero market share.

Interestingly, the DarkNMid liquidity provider is more aggressive (i.e., he sets prices that respond less to additional quantity) if large orders are more likely. This is because he makes more profits from large investors, even though it also increases his inventory cost.

6.3 Urgency elasticity of venue market shares

To establish the pecking order hypothesis, we are interested in how the market share of each of the three venues responds to a change in investor urgency. We first compute the expected market shares in equilibrium and then rank the venues according to their market share elasticities with respect to the urgency parameter, γ.

The focus on market shares as opposed to raw volume is consistent with existing empirical studies

of dark pools and fragmentation (see, for example, O'Hara and Ye, 2011; Buti, Rindi, and Werner, 2011; and Ready, 2014) as well as the empirical tests conducted in Section 5. From this point on we focus on the second case of Proposition 1 because it implies a strictly positive market share for all venues, as in reality.

A positive γ shock is naturally thought of as the model equivalent of a shock to VIX or EPS in the empirical approach. It raises investors' opportunity cost of not trading. One could argue that the elevated volatility following VIX and EPS shocks also raises inventory cost for intermediaries in DarkNMid and Lit. Therefore their opportunity cost of not trading is raised as well. The additional volume that follows these shocks, however, leads us to believe that at least part of the shock is attributable to a disproportionately large shock to investors as compared to intermediaries. One could think of γ as a reduced-form representation of a risk-aversion shock experienced by end investors in excess of that by intermediaries. This channel is perhaps most eloquently stated in Campbell, Grossman, and Wang (1993, p. 924):

> "We consider an economy in which there exist two assets: a risk-free asset and a risky asset ('stock'). We assume that innovations in the stock price are driven by three random variables: (i) the innovation to the current dividend, (ii) the innovation to information about future dividends, and (iii) the innovation to the time-varying risk aversion of a subset of investors. Shock (i) causes the payoff to the stock to be stochastic so that a premium is demanded by investors for holding it. Shock (iii) generates changes in the market's aggregate risk aversion, which cause the expected return on the stock to vary. Shock (ii) is in the model so that prices and dividends do not fully reveal the state of the economy and *volume* provides additional information." [Emphasis added]

A γ shock in our model corresponds to shock (iii) in the "canonical" Campbell, Grossman, and Wang (1993) model because it captures all the effects we empirically document. Shocks (i) and (ii) generate volatility and a risk premium, but they do not generate trading. Reallocation of assets only happens when risk aversion in one group changes relative to another group, i.e., a type (iii) shock.

Moreover, shock (iii) in and of itself generates elevated volatility as it triggers a price change due to changing "aggregate risk aversion." A type (iii) shock therefore becomes a sufficient condition to generate the empirical pattern of elevated volatility *and* volume.

The expected volume in the three venues and the total volume are given by

$$\bar{v}_M = \phi^2(2x_M(Q)) + (1 - \phi)^2(2x_M(0)) + 2\phi(1 - \phi)(2x_M(0)) = \frac{2\phi^2\beta}{(1 - \phi)\gamma}, \tag{12}$$

$$\bar{v}_N = \phi^2(2x_N(Q)) + (1 - \phi)^2(2x_N(0)) + 2\phi(1 - \phi)(x_N(Q) + x_N(0)) = 2\frac{\phi\beta}{\delta}, \tag{13}$$

$$\bar{v}_L = \phi^2(2x_L(Q)) + 2\phi(1 - \phi)x_L(Q) = 2\phi\left(Q - \frac{\beta}{(1 - \phi)\gamma} - \frac{\beta}{\delta}\right), \tag{14}$$

$$\bar{v} = \bar{v}_M + \bar{v}_N + \bar{v}_L = 2\phi Q - \frac{2\phi\beta}{\gamma}. \tag{15}$$

Note that in the above calculation we double-count volume in DarkMid. In practice, operators of DarkMid typically act as buyer to the seller and seller to the buyer, so one match shows up as two trades. Our pecking order results are not affected if DarkMid volume is single-counted.

The volume shares of different venues are defined as,

$$s_i := \frac{\bar{v}_i}{\bar{v}}, \quad \text{for } i \in \{M, N, L\}.$$

Signing partial derivatives of volume shares with respect to γ yields the model's main proposition. It formalizes this paper's pecking order hypothesis, depicted in Panel (b) of Figure 2.

Proposition 2 (Venue share and urgency). *As investor urgency increases, lit volume share increases and dark volume share decreases. Furthermore, DarkMid is more sensitive to urgency than DarkNMid:*

$$\frac{\partial s_M/s_M}{\partial \gamma/\gamma} < \frac{\partial s_N/s_N}{\partial \gamma/\gamma} < 0 < \frac{\partial s_L/s_L}{\partial \gamma/\gamma}.$$

The candidate model of this section is meant to be simple but yet able to demonstrate that the

pecking order hypothesis can emerge endogenously as an equilibrium outcome. There are a number of potential extensions to enrich the model. One direction is to incorporate asymmetric information. Another is to directly model the desired trade size Q of large investors from first principles. Yet another interesting but technically difficult direction is to consider a fully dynamic model in which investors can route orders across the three venue types. These potential extensions are out of the scope of the empirical focus of this paper and left for future research.

6.4 The welfare implications of market fragmentation

In this subsection we examine the welfare implications of market fragmentation across venue types in the same modeling framework. Welfare is measured by the overall costs, including explicit transaction costs and implicit inventory costs, to the representative buyer and seller. A model is important for measuring welfare because one cannot directly observe counterfactuals in the data, namely investors' costs without dark venues.

We examine four market structures: {Lit} (L), {DarkMid, Lit} (ML), {DarkNMid, Lit} (NL), and {DarkMid, DarkNMid, Lit} (MNL). In each market structure, we can solve the equilibrium as in the following proposition.

Proposition 3. *Suppose that $Q > \Delta$. The following strategies make up an equilibrium to the trading game in the various market structures: L, ML, NL, and MNL.*

1. *An investor who draws a zero liquidity shock does not send any order to any of the venues that are available; that is, $x^{\pm}(0) = 0$ for all market structures.*

2. *If DarkNMid is open, the DarkNMid liquidity provider sets $\delta = (1 - \phi)\eta$.*

3. *An investor who draws the nonzero shock Q sends orders to venues according to the following table:*

Market structure	DarkMid	DarkNMid	Lit
MNL	$\frac{\beta}{(1-\phi)\gamma}$	$\frac{\beta}{(1-\phi)\eta}$	$Q - \frac{\beta}{(1-\phi)\gamma} - \frac{\beta}{(1-\phi)\eta}$
ML	$\frac{\beta}{(1-\phi)\gamma}$		$Q - \frac{\beta}{(1-\phi)\gamma}$
NL		$\frac{\beta}{\delta}$	$Q - \frac{\beta}{\delta} - \frac{\beta}{\gamma}$
L			$Q - \frac{\beta}{\gamma}$

The proof of Proposition 3 follows the same logic as the proof of Proposition 1 and is omitted.

These strategies allow us to calculate the expected costs to investors, including explicit execution costs paid to liquidity providers and the implicit inventory cost for unexecuted quantities. We use "*C*" to denote these costs. Straightforward calculation yields:

$$C_{MNL} = \phi^2 \left(2x_L(Q)\beta + 2\frac{\delta x_N(Q)^2}{2} \right) + (1-\phi)^2 \cdot 0 + 2\phi(1-\phi)\left[x_L(Q)\beta + \frac{\delta x_N(Q)^2}{2} + \frac{\gamma}{2}\left(\frac{\beta}{(1-\phi)\gamma} \right)^2 \right]$$

$$= 2\phi\beta Q - \frac{\phi\beta^2}{(1-\phi)\gamma} - \frac{\phi\beta^2}{(1-\phi)\eta}, \tag{16}$$

$$C_{ML} = 2\phi\beta Q - \frac{\phi\beta^2}{(1-\phi)\gamma}, \tag{17}$$

$$C_{NL} = 2\phi\beta\left(Q - \frac{\beta}{\delta} - \frac{\beta}{\gamma} \right) + 2\phi\delta\frac{(\beta/\delta)^2}{2} + 2\phi\frac{\gamma}{2}(\beta/\gamma)^2 = 2\phi\beta Q - \frac{\phi\beta^2}{(1-\phi)\eta} - \frac{\phi\beta^2}{\gamma}, \tag{18}$$

$$C_L = 2\phi\beta\left(Q - \frac{\beta}{\gamma} \right) + 2\phi\frac{\gamma}{2}(\beta/\gamma)^2 = 2\phi\beta Q - \frac{\phi\beta^2}{\gamma}. \tag{19}$$

From these costs we can calculate the welfare cost of shutting down one or both types of dark venues. We use the current market structure, MNL, as the benchmark.

$$C_L - C_{MNL} = \frac{\phi\beta^2}{(1-\phi)\eta} + \frac{\phi^2\beta^2}{(1-\phi)\gamma} = \frac{\beta}{2}(\bar{v}_M + \bar{v}_N), \tag{20}$$

$$C_{NL} - C_{MNL} = \frac{\phi^2\beta^2}{(1-\phi)\gamma} = \frac{\beta}{2}\bar{v}_M, \tag{21}$$

$$C_{ML} - C_{MNL} = \frac{\phi\beta^2}{(1-\phi)\eta} = \frac{\beta}{2}\bar{v}_N. \tag{22}$$

41

Interestingly, the welfare costs of shutting down dark venues is equal to $\beta/2$, a half of the half-spread β, multiplied by the volume attracted by these dark venues. This result is not obvious *ex ante*. For example, the actual cost saving in DarkMid turns out to be only a half of the dollar value executed, $\beta \bar{v}_M$, because of the nonexecution risk in DarkMid.

As the trading volumes and the spread in Lit can be directly read from the data, it is unnecessary to estimate the structural parameters like preference γ and η and order size Q. A quantitative assessment of how fragmentation affects welfare becomes relatively straightforward.

Dollar-value measure. We first use a dollar-value measure to quantify welfare improvement generated by the availability of dark venues, i.e., the saved cost, $C_L - C_{LNM}$.

Using equation (20), the aggregate daily dollar cost saving for all 117 of the sample stocks and over all 390×21 minutes is

$$\frac{1}{21} \sum_{t=1}^{390 \times 21} \sum_{j=1}^{117} \frac{\beta(j,t)P(j,t)}{2} \left(v_M(j,t) + v_N(j,t)\right), \tag{23}$$

where $P(j,t)$ is the price of the stock j at the end of minute t. Note that this calculation essentially scales the dollar spread by dark volumes. This measure of cost savings amounts to $40,999 per day for our sample stocks.

Although our sample only includes 117 stocks, a rather small fraction of the entire U.S. equity market, it provides a fair representation of large-cap, medium-cap, and small-cap stocks (see Section 3). To the extent that our sample reasonably represents the composition of the universe of U.S. stocks, one way to gauge the cost savings across the entire market is to simply scale the above estimate by the ratio of trading volume of our sample to the total equity volume.[17]

Our sample has an average daily trading volume of 0.56 billion shares. The U.S. equity daily volume is approximately 7.80 billion shares, as of October 2014.[18] Hence, scaling up the aggregate

---

[17] A sufficient but not necessary condition for the validity of this approach is that conditional on the market cap tercile of a stock, its dollar spread is uncorrelated with volume in DarkMid and DarkNMid.

[18] The estimate of daily equity trading volume in U.S. is obtained from Tabb Liquidity Matrix: http://tabbforum.com/liquidity-matrix. Since we extrapolate from a sample of October 2010 to the broad market of October 2014, this extrapolation implicitly assumes that the average characteristics of stocks do not move substantially

daily costs savings by a factor of 7.80/0.56 yields a broad-market cost saving—due to the presence of DarkMid and DarkNMid—of $5.71 million per day, or $1.43 billion per year. Of course, this is a rough estimate and likely to be noisy.

This cost saving is economically sizeable. For instance, Berk and van Binsbergen (2014) find that an average mutual fund in their sample adds $3.2 million per year, gross of fees. Since their sample consists of about 6000 funds, the total value added amounts to $19.2 billion per year. Thus, the cost saved by DarkMid and DarkNMid is 7.4% relative to the value added by all mutual funds in the Berk and van Binsbergen (2014) sample.

Relative measures. We consider two additional relative measures to quantify the cost saving. The first measure is the saved cost due to dark venues, $C_L - C_{LNM}$, as a fraction of the total spread cost in Lit, $\beta \bar{v}_L$. We have

$$\frac{C_L - C_{LNM}}{\bar{v}_L \beta} = \frac{\bar{v}_M + \bar{v}_N}{2\bar{v}_L}. \tag{24}$$

Calculating this ratio stock by stock and taking the average, we get 4.6%.[19]

The second relative measure is $C_L - C_{LNM}$ as a fraction of the total cost of execution and inventory, C_{LNM}, under the current market structure. We have

$$\frac{C_L - C_{LMN}}{C_{LMN}} = \frac{\beta(\bar{v}_M + \bar{v}_N)}{2\phi\beta Q + \beta\bar{v}_L} = \frac{\bar{v}_M + \bar{v}_N}{2\phi Q + \bar{v}_L}. \tag{25}$$

The total latent trading interest $2\phi Q$ is not observed in the execution data, but we can bound it by observing that $2\phi Q > \bar{v}_L + \bar{v}_M + \bar{v}_N$ (see equation (15)). Thus,

$$\frac{C_L - C_{LMN}}{C_{LMN}} < \frac{\bar{v}_M + \bar{v}_N}{\bar{v}_M + \bar{v}_N + 2\bar{v}_L}. \tag{26}$$

---

over time.

[19] Alternatively, one can directly use the average volume reported in Table 2 and obtain

$$\frac{C_L - C_{LNM}}{\bar{v}_L \beta} = \frac{\bar{v}_M + \bar{v}_N}{2\bar{v}_L} \approx \frac{0.259 + 0.941}{2 \times 8.931} \approx 6.7\%.$$

This estimate is higher than the average of stock-by-stock estimates because of the Jensen's Inequality effect.

43

Again, calculating this ratio stock by stock and taking the average, we get 3.8%.[20]

Discussion. There are three caveats to our dollar-value and relative estimate if they are to be inter-preted for policy. First, in the model the Lit spread β is invariant to market structure, which is a direct consequence of the assumption that no one has superior information about the asset fundamentals. If adverse selection were present, we would expect the two dark venues to attract disproportionately more uninformed order flows (see Zhu 2014 and the empirical evidence discussed there). Thus, if the dark venues were closed, we would expect the lit spread to become smaller on average, reducing the investors' explicit transaction costs in Lit and the implicit delay costs of unexecuted orders.

Second, our measure of fragmentation is at the level of venue types, but not at individual venues. Our estimates do not address the cost of fragmented liquidity within dark venues of the same type, such as multiple DarkNMid venues. Third, our model does not address the other three dark venue types: DarkRetail, DarkOther, and DarkPrintB, for reasons discussed in Section 5.2. The inclusion of these three types of dark venues in our calculation will likely increase the estimated cost savings due to dark venues.

7 Conclusion

In this paper, we propose and test a pecking order hypothesis for the dynamic fragmentation of U.S. equity markets. The hypothesis posits that investors disperse their orders into a series of venue types, sorted along a pecking order. The position of venue types on the pecking order depends on the trade-off between cost (price impact) and immediacy (execution certainty). On the top of the pecking order are low-cost, low-immediacy venues such as midpoint dark pools, whereas in the bottom are high-cost, high-immediacy venues such as lit exchanges; in the middle of the pecking order are non-

---

[20] Alternatively, one can directly use the average volume reported in Table 2 and obtain

$$\frac{C_L - C_{LMN}}{C_{LMN}} < \frac{\bar{v}_M + \bar{v}_N}{\bar{v}_M + \bar{v}_N + 2\bar{v}_L} \approx \frac{0.259 + 0.941}{0.259 + 0.941 + 2 \times 8.931} \approx 6.3\%.$$

midpoint dark pools. A positive shock to investors' urgency to trade tilts their order flows from the top of the pecking to the bottom; therefore, the elasticities of venue market shares to urgency shocks are progressively less negative and more positive further down the pecking order.

We test the pecking order hypothesis using a unique dataset that identifies the minute-by-minute trading volumes in five different types of dark venues in the U.S. equity markets. A panel VARX model characterizes the dynamic interrelation between dark trading volumes, HFT participation in lit venues, and market conditions, as well as two exogenous shocks to investors' demand for immediacy: VIX and earnings surprises.

Consistent with the pecking order hypothesis, we find that an upward shock to VIX substantially reduces the market share of midpoint dark pools, moderately reduces the shares of non-midpoint dark pools, but increases the share of lit venues. After shocks to earnings surprises, the share of midpoint pools also declines significantly. Our results offer a pecking order representation for the dynamic fragmentation of U.S. equity markets.

Finally, we propose a simple theoretical model that formalizes the pecking order hypothesis. The model also allows us to quantitatively gauge the impact of dark venues on investors' welfare, measured by their explicit transaction costs and implicit inventory costs. A simple calibration suggests that DarkMid and DarkNMid, two types of dark venues that we focus on, save investors $1.43 billion of costs per year. This amounts to 7.4% of annual gross dollar value added by mutual funds as estimated by Berk and van Binsbergen (2014).

Appendix

A   Details on the implementation of the panel VARX model

In this appendix we discuss the details of the panel VARX model.

The estimation is implemented via OLS by stacking the observations associated with different stocks into a single vector. The stock fixed effect is accounted for by adding dummy variables to the set of regressors. Lags of the variables are only constructed intraday.

The optimal numbers of lags p and r are chosen according to Bayesian Information Criterion (BIC). Specifically, for each of the 117 stocks, the VARX model is estimated for all pairs of $(p, r) \in \{1, 2, ..., 10\} \times \{1, 2, ..., 5\}$. Then the best (according to BIC) model is chosen as (p_j, r_j) for stock j. That is, we confine the search of the optimal lags within 10 and 5 lags, respectively, for endogenous and exogenous variables. The above procedure generates 117 pairs of optimal lags of $\{p_j\}_{j=1}^{117}$ and $\{r_j\}_{j=1}^{117}$. There are 16 p_j that are found to be 1, 93 to be 2, and the other 8 to be 3. All r_j are 1. We hence choose $p = 2$ and $r = 1$ for parsimony.

The standard errors for panel data estimators should account for potential correlation through time and across stocks. One standard way to account for these issues is to do "double-clustering" (Petersen, 2009). The laborious (but most flexible) way of implementing such clustering is by calculating

$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = (X'X)^{-1} V (X'X)^{-1} \text{ with } v_{ij} = \sum_{kt,ls} x_{ikt} \hat{\varepsilon}_{ikt} \hat{\varepsilon}_{jls} x_{jls} \times 1_A(ktls), \tag{27}$$

where $i, j \in \{1, \ldots, N\}$ where N is the number of regressors, $k, l \in \{1, \ldots, J\}$ where J is the number of stocks, and $s, t \in \{1, \ldots, T\}$ where T is the number of time periods. $1_A(ktls)$ is the indicator function where the subset A of the index value space identifies which auto- or cross-correlations a researcher worries about. If error terms are independent and identically distributed, then the indicator function equals one if $k = l$ and $t = s$. The subset A for an indicator function in a standard double-clustering is such that:

$$1_A(ktls) = \begin{cases} 1 & \text{if } k = l \text{ or } s = t, \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

A researcher can easily be more conservative and also account for non-zero cross-autocorrelations by also including changing the $s = t$ condition by, say, $|s - t| \leq 5$.

The cumulative impulse response function is most easily calculated by stacking the estimated Φ matrices, as any VAR can always be written as a first-order VAR. Consider, for example, a VAR with two lags. This VAR can be written as

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix}^t \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}. \tag{29}$$

The t-period cumulative impulse response of the jth variable to a unit impulse in the ith variable is the jth element of the vector

$$\begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix}^t e_k, \tag{30}$$

where I is the identity matrix and e_k is the unit vector where the kth element is one and all other elements are zeros.

Confidence intervals on the impulse response function (IRF) are obtained through simulation. The IRF is a non-linear transformations of the VARX coefficient estimates denoted by $\hat{\theta}$. Each simulation involves a draw from the multivariate normal distribution $N\left(\hat{\theta}, \Sigma_{\hat{\theta}}\right)$, where $\Sigma_{\hat{\theta}}$ is the estimated double-clustered covariance matrix of the coefficients. Note that this distribution is asymptotically true given the assumption that the VARX model is correctly specified with normal residual terms. We perform 10,000 independent draws of the coefficients $\hat{\theta}$ and for each draw compute the IRFs at all lags. Thus, we obtain, for each IRF, an i.i.d. sample of size 10,000. The confidence bounds are then chosen at the 2.5 and the 97.5 (or 0.5 and 99.5) percentiles of the simulated IRFs. The significance levels shown in table 3 are based on whether the estimates exceed the confidence bounds found above.

B   Transformation between logarithms and levels of market share variables

In the VARX model implementation, the variables are log-transformed (except *EpsSurprise*). Log-transformation has several advantages. For example, the strictly positive variables (e.g. volume, spread, depth, etc.) are converted to a possibly negative support; the concavity in logarithm discourages the abnormal effects of outliers; the estimation coefficients can be readily interpreted as elasticity. The key variables of our focus are the (logged) trading volumes in the five dark venues and the lit venue, denoted by $\log v_1$, ..., $\log v_5$, and $\log v_6$, where the first five are for the five dark venues and the last v_6 is for the trading volume in the lit. (Each of these variables has stock-day-minute granularity.) For the pecking order hypothesis, it is however useful to think in terms of market shares, defined as

$$s_j = \frac{v_j}{\sum_j v_j}$$

for $j \in \{1, ..., 6\}$.

The purpose of this appendix is to derive the closed-form, exact transformation formula from a shock in trading volume in one venue to the response of all market shares. Specifically, given a shock of $\Delta \log v_i$, we want to know the immediate response Δs_j, for *all* $j \in \{1, ..., 6\}$. Reverse directions from Δs_i to $\Delta \log v_j$ will also be dealt with. These formulas are used in generating the impulse responses in testing the pecking order hypothesis.

In the derivation below, we shall use the following additional notations. Let v be the total volume: $v = \sum_j v_j$. We shall use a superscript of "+" to denote the variables after a shock; for example, $\log v_j^+ = \log v_j + \Delta \log v_j$. Similarly, while s_j denotes the market share of venue j, s_j^+ denotes the market share after the shock. In the IRF exercise, the pre-shock values will be chosen as the stock-day-minute average across all raw sample observations. Consider the following cases.

From $\Delta \log v_i$ to Δs_j. By construction, $\log v_i^+ = \log v_i + \Delta \log v_i$. Taking the exponential on both sides gives the *level* of the post-shock trading volume: $v_i^+ = v_i \exp \Delta \log v_i$. The post-shock market share by construction is

$$s_i^+ = \frac{v_i^+}{v^+} = \frac{v_i^+}{v + (v_i^+ - v_i)}.$$

Substituting with the expression of v_i^+ and then subtracting $s_i = v_i/v$ yields

$$\Delta s_i = s_i^+ - s_i = ... = s_i \cdot \left(e^{\Delta \log v_i - \Delta \log v} - 1 \right) \tag{31}$$

where $\Delta \log v = \log v^+ - \log v = \log(\sum_{j \neq i} v_j + v_i^+) - \log v$. The above formula actually applies to both the venue i whose volume is shocked and any other venue $j \neq i$ whose volume is not shocked. The only difference is, as can be seen after substituting the index i with a different j, that $\Delta \log v_j = 0$ for $j \neq i$. Finally, we can immediately derive the dark volume share change as the complement of the change in the lit share: $\sum_{j \leq 5} \Delta s_j = -\Delta s_6$, simply because the identity of $s_6 = 1 - \sum_{j \leq 5} s_j$.

From $\Delta \log v$ to Δs_i, assuming proportionally scaling across all venues. Now we shock the total volume such that $\log v^+ = \log v + \Delta \log v$ and make the assumption that the increase in volume is proportionally scaled across all venues. That is, for each venue i, $v_i^+ - v_i = s_i \cdot (v^+ - v)$, or

$$v_i^+ = v_i + s_i \Delta v = s_i v + s_i \Delta v = s_i \cdot (v + \Delta v) = s_i v^+.$$

Taking logarithm on both sides gives $\log v_i^+ = \log s_i + \log v^+$. Substitute $\log s_i$ with $\log s_i = \log(v_i/v) = \log v_i - \log v$ and then

$$\log v_i^+ = \log v_i - \log v + \log v^+ \implies \Delta \log v_i = \Delta \log v. \tag{32}$$

Applying equation (32) to (31) immediately gives $\Delta s_i = 0$. Clearly, this holds for all $i \in \{1, ..., 6\}$.

From Δs_i to $\Delta \log v_j$ by shocking $\log v_i$ and proportionally offsetting in other venues, without changing total volume v. Finally we do the reverse. Suppose we shock s_i by Δs_i. Such a change in market share must be driven by some change(s) in trading volume(s). Here a particular change is considered: Let v_i change in the same direction as s_i but all other $v_{j \neq i}$ move in the other direction so that the total volume does not change, i.e. $v = v^+$. We want to know, given the size of Δs_i, what are the sizes of $\Delta \log v_j$ for all $j \in \{1, ..., 6\}$.

First, consider $j = i$. By construction, $\Delta s_i = v_i^+/v^+ - s_i$. Because the total volume is assumed to be unchanged, we have $v_i^+ = (s_i + \Delta s_i)v$. This enables the second equality below:

$$\Delta \log v_i = \log v_i^+ - \log v_i$$

$$= \log(s_i + \Delta s_i) + \log v - \log v_i = \log(s_i + \Delta s_i) + \log \frac{v}{v_i} = \log(s_i + \Delta s_i) - \log s_i = \log\left(1 + \frac{\Delta s_i}{s_i}\right).$$

$$(33)$$

Consider next $j \neq i$. To offset Δv_i, summing over all $j \neq i$ gives $\sum_{j \neq i} \Delta v_j = -\Delta v_i$. Because the changes are proportional according to s_j, we have

$$v_j^+ = v_j - \frac{s_j}{\sum_{h \neq i} s_h}\Delta v_i = s_j \cdot \left(v - \frac{\Delta v_i}{\sum_{h \neq i} s_h}\right).$$

Take logarithm on both sides and expand $\log s_j = \log v_j - \log v$ to get

$$\log v_j^+ - \log v_j = \Delta \log v_j = -\log v + \log\left(v - \frac{\Delta v_i}{\sum_{h \neq i} s_h}\right) = \log\left(1 - \frac{\Delta v_i/v}{\sum_{h \neq i} s_h}\right) = \log\left(1 - \frac{\Delta s_i}{\sum_{h \neq i} s_h}\right),$$

$$(34)$$

where the last equality follows because the total volume is assumed to be unchanged: $\Delta v_i/v = v_i^+/v - v_i/v = v_i^+/v^+ - v_i/v = s_i^+ - s_i = \Delta s_i$.

C Proofs

C.1 Proof of Proposition 1

We directly verify the strategies. Without loss of generality, we consider the strategy of a buyer, who might get a zero liquidity shock ($z^+ = 0$) or a nonzero liquidity shock ($z^+ = Q$). We refer the buyer in these two conditions by "zero buyer" and "nonzero buyer," respectively.

The first case, $Q \leq \Delta$

Clearly, under these strategies, a zero buyer does not have any incentive to send an order to any of the venues and therefore abstains from trading. The nonzero buyer's cost is

$$-\pi^+(Q) = \frac{\gamma}{2}\mathbb{E}[(Q - V_M(Q) - x_N(Q) - x_L(Q))^2] + \frac{\delta}{2}x_N(Q)^2 + \beta x_L(Q). \tag{35}$$

Verify $x_N(Q)$. The first order condition with respect to $x_N(Q)$ yields

$$0 = \delta x_N(Q) - \gamma(Q - \mathbb{E}[V_M(Q)] - x_N(Q) - x_L(Q)), \tag{36}$$

which is easily verified by substituting the conjectured strategies.

Verify $x_L(Q) = 0$. The marginal cost of buying the last unit in DarkNMid is

$$\delta x_N(Q) = Q\frac{\delta(1 - \phi)\gamma}{\delta + (1 - \phi)\gamma} < \beta. \tag{37}$$

So the nonzero buyer has no incentive to deviate to Lit.

Verify $x_M(Q)$. Suppose the nonzero buyer chooses some other order size, y, in DarkMid. Note that $x_M(Q) = Q - x_N(Q)$. Clearly, the nonzero buyer has no incentive to deviate to $y > x_M(Q)$; this deviation does not change this payoff regardless of the size of the seller. Deviating to $y \in [0, x_M(Q))$ is also suboptimal: if the seller is "nonzero," this deviation reduces matched volume; if the seller is "zero," this deviation does not change the payoff.

The second case, $Q > \Delta$

The optimality of the zero investor's strategy is verified as before.

In this case, the conjecture is that $x_L(Q) > 0$. This means the first order condition of $-\pi^+$ with respect to $x_N(Q)$ and $x_L(Q)$ must both hold with equality. That is,

$$\delta x_N(Q) - \gamma(Q - \mathbb{E}[V_M(Q)] - x_N(Q) - x_L(Q)) = 0, \tag{38}$$

$$\beta - \gamma(Q - \mathbb{E}[V_M(Q)] - x_N(Q) - x_L(Q)) = 0. \tag{39}$$

It is easy to verify that these two first order conditions hold at the conjectured strategies. To verify the optimality of $x_M(Q)$, we note $x_M(Q) = Q - x_N(Q) - x_L(Q)$ and use the same argument as before.

Calculating the equilibrium δ

The last step is to calculate the equilibrium δ, such that the DarkNMid liquidity provider makes zero expected profit, net of inventory costs. Thus, the expected profit is equal to the expected inventory cost, i.e.,

$$\phi^2 \delta x_N(Q)^2 + (1-\phi)^2 \delta x_N(0)^2 + 2\phi(1-\phi)\left(\frac{\delta}{2} x_N(Q)^2 + \frac{\delta}{2} x_N(0)^2\right) = 2\phi(1-\phi) \cdot \frac{\eta}{2} (x_N(Q) - x_N(0))^2. \tag{40}$$

From this we solve

$$\delta = \eta \frac{\phi(1-\phi)(x_N(Q) - x_N(0))^2}{\phi x_N(Q)^2 + (1-\phi)x_N(0)^2} = \eta(1-\phi), \tag{41}$$

where the last equality follows from the equilibrium strategy $x_N(0) = 0$.

C.2 Proof of Proposition 2

We first calculate the partial derivatives of expected volumes with respect to γ. For each of the venue:

$$\frac{\partial \bar{v}_M}{\partial \gamma} = -\frac{2\phi^2\beta}{(1-\phi)\gamma^2} < 0$$

$$\frac{\partial \bar{v}_N}{\partial \gamma} = 0$$

$$\frac{\partial \bar{v}_L}{\partial \gamma} = \frac{2\phi\beta}{(1-\phi)\gamma^2} > 0$$

and for the total volume:

$$\frac{\partial \bar{v}}{\partial \gamma} = \frac{2\phi\beta}{\gamma^2} > 0.$$

Recall the definition of the market shares: $s_i := \bar{v}_i/\bar{v}$ for $i \in \{M, N, L\}$. Following the above calculation, we have 1) $\partial s_M/\partial \gamma < 0$ and 2) $\partial s_N/\partial \gamma < 0$, because while the denominator $\bar{v}$ increases in γ, the denominators (weakly) decrease. Then, by the accounting identity of $s_M + s_N + s_L = 1$ (hence $\partial(s_M + s_N + s_L)/\partial \gamma = 0$), we have 3) $\partial s_L/\partial \gamma > 0$. That is, both dark venue shares reduces and the lit share grows, as urgency γ increases.

It remains to rank the volume share sensitivity to urgency. Define the sensitivity as the elasticity of market share to urgency: $(\partial s_i/\partial \gamma)/(s_i/\gamma)$. This definition matches our empirical approach. Comparing DarkMid and DarkNMid gives

$$\frac{\partial s_N}{\partial \gamma}\frac{\gamma}{s_N} - \frac{\partial s_M}{\partial \gamma}\frac{\gamma}{s_M} = \frac{\gamma}{\bar{v}^2}\left(\overbrace{\frac{\partial \bar{v}_N}{\partial \gamma}\frac{\bar{v}}{s_N}}^{=0} - \frac{\partial \bar{v}}{\partial \gamma}\frac{\bar{v}_N}{s_N} \right) - \frac{\gamma}{\bar{v}^2}\left(\frac{\partial s_M}{\partial \gamma}\frac{\bar{v}}{s_M} - \frac{\partial \bar{v}}{\partial \gamma}\frac{\bar{v}_M}{s_M} \right)$$

$$= \frac{\gamma}{\bar{v}^2}\left[\frac{\partial \bar{v}}{\partial \gamma}\underbrace{\left(\frac{\bar{v}_M}{s_M} - \frac{\bar{v}_N}{s_N} \right)}_{=\bar{v}-\bar{v}=0} - \frac{\partial s_M}{\partial \gamma}\frac{\bar{v}}{s_M} \right] > 0.$$

Hence, the sensitivities of volume shares to urgency can be ranked as stated in proposition 2.

References

Berk, Jonathan B. and Jules H. van Binsbergen. 2014. "Measuring Skill in the Mutual Fund Industry." Working paper.

Boehmer, Ekkehart, Kingsley Fong, and Juan Wu. 2014. "International Evidence on Algorithmic Trading." Manuscript, EDHEC Business School.

Boni, Leslie, David C. Brown, and J. Chris Leach. 2012. "Dark Pool Exclusivity Matters." Working paper.

Boulatov, Alex and Thomas J. George. 2013. "Hidden and Displayed Liquidity in Securities Markets with Informed Liquidity Providers." *Review of Financial Studies* 26 (8):2095–2137.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. "High Frequency Trading and Price Discovery." *Review of Financial Studies, forthcoming* .

Brolley, Michael. 2014. "Should Dark Pools Improve Upon Visible Quotes? The Impact of Trade-at Rules." Working paper, University of Toronto.

Buti, Sabrina, Barbara Rindi, and Ingrid M. Werner. 2011. "Diving into Dark Pools." Working paper, Fisher College of Business, Ohio State University.

———. 2014. "Dark Pool Trading Strategies, Market Quality and Welfare." Working paper.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang. 1993. "Trading Volume and Serial Correlation in Stocks Returns." *The Quarterly Journal of Economics* 108:905–939.

Chowdhry, Bhagwan and Vikram Nanda. 1991. "Multimarket Trading and Market Liquidity." *Review of Financial Studies* 4:483–511.

Comerton-Forde, Carole and Tālis Putniņš. 2013. "Dark trading and price discovery." Working paper.

de Winne, Rudy and Catherine D'Hondt. 2007. "Hide-and-Seek in the Market: Placing and Detecting Hidden Orders." *Review of Finance* 11:663–692.

Degryse, Hans, Frank de Jong, and Vincent van Kervel. 2014. "The Impact of Dark Trading and Visible Fragmentation on Market Quality." *Review of Finance (forthcoming)* .

Degryse, Hans, Mark Van Achter, and Gunther Wuyts. 2009. "Dynamic Order Submission Strategies with Competition between a Dealer Market and a Crossing Network." *Journal of Financial Economics* 91:319–338.

Egginton, Jared, Bonnie F. Van Ness, and Robert A. Van Ness. 2013. "Quote Stuffing." Manuscript, University of Mississippi.

Foley, Sean, Katya Malinova, and Andreas Park. 2013. "Dark Trading on Public Exchanges." Working paper.

Foley, Sean and Tālis Putniņš. 2014. "Should we be afraid of the dark? Dark trading and market quality." Working paper.

Grossman, Sanford J. and Merton H. Miller. 1988. "Liquidity and Market Structure." *Journal of Finance* 43:617–633.

Hagströmer, Björn and Lars Nordén. 2013. "The Diversity of High Frequency Traders." Manuscript, Stockholm University School of Business.

Hasbrouck, Joel. 2013. "High Frequency Quoting: Short-Term Volatility in Bids and Offers." Manuscript.

Hatheway, Frank, Amy Kwan, and Hui Zheng. 2013. "An empirical analysis of market segmentation on U.S. equities markets." Working paper.

Hendershott, Terrence and Charles M. Jones. 2005. "Island Goes Dark: Transparency, Fragmentation, and Regulation." *Review of Financial Studies* 18:743–793.

Hendershott, Terrence and Haim Mendelson. 2000. "Crossing Networks and Dealer Marekts: Competition and Performance." *Journal of Finance* 55 (5):2071–2115.

Jovanovic, Boyan and Albert J. Menkveld. 2011. "Middlemen in Limit-Order Markets." Manuscript, VU University Amsterdam.

Kim, Oliver and Robert E. Verrecchia. 1994. "Market Liquidity and Volume around Earnings Announcements." *Journal of Accounting and Economics* 17:41–67.

Kinney, William, David Burgsthler, and Roger Martin. 2002. "Earnings Surprise "Materiality" as Measure by Stock Returns." *Journal of Accounting Research* 40 (5):1297–1329.

Kwan, Amy, Ronald Masulis, and Thomas McInish. 2014. "Trading Rules, Competition for Order Flow and Market Fragmentation." *Forthcoming, Journal of Financial Economics* .

Nagel, Stefan. 2012. "Evaporating Liquidity." *Review of Financial Studies* 25 (7):2005–2039.

Nimalendran, Mahendrarajah and Sugata Ray. 2014. "Informational Linkages Between Dark and Lit Trading Venues." *Journal of Financial Markets* 17:230–261.

O'Hara, Maureen and Mao Ye. 2011. "Is Market Fragmentation Harming Market Quality?" *Journal of Financial Economics* 100:459–474.

Pagano, Marco. 1989. "Trading Volume and Asset Liquidity." *Quarterly Journal of Economics* 104:255–274.

Parlour, Christine A. 1998. "Price Dynamics in Limit Order Markets." *Review of Financial Studies* 11:789–816.

Petersen, Mitchell A. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies* 22:435–480.

Ready, Mark J. 2014. "Determinants of Volume in Dark Pool Crossing Networks." Working paper, University of Wisconsin-Madison.

SEC. 2010. "Concept Release on Equity Market Structure." Release No. 34-61358; File No. S7-02-10.

———. 2013. "Equity Market Structure Literature Review Part I: Market Fragmentation."

Seppi, Duane J. 1997. "Liquidity Provision with Limit Orders and a Strategic Specialist." *Review of Financial Studies* 10:103–150.

Tuttle, Laura. 2014. "OTC Trading: Description of Non-ATS OTC Trading in National Market System Stocks." Memorandum, SEC.

Ye, Mao. 2011. "A Glimpse into the Dark: Price Formation, Transaction Costs and Market Share of the Crossing Network." Working paper, University of Illinois.

Zhu, Haoxiang. 2014. "Do Dark Pools Harm Price Discovery?" *Review of Financial Studies* 27 (3):747–789.