Dark trading and price discovery

Carole Comerton-Forde [a] and Tālis J. Putniņš [b] *
[a] *Department of Finance, University of Melbourne*
[b] *University of Technology Sydney* and *Stockholm School of Economics in Riga*

12 September 2014

Abstract

Regulators around the world are concerned that growth in dark trading may harm price discovery. We show that block and non-block dark trades affect price discovery differently. We find no evidence that block trades in the dark impede price discovery. In contrast, high levels of non-block dark trading harms price discovery and reduces the informational efficiency of prices, while low levels of non-block dark trading can be beneficial. One reason non-block dark trading can be harmful is that the lack of pre-trade information reduces the market's ability to infer and incorporate private information. Uninformed trades are more likely to execute in the dark, which increases adverse selection risk and bid-ask spreads in the transparent exchange.

JEL classification: G14

Keywords: dark pool, price discovery, information share, efficiency

1. Introduction

Technology has transformed trading and new trading venues, known as dark pools, have emerged. Although the term "dark trading" is new, the concept is not. Dark trading of old included block orders managed by upstairs brokers and orders in the pockets of floor brokers not yet revealed to the market. Today's dark pools systematically match orders without providing any pre-trade transparency. Technology has facilitated rapid growth in dark trading around the world. For example, dark trading's share of US consolidated volume has grown from 17% in July 2008 to 37% in June 2014 (Rosenblatt Securities). Technology has also changed the nature of dark trading, with block executions becoming less significant than non-block dark executions, due to market participants' increasing use of algorithms to execute dark trades.[1]

Many regulators and stock exchanges have expressed concern that excessive growth in dark trading may harm price discovery. For example, in a recent speech SEC Chairman White states "…we must continue to examine whether dark trading volume is approaching a level that risks seriously undermining the quality of price discovery provided by lit venues." Over the last five years many regulators have undertaken public consultations or proposed new regulations on dark trading. However, to date, only the Canadian and Australian regulators have implemented new rules.[2] The extensive consultation and subsequent lack of action by regulators reflects the uncertainty about the real costs and benefits of dark trading and the competing interests of the different participants in the market. In many cases this uncertainty is compounded by an inability to accurately identify and measure dark trading, making it difficult to assess its impact. This has also limited academic research on dark trading.

This paper is the first to focus empirically on the effect of dark trading on price discovery. This focus is consistent with the intense global regulatory concern about the

---

[1] Seppi (1990) reports that block trades accounted for roughly half of the NYSE trading volume in 1989, compared with current NYSE statistics which report 25% of current volumes are blocks. Tuttle (2013) shows that during May 2012 the distribution of trades sizes executed in dark Alternative Trading Systems is very similar to that of exchanges.

[2] For example, in November 2009 the US SEC proposed rules on the "Regulation of Non-Public Trading Interest" but to date no rule changes have been made in the US. In Europe, MiFID II proposes the introduction of a double cap on dark trading with a 4% cap on trading in a single dark venue and an 8% cap on total dark trading across all venues. The Canadian and Australian securities regulators imposed price improvement requirements for non-block dark trades on 15 October 2012 and 27 May 2013 respectively.

Electronic copy available at: http://ssrn.com/abstract=2183392

impact of dark trading on price discovery. For expositional ease, for the remainder of the paper we refer to non-block dark trading as 'dark' trading, and block dark trading as 'block' trading. We address three questions. First, what is the effect of dark trading on price discovery? Second, what is the effect of block trading on price discovery? Third, given that regulators are typically only concerned about high levels of dark trading, we examine whether the association between dark trading and price discovery is non-linear and whether there is a threshold or 'tipping point' above which dark trading is harmful.[3]

We overcome the issues in accurately observing and measuring dark trading in US and other markets by using equities data from the Australian Securities Exchange (ASX), which does not suffer from these data deficiencies. Our data enables us to precisely identify and measure all dark and block trading over a long time-series for a broad cross-section of stocks. The data are highly granular. All orders and trades are time-stamped to the millisecond and the time-stamps are consistent across the different trading mechanisms. The data also enable us to distinguish between different types of trades, allowing us to assess differences between the traditional 'upstairs' block trades and smaller dark trades executed without negotiation. This distinction is particularly important because in most markets regulators apply different rules to the two types of dark trading because they are thought to have different market impacts.

Our results show that dark and block trades have different impacts on informational efficiency. Low levels of dark trading can be beneficial, but high levels are harmful to informational efficiency. The deterioration in informational efficiency begins to occur when dark trading in a given stock exceeds approximately 10% of dollar volume. The change in informational efficiency is economically meaningful. We address the endogeneity of dark trading by using instrumental variables, and, therefore, provide evidence on the causal relation between dark trading and informational efficiency. Our results are robust to a number of control variables, hold in both large and small stocks, and in early and later parts of our sample. Our main analysis uses informational efficiency metrics calculated from intra-day observations. In robustness tests we find similar results using lower frequency metrics.

---

[3] For example, in Testimony before the Senate Banking, Housing and Urban Affairs Committee, Nasdaq OMX Chief Economist, Frank Hatheway, states that based on Nasdaq's empirical analysis "execution quality begins to deteriorate when stocks experience dark trading in excess of 40 percent of total volume."

3

In contrast, we find no evidence that block trades negotiated away from the exchange without pre-trade transparency harm informational efficiency. In fact, having some block trades execute away from the lit market (up to approximately 40% of dollar volume) can be beneficial to informational efficiency. This may be due to upstairs block brokers tapping into liquidity that would not otherwise be expressed in the limit order book. Furthermore, although block trades tend to be relatively uninformed, their large size can cause temporary price pressure if they are executed in the limit order book.

Theory suggests three reasons why high levels of dark trading can harm price discovery: (i) decreased transparency; (ii) segmentation of informed and uninformed traders; and (iii) increased fragmentation, which changes the way traders submit orders. We find evidence consistent with all three mechanisms contributing to the deterioration of informational efficiency at high levels of dark trading. First, our results are consistent with the view that although dark trades tend to be less informed than lit trades, they are not completely uninformative. The lack of pre-trade information on dark order flow may decrease the timeliness and accuracy with which the market is able to incorporate the information contained in the order flow, thereby harming price discovery. Our finding that high levels of dark trading harm price discovery while low levels may even be beneficial is consistent with the literature on the effects of transparency; in particular the notion that market quality is an increasing concave function of transparency (Eom, Ok and Park, 2007).

Second, our results support the hypothesis that dark trading leads to partial segmentation of informed and uninformed traders, as predicted by Zhu (2014). We find that orders executed in the dark tend to be less informed than orders executed in the lit market, consistent with informed traders facing lower execution probabilities in the dark than uninformed traders. By disproportionately reducing the number of uninformed trades in the lit market, dark trading increases adverse selection risk in the lit market, leading to wider bid-ask spreads, consistent with Zhu (2014). The reduction in uninformed traders in the lit market, accompanied by wider spreads, reduces incentives for costly information acquisition given that informed traders are less able to trade in the dark than uninformed traders. Therefore, dark trading may decrease the aggregate amount of information produced about fundamental values.

4

Third, we find that as dark trading increases, order book quotes take on a more important role in impounding new information compared to trade prices, consistent with liquidity providers in the lit market becoming increasingly informed. This result is consistent with the prediction of Ye (2012) that informed traders scale back the aggressiveness with which they submit orders to the lit market when they also trade in the dark. In Ye's model the reduction in informed traders' aggressiveness impedes price discovery, and therefore may also contribute to the decrease in informational efficiency that occurs at high levels of dark trading. Although Zhu (2014) and Ye (2012) make different predictions about the effect of dark trading on price discovery, our empirical results support aspects of both models.

This paper is also related to recent empirical studies of dark trading, most of which analyze the relation between dark trading and liquidity. For example, Buti, Rindi and Werner (2011) use data from 11 out of 32 US dark pools and conclude that dark pools improve spreads, depth and short-term volatility. Ready (2013) examines the determinants of volume in two block dark pools. He finds that these pools execute a lower fraction of institutional volume in stocks with higher levels of adverse selection. Nimalendran and Ray (2014) examine data from one of the 32 US dark pools and find that trading in the dark pool is associated with increased spreads and price impact on the quoting exchanges. Kwan, Masulis and McInish (2014) use a comprehensive dataset of US off-exchange trading classified into five types of dark venues to show that the un-level playing field between dark venues and exchanges increases fragmentation and has a detrimental impact on liquidity.

Our paper extends the recent empirical literature on dark trading and liquidity with an analysis of dark trading and price discovery. While both liquidity and price discovery are fundamentally important functions of the market, theory suggests that the impact of dark trading on price discovery can be entirely different to the effects on liquidity (e.g. Zhu (2014) predicts opposite effects on liquidity and price discovery).

2. Theory and previous literature

The impact of dark and block trading on price discovery is a complex issue because they simultaneously affect: (i) the level of transparency; (ii) fragmentation of

order flow across multiple trading venues; and (iii) segmentation of informed and uninformed order flow. This section reviews the theory on each of these characteristics, focussing on the effects on price discovery. Dark and block trading are considered separately because the literature suggests that they have different impacts and because market regulators tend to view these two forms of non-transparent trading differently.

2.1 Dark trading

Dark and lit trades differ in their levels of pre-trade transparency, but not post-trade transparency.[4] For example, a limit order submitted to a lit exchange is immediately visible to all market participants and thus has an immediate price impact as market participants revise their beliefs about the fundamental value.[5] In contrast, if the limit order instead rests in a dark market, no one except the order submitter can observe the order and none of the information contained in the limit order can be impounded into prices until a trade occurs. If the limit order does not execute, the market will never know about the order. Even if the order eventually executes and market participants observe the dark trade printed to the tape, the market still knows less about the order than if the order had been sent to the lit market (e.g., the time at which the limit order was submitted, the original size of the limit order, any revisions to the order price, and the venue where it was executed, all of which can be informative). Furthermore, market participants can usually determine the direction of trades in a lit market (the trade initiator) because trades generally execute at the best bid or ask price. In contrast, trades in the dark can occur within the spread, making it difficult for the market to infer the trade direction. As all of these examples illustrate, market participants observe less information about order flow sent to dark venues than order flow sent to the lit market. Because order flow conveys information (e.g., in Roşu (2013) and Kaniel and Liu (2006), both market and limit orders convey information), dark trading may have a negative impact on price discovery.

---

[4] Pre-trade transparency is the degree to which information is available to market participants about buying or selling interest, including quotes to buy or sell and the volumes available at the quotes. Post-trade transparency is the degree and timeliness of information about trades after they execute.
[5] Hautsch and Huang (2012), among others, find empirical evidence that limit orders have price impacts.

The conjecture that a reduction in transparency can harm price discovery is supported by the literature. Most, but not all, previous studies argue that transparency benefits liquidity and price discovery. For example, in various auction and dealer markets examined by Pagano and Röell (1996), pre-trade transparency allows market makers to learn information from trades more quickly, leading to more efficient prices and lower trading costs for uninformed traders. In Baruch's (2005) model, making the limit order book transparent increases the ability of market participants to compete with the specialist in liquidity provision. The increased competition increases liquidity and improves price discovery. In contrast, Boulatov and George (2013) find that hiding liquidity providing orders causes *more* aggressive competition among informed traders in providing liquidity, which improves price discovery. The empirical evidence, although not unanimous, also tends to support the view that pre-trade transparency in most circumstances has positive effects on price discovery.[6]

Eom et al. (2007) argue that market quality is an increasing concave function of pre-trade transparency, or equivalently, a *decreasing* concave function of pre-trade *opaqueness*. Dark trading increases pre-trade opaqueness in a market and therefore high levels of dark trading may harm price discovery. If market quality is a decreasing concave function of pre-trade opaqueness, low levels of dark trading should not be harmful to price discovery and may even be beneficial.

In addition to reduced transparency, the launch of new, automated dark venues has fragmented order flow and trading activity. Market participants typically use smart order routers to consider multiple venues when trying to get their orders filled and rest orders in multiple venues. Fragmentation has both positive and negative effects on market quality, including price discovery. Network externalities suggest there are benefits to consolidation.[7] When more traders use a particular market, the market's ability to match buyers and sellers increases. Consequently, trading costs decrease,

---

[6] For example, Boehmer, Saar and Yu (2005) provide empirical support for Baruch's (2005) model examining the increase pre-trade transparency resulting from the introduction of NYSE's OpenBook. However, Madhavan, Porter and Weaver (2005) report that a similar increase in pre-trade transparency in Toronto decreased liquidity. They argue that too much pre-trade transparency makes traders reluctant to post limit orders because of the increased "free option" cost. Hendershott and Jones (2005) examine trading in three exchange-traded funds and find that when Island ECN suspends the display of the limit order book overall trading costs increase and price discovery deteriorates.
[7] See, e.g., Chowdry and Nanda, 1991; Madhavan, 1995; and Hendershott and Mendelson, 2000.

which attracts more traders. Improved liquidity incentivises production of costly information and facilitates arbitrage, and thus can increase the informativeness of prices (e.g., Kyle, 1984; Chordia, Roll and Subrahmanyam, 2008). Fragmentation can also harm liquidity and price discovery by increasing search costs and thus decreasing competition between liquidity providers (e.g., Yin, 2005). Fragmentation can also benefit markets through increased competition between trading venues, which can lower trading costs (e.g., Foucault and Menkveld, 2008; Colliard and Foucault, 2012). The empirical evidence on the effects of fragmentation is mixed, mirroring the various opposing effects predicted by theory.[8]

Segmentation refers to the tendency for different types of traders to use different markets and can impact on price discovery. There are several reasons why dark trading can lead to segmentation of informed and uninformed order flow. First, at any point in time, informed traders are more likely than uninformed traders to cluster on one side of the market (either buying or selling). Consequently, informed traders face lower execution probability in the dark than uninformed traders (Zhu, 2014). Second, in some jurisdictions, including the US and Australia, dark venues are subject to lower regulatory requirements regarding fair access and consequently can discourage or exclude relatively informed order flow (Boni, Brown and Leach, 2012). Third, dark trading can make it easier for brokers to internalize order flow. Internalization of uninformed order flow is more profitable for a broker than informed order flow due to the differences in adverse selection costs. Therefore, internalization can also lead to disproportionately large share of uninformed trades taking place in the dark. A disproportionately high proportion of uninformed trades in the dark implies an increased concentration of informed traders in the lit market. It is therefore expected that adverse selection costs and spreads in the lit market increase with the level of dark trading, as predicted by Zhu (2014).

---

[8] For example, Hendershott and Jones (2005) find that fragmentation in trading for three ETFs harms liquidity and price discovery, whereas O'Hara and Ye (2011) use off-exchange volume of US stocks as a proxy for fragmentation and find that more fragmented stocks have lower transaction costs and better informational efficiency. Degryse, de Jong and van Kervel (2013) argue that fragmentation of trading across lit venues can have different impacts than fragmentation stemming from dark venues. Using a sample of Dutch stocks they show that fragmentation across visible order books improves consolidated liquidity, whereas dark trading has a detrimental effect. The different effects of lit and dark fragmentation may also be due to the other factors, including transparency and segregation of order flow.

The effects of segmentation on price discovery, however, are less clear. Different theories show that a higher concentration of informed trading can have a positive, zero or negative effect on price discovery, depending on a number of factors. A substantial decrease in uninformed traders in the lit market could harm price discovery by reducing the profitability of acquiring unique private information (e.g., Kyle, 1981, 1984, 1989). Alternatively, if all informed traders have the same piece of private information as in Zhu (2014), fewer uninformed traders in the lit market could improve price discovery. Therefore, the impact of the level of informed trading in lit venues on price discovery is ultimately an empirical question.

2.1 Block trading

In many markets large block trades are negotiated manually between brokers in the 'upstairs' market. From the perspective of a trader that participates only in the downstairs market, an upstairs block trade has no pre-trade transparency (similar to dark trades). However, from the perspective of the upstairs market trade counterparty, the block trade has greater pre-trade transparency than a lit trade, because in negotiating upstairs trades, brokers are able to signal the likely motivation for the trade, thereby reducing adverse selection risks and execution costs for large liquidity-motivated trades (Madhavan and Cheng, 1997; Bessembinder and Venkataraman, 2004).

Differences in the nature of upstairs block trading and dark trading imply they should have different consequences for price discovery. First, the upstairs market is able to facilitate trades that would not be possible in the downstairs market (Madhavan and Cheng, 1997; Bessembinder and Venkataraman, 2004). Such trades are made possible by the upstairs brokers' ability to tap into unexpressed liquidity of large institutional traders, thereby expanding the total liquidity available to the market (Grossman, 1992), and negotiate prices outside the limit order book quotes (Bessembinder and Venkataraman, 2004). In contrast, if a market for dark trades did not exist, most of the dark trades would simply execute on the lit market. By expanding the total available liquidity and facilitating trades that would not be possible in the limit order book, block trading is likely to benefit price discovery by providing additional information about the fundamental value to market participants (on a post-trade basis). Furthermore, upstairs

block trades that would have been sent to the downstairs market had they not been able to tap into unexpressed liquidity in the upstairs market, would 'walk the book', creating substantial, temporary price distortions (Bessembinder and Venkataraman, 2004). Therefore, block trades are not expected to harm price discovery.

3. Institutional setting

During our sample period, February 2008 to October 2011, the ASX was the only stock exchange operating in Australia. The ASX is one of the top ten equity markets in the world by market capitalization. There are approximately 2,200 companies listed on the ASX with a market capitalization of around AUD 1.5 trillion.

The ASX operates a transparent central limit order book (CLOB) where orders are matched based on price-then-time priority. The CLOB is anonymous, but the brokerage firms associated with each trade are reported to the market on T+3. There are no official market makers. During our sample period, the ASX Operating Rules provided two exceptions that allowed trades to be executed away from the CLOB with reduced pre-trade transparency, provided that the trades are reported to the market immediately. These exceptions include:

   i.   *Block trades* that must have a minimum value of $1 million or comprise a portfolio of trades with a combined value of at least $5 million. These trades may be negotiated away from the CLOB at any price.

   ii. *Dark trades* that have no minimum size requirement but must occur at or within the prevailing best bid or ask price. Dark trades comprise trades executed in Australian dark pools, including the ASX operated dark pool named *Centre Point* and a large number of broker operated dark pools, as well as trades manually matched away from the lit market. *Centre Point* (launched in June 2010) executes orders in price-then-time priority at the midpoint of the bid-ask spread on the CLOB. *Centre Point* orders do not interact with orders on the CLOB. At the beginning of our sample in February 2008 there were four broker-operated dark pools in Australia. This number grew to 16 by the end of our sample period.

Further details about these exceptions and dark pools in Australia are provided in the Internet Appendix.[9]

There were a number of other institutional changes that impacted dark and block trading during our sample period. On 30 November 2009 the ASX Operating Rules changed to remove the '10 second rule', which had required brokers to place an order in the CLOB for 10 seconds before executing a dark trade. This change made it easier for brokers to execute dark trades, especially when using algorithms. On 28 June 2010, the ASX Operating Rules were amended to allow dark trades to be executed at the midpoint of the best bid and ask price, as well as at the best prices.[10] On 1 July 2010, the ASX reduced its trading fees for all trade types, with a larger fee decrease for CLOB trades than for block and dark trades. Further details about these changes are provided in the Internet Appendix.

There is one important difference between the US and Australian markets that is worth noting for readers unfamiliar with the Australian market. Unlike the US market where almost all marketable retail order flow is routed to wholesale market makers,[11] in Australia, retail order flow is almost exclusively executed on the ASX.[12] This is likely influenced by the fact that payment for order flow is not permitted in Australia. This difference means that the dark order flow examined in this paper is more similar to the dark order flow executed in US dark pools, rather than the dark order flow executed by US wholesale market makers.

4. Data and descriptive statistics

Our sample comprises the constituents of the All Ordinaries Index, which includes the 500 largest (by market capitalization) ASX-listed stocks and accounts for over 95% of the total market capitalization. Our sample period extends from 1 February 2008 to 30 October 2011. The end of the sample period is chosen to avoid confounding

---

[9] The Internet Appendix is available at http://goo.gl/yIaOJR

[10] Unlike the US, dark pools may only trade at the midpoint, or at the same tick sizes as the exchange. Therefore the competition based on tick size discussed in Kwan et al. (2014) is not relevant in the Australian market.

[11] See SEC Concept Release on Equity Market Structure (2010) for further details.

[12] ASIC estimates that in September 2010 only 4% of retail order flow was executed away from the exchange.

effects from fragmentation in lit liquidity resulting from the launch of a second lit exchange, Chi-X, on 31 October 2011.

We obtain millisecond-stamped data on all trades and all CLOB and *Centre Point* orders (including order entry, amendment and cancellation messages) for our sample from the *AusEquities* database maintained by the *Securities Industry Research Centre of Asia-Pacific*. During our sample period, all trades are required to be reported to the exchange immediately. As a result, we have a single consolidated source for all trade types: lit, dark and block trades. Therefore, we minimize issues which arise in the US and other markets due to inconsistencies in time-stamps across different trading venues and inaccuracies with classification of dark and lit trade types.[13] For most of our analysis, we restrict our sample to the ASX continuous trading hours of approximately 10:00 to 16:12.[14]

We precisely classify CLOB trades and *Centre Point* trades as buyer- or seller-initiated by tracing trades back to their originating orders using the order identifiers recorded in the data. We define the trade initiator as the counterparty to a trade that submits their order last. In cases where we do not observe the original orders (block trades and dark trades other than those executed on *Centre Point*) we classify trades as buyer- (seller-) initiated if the trade price is above (below) the prevailing CLOB midquote.

Table 1 reports descriptive statistics for the stocks in our sample. The average stock-day has 1,050 trades with a total value of 9.91 million AUD. The median stock-day has substantially fewer trades with around 270 with a total value of 0.7 million AUD. Table 1 also reports that the average (median) company in the sample has a market capitalization of 2.75 billion (422 million) AUD. The average spread of 129 bps is considerably higher than the median spread of 67 bps. On average approximately 60% of

---

[13] For example, in the US the Trade Reporting Facility (TRF) is often used to proxy for dark trades. However, the TRF includes trade reports for ECNs, which are lit trading venues. Prior to BATS and Direct Edge being registered as exchanges, these lit venues accounted for substantial fractions of TRF volume. The misclassification also applies to lit trades, as exchange volumes include executions resulting from the execution of hidden orders on exchange. SEC statistics indicate that these account for around 10 to 12% of exchange volume, with substantial variation across exchanges.

[14] Opening call auctions take place at a random time within a 30 second window, and stocks commence trading in batches between 10:00 and 10:09. Closing call auctions take place in a single batch between 16:10 and 16:12 at a random time within a 60 second window.

the time stocks trade at the minimum possible spread of one tick (0.01 AUD for stock prices greater than 2 AUD). An average stock-day has around 4.6 quote messages per trade.

<center>< Insert Table 1 here ></center>

Figure 1, Panel A, provides a time-series of dark, block and total trading on the ASX over the months February 2008 to October 2011. The combined proportion of dollar volume executed using dark and block trades is approximately 18% and does not exhibit a clear trend over the period. However, from early 2010 there is an upward trend in the share of dark trades and a downward trend in the share of block trades. This reflects the growth in the number of dark pools. It also indicates that dark trades may have been used as a substitute for block trades as brokers increased their use of algorithms. In untabulated cross-sectional analysis we find that larger stocks exhibit higher levels of dark and block trading. We also find that high-volume days have higher levels of dark trading.

<center>< Insert Figure 1 here ></center>

Figure 1, Panel B, shows that average trade sizes decline substantially over our sample period for all trade types, although the rate of decline is greatest for dark trades. The average size of dark trades declines from approximately 150,000 AUD to 10,000 AUD, likely due to the increasing use of algorithms to manage executions in dark pools. Similarly, the average size of lit trades declines from approximately 13,000 AUD to 5,000 AUD, again due to increasing use of execution algorithms and growth in high-frequency trading.

5. Empirical approach

Our empirical approach involves: (i) estimating a variety of price discovery characteristics for each stock-day in our sample using intraday data; and (ii) relating the price discovery characteristics to dark and block trading via stock-day panel regressions. For the panel regressions we take two approaches: (i) one-stage OLS regression with a

<center>13</center>

range of control variables and fixed effects (stock and time); and (ii) two-stage least squares (2SLS) instrumental variables (IV) regressions.

While the IV models have the advantage of explicitly addressing the potential endogeneity of dark trading, we also report the one-stage OLS regression results for several reasons.[15] First, anecdotal accounts of how traders choose to execute trades indicate that endogeneity concerns are likely to be more severe in causally relating dark trading and liquidity, than relating dark trading and price discovery. Our empirical results support this view; the results from our IV models are qualitatively similar to the one-stage OLS results (in some cases stronger in magnitude, suggesting that endogeneity may act against us finding a significant result). Second, the one-stage OLS regression models are simpler and likely to have higher statistical power. Therefore, if the *potential* for endogeneity does not ultimately have a large impact on our estimates, the one-stage OLS regression models may be preferable due to their higher precision.

The general form of our panel regressions is:

$$y_{id} = \alpha + \beta_{DARK} DARK_{id} + \beta_{BLOCK} BLOCK_{id} + \sum_{j=1}^{6} \delta_j C_{jid} + \varepsilon_{id} \qquad (1)$$

where y_{id} is one of the price discovery characteristics for stock i on day d, and C_{jid} is a set of j control variables including log market capitalization, log bid-ask spread, the proportion of the trading day for which the stock's spread is constrained to one tick, log total dollar volume, midquote volatility (standard deviation of 1-minute midquote returns) and the messages-to-trades ratio, which serves as a proxy for algorithmic trading. $DARK_{id}$ and $BLOCK_{id}$ measure the dollar volume of dark and block trades, respectively, as a percentage of the stock-day's total dollar volume.[16] In the 2SLS IV models $DARK_{id}$ and $BLOCK_{id}$ are replaced with fitted values from first-stage regressions. In the first stage regressions, $DARK_{id}$ and $BLOCK_{id}$ are regressed on the set of instrumental

---

[15] The form of endogeneity that may be a concern is traders conditioning their order submission strategies (whether to send an order to the dark or to the lit market, which in turn determines the share of dark trading) on the price discovery characteristics. For example, Buti, Rindi and Werner (2013) show that orders are more likely to be sent to a dark pool when liquidity is higher.

[16] We use the share of dollar volume in our primary specification and in robustness tests we re-estimate the regressions using the share of trades instead of dollar volume. We also estimate the regressions using log-transformed $DARK_{id}$ and $BLOCK_{id}$ metrics, which reduce the influence of extreme values in dark and block volumes. The results are similar and our conclusions hold under both measures.

variables and the control variables, for each stock. F-tests of the null hypothesis that the instruments do not enter the first stage regression show that our tests do not suffer from the problem of weak instruments.[17]

We use two different sets of instruments for robustness, and find consistent results across the two sets. Our first set is based on market structure changes that are exogenous with respect to price discovery, but influence the amount of dark trading. This is similar to the approach used in other studies, such as Hendershott, Jones and Menkveld (2011), except that our instruments contain only time-series variation because they affect all stocks at the same points in time. The market structure changes include the removal of the 10-second rule on 30 November 2009, which made it easier to execute dark trades. We construct a dummy variable ($D_t^{NO-10SEC-RULE}$) that takes the value 1 after the change, and 0 before. The change in ASX trading fees on 1 July 2010 occurred largely in anticipation of competition from other market operators, and changed the relative explicit costs of trading in the dark compared to trading in the CLOB. Similarly, the launch of ASX's own dark pool, *Centre Point*, is unlikely to have been motivated by price discovery characteristics, but had an impact on the amount of dark trading. Because the change in trading fees and the launch of *Centre Point* took place three days apart, we construct a single dummy variable ($D_t^{NEW-FEES}$) that takes the value of 1 after both changes occurred, and 0 before. Finally, the growth in the number of dark pools from four at the start of the sample period to 16 at the end has increased the ability to automate dark executions. We construct an instrumental variable that measures the number of dark pools in operation, $DarkVenues_t$, as well as its square, $DarkVenues_t^2$.[18] Together these four variables form our first set of instruments. When using the first set of instruments we also control for a time trend in both the first- and second-stage regressions to

---

[17] Bound, Jaeger and Baker (1995, p. 446) state that "F statistics close to 1 should be cause for concern". When instrumenting for $DARK_{id}$, the average F-statistics for the first and second set of instruments are 5.69 and 4.89, respectively, both well in excess of levels that warrant concern. When instrumenting for $BLOCK_{id}$, the average F-statistics for the first and second set of instruments are 2.96 and 65.58 – the former being the lowest of the four F-statistics because the first set of instrumental variables are chosen specifically to instrument for the level of *dark* trading, not necessarily *block* trading.

[18] Our results are robust to different combinations of these variables, including omitting the variable $DarkVenues_t^2$.

minimize the possibility that the instruments pick up any general trends in dark and block trading.

For our second set of instruments we follow Hasbrouck and Saar (2013) and Buti et al. (2011) and instrument the level of dark trading in a stock-day with the average level of dark trading on that day in all other stocks in the corresponding size (market capitalization) quartile. This variable meets the requirements for an instrument because the level of dark trading in other stocks is correlated with the level of dark trading in a particular stock (95% confidence interval for the pooled Pearson correlation coefficient is 0.154 to 0.160), and dark trading in other stocks is unlikely to be driven by the nature of price discovery in the particular stock. Similarly, we instrument the level of block trading with the average level of block trading on that day in all other stocks in the corresponding size quartile. When using the second set of instruments we include an additional control variable in both the first- and second-stage regressions ($Y_{OtherStocks}$), which measures the average of the corresponding dependent variable on that day for all other stocks in the corresponding size quartile. This variable helps the instruments isolate the causal effect of dark and block trading, by removing potential reverse causality that can arise from cross-sectional commonality in market characteristics.

6. Informational efficiency

We start with the key question of interest to regulators, which is also the source of conflicting theoretical predictions; namely, how does dark and block trading impact the absolute amount of information that is impounded in prices? We use three types of informational efficiency measures commonly used in empirical studies: autocorrelation-based measures, variance ratios, and measures of short-term return predictability using lagged market returns. We calculate the informational efficiency measures each stock-day using intra-day data. Details of the calculation and interpretation of these measures are provided in Appendix A.

Using high-frequency informational efficiency metrics is important to maximize the statistical power of our tests. Rösch, Subrahmanyam and van Dijk (2013) provide evidence that such informational efficiency metrics measured at intraday horizons are highly correlated with low-frequency measures of informational efficiency, and are

different from liquidity measures. Anderson, Eom, Hahn and Park (2013) find that partial price adjustment (slow price adjustment and overshooting), which implies a degree of informational inefficiency, is a major source of positive and negative autocorrelations. In robustness tests we confirm that our results hold at lower frequencies (estimating the measures for each stock-month using daily data), although such tests have lower statistical power and less precision.

Table 2 reports panel regression estimates of the relation between the share of dark and block trading and informational efficiency, using one-stage OLS. The general pattern that emerges is that an increase in the share of dark trading, all else equal, is associated with deterioration in informational efficiency. The coefficients of $DARK_{id}$ are positive for all measures of informational inefficiency, and statistically significant for all specifications with and without fixed effects. The R^2s, which exclude the variation explained by the fixed effects, tend to be lower for specifications that include stock fixed effects than time fixed effects, indicating that the explanatory variables are able to explain a greater fraction of the cross-sectional variation in informational efficiency than the time-series variation. Overall, the results in Table 2 provide evidence that informational efficiency is harmed by dark trading.

< Insert Table 2 here >

Block trading, however, according to our evidence is not detrimental to informational efficiency. For all three informational efficiency measures the coefficients of $BLOCK_{id}$ are negative and they are statistically significant for two of the three measures, with and without fixed effects. This provides evidence that trading large blocks off-exchange may even be beneficial to the efficiency of the lit market. Previous studies identify two possible reasons why block trades may have a different effect on the market compared to dark trades. First, through the unique role of upstairs brokers as 'information repositories' block trades are able to tap into additional liquidity that would not otherwise be expressed in the limit order book, thereby expanding aggregate liquidity (Grossman, 1992; Bessembinder and Venkataraman, 2004). Second, block trades are largely uninformed, but due to their size they would cause significant temporary price

distortions if submitted to the limit order book. By being able to credibly signal the likely motivation for the trade in the upstairs market, a block trade's counterparty faces lower adverse selection risk, allowing the block trade to occur with a smaller price impact (Bessembinder and Venkataraman, 2004).

Turning to the 2SLS IV models of the impact of dark and block trading on informational efficiency (reported in Table 3), we find similar results. Across all three informational inefficiency measures dark trading is associated with a statistically significant deterioration in informational efficiency, whereas block trading is estimated to have the opposite effect. These results have a similar level of statistical significance as the one-stage OLS models and tend to have larger impacts. This provides evidence that our results relating dark trading to deterioration in informational efficiency are not driven by traders choosing to execute in the dark when informational efficiency is poor. If anything, the magnitudes suggest that endogeneity may work against finding significant results in our OLS regressions. These results provide evidence of a causal link from dark trading to deterioration of informational efficiency. The positive and highly significant coefficients on the control variable $Y_{OtherStocks}$ show significant cross-sectional commonality in informational efficiency. For two of the three efficiency measures the time trend is insignificant and for the third it is negative and marginally significant. We obtain similar results in specifications that omit the control variables.

< Insert Table 3 here >

The autocorrelations and variance ratios at the different frequencies (the components of $Autocorrelation_{Factor}$ and $VarianceRatio_{Factor}$) provide results that are consistent with those using $Autocorrelation_{Factor}$ and $VarianceRatio_{Factor}$. We also examine the autocorrelations and variance ratios without the absolute value transformation, i.e., allowing them to take positive and negative values. In our sample the autocorrelations and variance ratios tend to be negative – the pooled means are statistically significantly negative (ranging from -0.04 to -0.15) and even the 75th percentile values are negative for all of the different frequencies. The estimated effect of dark trading is to make the autocorrelations and variance ratios more negative (these

18

results are not tabulated), consistent with a decrease in informational efficiency. An interpretation of these results is that prices (midquotes) tend to overreact to new information or order flow and subsequently reverse the overreaction (Anderson et al., 2013), and high levels of dark trading tend to exacerbate the inefficient overreactions and reversals. Negative autocorrelations also arise from imperfect risk-bearing capacity of liquidity providers (e.g., Ho and Stoll, 1981). Thus, some of the decrease in informational efficiency associated with high levels of dark trading may be due to decreases in liquidity.

The effects of dark and block trading on informational efficiency may be nonlinear in their share of volume. For example, Eom et al. (2007) argue that market quality is an increasing concave function of transparency. This implies that an increase in dark trading from a low level is likely to have a smaller effect on market quality (and may even improve market quality) than the same magnitude increase from a relatively high level of dark trading. To investigate this possibility we estimate an alternative version of the stock-day panel regression in which we replace the continuous variables $DARK_{id}$ and $BLOCK_{id}$ with a series of dummy variables that measure dark trading (D_{id}^{range}) and block trading (B_{id}^{range}) over various ranges:

$$
\begin{aligned}
y_{id} = \alpha &+ \beta_1 D_{id}^{0-5\%} + \beta_2 D_{id}^{5-10\%} + \beta_3 D_{id}^{10-20\%} + \beta_4 D_{id}^{20-30\%} + \beta_5 D_{id}^{30-40\%} + \beta_6 D_{id}^{>40\%} \\
&+ \gamma_1 B_{id}^{0-5\%} + \gamma_2 B_{id}^{5-10\%} + \gamma_3 B_{id}^{10-20\%} + \gamma_4 B_{id}^{20-30\%} + \gamma_5 B_{id}^{30-40\%} + \gamma_6 B_{id}^{>40\%} + \sum_{j=1}^{6} \delta_j C_{jid} + \varepsilon_{id}
\end{aligned} \quad (2)
$$

The omitted, reference category corresponds to zero dark and zero block trading. As an example of how the dummy variables are defined, $D_{id}^{0-5\%}$ takes the value 1 if the share of stock-day id's dollar volume executed in the dark is $0 < DARK_{id} \leq 5\%$, and 0 otherwise. The dummy variable $B_{id}^{10-20\%}$ takes the value 1 if the share of stock-day id's dollar volume executed as block trades is $10\% < BLOCK_{id} \leq 20\%$. Therefore, the coefficients of the dummy variables estimate the effect of different levels of dark/block trading relative to the case of no dark/block trading. This specification is able to characterize many forms of non-linearity that would be difficult to fit with a polynomial. Our robustness tests indicate that the results are not sensitive to the choice of ranges. The

patterns are similar when we use one-stage OLS and 2SLS IV approaches, and, therefore, to save space we only report results using the one-stage OLS approach.

Figure 2 plots the coefficients of the dummy variables for each of the three informational efficiency metrics together with error bounds corresponding to +/- two standard errors.[19] Panel A suggests that low levels of dark trading are not harmful to informational efficiency (they may even be beneficial), but as dark trading increases it eventually reaches a 'tipping point' after which it has a negative impact. Specifically, after controlling for other stock characteristics, when dark trading accounts for approximately 10% of total dollar volume its impact on informational efficiency is very close to zero, i.e., it neither harms nor benefits informational efficiency. However, levels of dark trading above 10% of dollar volume are associated with lower informational efficiency compared to zero dark trading.

To illustrate the economic significance, an increase in dark trading from 10% to 20% of dollar volume is estimated to increase the informational inefficiency measures by 10% to 15% of a standard deviation using the one-stage OLS model, and 19% to 26% using the 2SLS IV models. A more modest increase in dark trading from 10% to 12.5% of dollar volume is expected to increase the informational inefficiency measures by 2% to 4% of a standard deviation using the one-stage OLS model, and 6% to 7% using the 2SLS IV models.

< Insert Figure 2 here >

Figure 2, Panel B, provides evidence that executing block trades away from the CLOB improves informational efficiency, but only up to a certain point. Maximum informational efficiency occurs around the point where block trades account for approximately 15% of total dollar volume. Beyond this level additional block trades tend to have a negative *marginal* impact on informational efficiency, although the *total* impact on informational efficiency remains positive until block trades account for approximately 40% of total dollar volume. Block trading at 15% of dollar volume is associated with

---

[19] The range covered by each dummy variable is reduced to a single point for the purpose of the plots by taking the mean of $DARK_{id}$ and $BLOCK_{id}$ for the stock-days that fall into the corresponding range. For example, for stock-days that have dark dollar volume greater than zero but less than or equal to 5%, the mean of $DARK_{id}$ is 1.7%. Therefore $D_{id}^{0-5\%}$ is plotted at the horizontal axis value of $DARK_{id} = 1.7\%$.

improvements in the informational efficiency measures of approximately 14% to 21% of a standard deviation using the one-stage OLS models, and 5% to 14% using the 2SLS IV models. In general, small amounts of block trading away from the lit market is good for informational efficiency, but as with dark trading: too much can be harmful.

Our finding that high levels of dark trading harm price discovery is not inconsistent with Zhu's (2014) prediction that in equilibrium, adding a dark crossing system alongside a lit exchange will improve price discovery, because *high* levels are not necessarily *equilibrium* levels. Therefore, it is interesting to examine for how many stocks the current levels of dark trading are harmful to price discovery. During the last ten months of our sample (January-October 2011) the median level of dark trading as a share of dollar volume was greater than 10% for 62 of the 498 stocks (12% of stocks). This illustrates that approximately 12% of stocks had levels of dark trading that were harmful to price discovery on most (>50%) trading days during the first 10 months of 2011. On average these stocks are larger, more actively traded and more likely to have a constrained spread than the other stocks in the sample. Approximately one third of the stocks in our sample have harmful levels of dark trading (>10% of dollar volume) on more than one quarter of the trading days in 2011. No stocks have block trading levels in excess of the 40% 'tipping point' for more than one quarter of the trading days in 2011. Overall, these results demonstrate that block trading on a typical day is below harmful levels in all stocks, and dark trading on a typical day is below harmful levels for most stocks in our sample.

The calculations above also illustrate an important point: the 'tipping points' suggested by our analysis correspond to stock-day levels, and, therefore, should not be compared to market-wide aggregates of dark trading. The tendency for high-volume stock-days to have higher levels of dark trading means that market-wide aggregates of dark trading (effectively volume weighted averages) are typically higher than the median and the equal-weighted mean levels of dark trading.

7. Informativeness of different types of trades and spreads

Theory provides several reasons why high levels of dark trading can harm price discovery, including the effects of reduced transparency, increased fragmentation and

segmentation of order flow. The remainder of our analysis aims to provide insights about these various channels, and test some of the more specific predictions made by models of dark and lit trading. This section examines the informativeness of lit, dark and block trades.

Ye (2012) argues that the mechanism via which dark trading harms price discovery is that an informed trader scales back the aggressiveness of his trading in the lit market to make larger profits in the dark. This means informed traders will execute a considerable share of their trades in the dark and therefore dark trades should be relatively informative. In contrast, Zhu (2014) predicts that informed traders face a lower execution probability in the dark compared to uninformed traders and therefore uninformed traders will execute a disproportionately higher share of their trades in the dark. It therefore follows from Zhu's model that dark trades should be less informative than lit trades. Dark trades may also be less informed than lit trades due dark pools deliberately excluding certain types of relatively informed traders, or due to broker internalization of relatively uninformed order flow.

There are also other reasons why both informed and uninformed traders might be attracted to relatively non-transparent trading venues; for uninformed traders the lack of transparency can help reduce "picking off" risks and exploitation by predatory traders, while for informed traders a lack of transparency can help prevent information leakage. Therefore, whether relatively more or less informed trades occur in the dark is an empirical question, and one that has implications for how price discovery occurs and how adverse selection risk changes in response to dark trading.

To measure the informativeness of different trade types (lit compared to dark and block) we adapt the Hasbrouck (1991) vector auto-regression (VAR) framework to our trade type partition. We calculate signed dollar volume of lit, dark and block trades, x_t^{LIT}, x_t^{DARK} and x_t^{BLOCK}, in every 1-second interval, t, for every stock-day. For each stock-day we estimate the following system:

$$x_t^{LIT} = \mu^{LIT} + \sum_{i=1}^{60} \phi_i^r r_{t-i} + \sum_{i=1}^{60} \phi_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \phi_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \phi_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{LIT}$$

$$x_t^{DARK} = \mu^{DARK} + \sum_{i=1}^{60} \theta_i^r r_{t-i} + \sum_{i=1}^{60} \theta_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \theta_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \theta_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{DARK}$$

$$x_t^{BLOCK} = \mu^{BLOCK} + \sum_{i=1}^{60} \lambda_i^r r_{t-i} + \sum_{i=1}^{60} \lambda_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \lambda_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \lambda_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{BLOCK} \qquad (3)$$

$$r_t = \mu^r + \sum_{i=1}^{60} \gamma_i^r r_{t-i} + \sum_{i=0}^{60} \gamma_i^{LIT} x_{t-i}^{LIT} + \sum_{i=0}^{60} \gamma_i^{DARK} x_{t-i}^{DARK} + \sum_{i=0}^{60} \gamma_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^r$$

where t indexes 1-second intervals (individual stock and date subscripts are suppressed) and r_t is the log-midquote change in the tth interval.

After estimating the above system for each stock-day, we calculate the informativeness of lit, dark and block volume as the cumulative impulse response (measured 60 seconds forward in time) of midquote returns for a shock of +$10,000 of signed lit, dark, and block volume, respectively, holding all other variables equal to their unconditional means. Following Hasbrouck (1991) we interpret the permanent price impact of order flow as a measure of the private information contained in the order flow. In order to minimize the effects of outliers we winsorize the permanent price impact measures by setting extreme positive and negative values to the 1st and 99th percentile values, for each stock and each date.

Table 4 reports permanent price impacts of lit, dark and block trades. The average permanent price impact of lit trades is larger than that of dark trades (3.62 bps per $10,000 for lit trades versus 3.31 bps for dark trades with equal weighting of stock-days, and 0.65 bps for lit versus 0.41 bps for dark with dollar volume weighting). The difference in lit and dark price impacts is statistically distinguishable at the 1% level using paired t-tests. The average permanent price impacts of lit and dark trades are both higher than the average permanent price impact of block trades (0.15 bps with equal weighting and 0.02 bps with dollar volume weighting) and these differences are also statistically significant at the 1% level.[20] The median permanent price impact for lit

---

[20] The price impacts presented in this section are all per $10,000 of volume. Because block trades are much larger than lit trades, the total price impact of a block trade is larger than the total price impact of a lit trade. The relatively low informativeness of block trades (per unit volume) is consistent with previous studies that find upstairs markets tend to be used by traders who can credibly signal that their trades are uninformed (e.g., Madhavan and Cheng, 1997; Bessembinder and Venkataraman, 2004; Booth, Lin, Martikainen and Tse, 2002).

trades (1.91 bps) is considerably higher than the medians for dark and block trades (0.03 bps and 0.01 bps) and the difference in medians is statistically significant at the 1% level using paired sign tests. Therefore, a 'typical' (median) lit trade contains considerably more private information per unit of volume than dark and block trades. On average, dark and block trades do contain some information, and in particular some dark trades are highly informed.

< Insert Table 4 here >

These results do not rule out either of the mechanisms modeled in Ye (2012) or Zhu (2014). On one hand, some dark trades contain considerable private information about the fundamental value. If the traders that are responsible for the privately informed dark trades also trade in the lit market, then it is plausible that as predicted by Ye (2012) they may scale back the aggressiveness of their trading in the lit market in order to avoid imposing a negative externality on their profits from trading in the dark. While the evidence in this section is consistent with the mechanism modeled by Ye, the evidence is rather indirect. We conduct some more specific tests in the next section.

On the other hand, the price impacts in Table 4 illustrate that 'typical' dark and block trades are less informed than 'typical' lit trades. This finding is consistent with the predictions of Zhu (2014) that a relatively larger proportion of uninformed trades will execute in the dark because they are less likely to cluster on one side of the market compared to informed trades.[21] Zhu predicts that the higher proportion of uninformed trades in the dark will leave behind a higher concentration of informed traders in the lit market, which will result in higher adverse selection risk and wider spreads in the lit market. Therefore, to provide some further evidence on the mechanism modeled by Zhu we examine how spreads in the lit market are impacted by dark trading.

< Insert Table 5 here >

---

[21] An alternative explanation why dark trades tend to be less informed compared to lit trades is that dark trading venues are more selective in what order flow they accept and thereby screen out some informed traders. For example, dark pools have selective membership whereby only relatively uninformed traders are allowed access (see Boni et al., 2012), and brokers that know their clients are able to be selective in what order flow they internalize and what they send to a lit market for execution. These explanations are not mutually exclusive.

Consistent with the predictions of Zhu (2014) and the notion that a larger share of uninformed trades will execute in the dark, Table 5 indicates that quoted spreads become wider in the lit limit order book as dark and block trading increase. The impact of dark trading on spreads in the lit market is highly statistically significant across all of our regression specifications: one-stage OLS without fixed effects, with stock fixed effects, with date fixed effects, and 2SLS IV regressions using two different sets of instruments. We find similar results using spread measured in cents, rather than spread expressed in basis points.

The magnitude of the increase in quoted spreads is also economically meaningful. For example, estimates using the one-stage OLS regression model with dummy variables for different levels of dark and block trading (equation (2)) show that increasing dark trading from zero to 10% of dollar volume is expected to increase quoted spreads by 11% after controlling for other factors. This means that for the average stock spreads will increase from 128 bps to 142 bps. A more modest increase in dark trading from 10% to 12.5% is expected to increase spreads by 2.2% (an increase of 2.8 bps for the average stock). Again, the 2SLS IV estimates are larger in magnitude; for example, an increase in dark trading from 10% to 12.5% is expected to increase spreads by 6.5% to 7.2%, depending on which set of instruments is used (an increase of 8.4 bps to 9.3 bps for the average stock). Similarly, an increase in block trading from 10% to 12.5% of dollar volume is expected to increase spreads by 1.7% using the one-stage OLS model and 1.7% to 4.0% using the 2SLS IV models. Wider spreads increase the costs of trading in the lit market, which can encourage order flow to migrate away from the lit market in a self-reinforcing spiral.

The segregation of order flow that is apparent from the differences in informativeness of different trade types may be one of the reasons why high levels of dark trading harm price discovery. For example, suppose information acquisition is endogenous and costly, and traders that choose to become informed receive unique noisy signals (e.g., the fundamental value plus an independent error). In such a setting a decrease in the amount of uninformed trading decreases the profitability of acquiring information. This leads to less information production in aggregate and therefore less

25

informative prices (e.g., Kyle, 1981, 1984, 1989; Admati and Pfleiderer, 1988).[22] As uninformed traders leave the lit market to trade in the dark, the profitability of acquiring information decreases (because informed traders tend to cluster on one side of the market they are not able to trade in the dark to the same extent as uninformed traders). Thus, fewer traders choose to become informed (and/or informed traders acquire less costly, less precise information), which reduces the aggregate amount of private information and the informativeness of prices.

The tendency for less informed traders to trade in the dark (and in doing so avoid interacting with some informed traders) has important welfare implications. Although our results provide evidence that high levels of dark trading widen spreads in the lit market, this does not necessarily increase trading costs in aggregate because higher trading costs in the lit market may be offset by lower trading costs in the dark. As dark trading activity increases, costs associated with non-execution and delayed execution decrease in the dark. Although the impact on aggregate trading costs is not clear, dark trading leads to redistributions (transfers) of trading costs across different types of traders. The increase in trading costs in the lit market is largely borne by the informed traders that are less able to trade in the dark.

In the seminal models of Kyle (1985) and Glosten and Milgrom (1985) uninformed market makers break even on average and informed traders profit from trading on their information. As a result, uninformed traders on average lose an amount equal to the informed traders' profits. The wealth transfer from uninformed traders to informed traders occurs through the trading costs faced by uninformed traders: the bid-ask spread in Glosten and Milgrom (1985) and trade prices away from fundamental value in Kyle (1985). Importantly, the wealth transfer from uninformed traders to informed traders compensates them for the costs of producing information and thereby providing price discovery. In fact, when information acquisition is costly, the absence of uninformed ('noise') traders can cause a complete breakdown of price discovery resulting

---

[22] Zhu (2014) arrives at a different prediction due to different assumptions about the nature of information acquisition. In Zhu's model all traders who choose to become informed receive an identical piece of information: exact knowledge of the fundamental value. Under this assumption, a decrease in the number of informed traders (due to fewer uninformed traders in the lit market) corresponds to a decrease in the degree of *competition* on the same set of private information, but no change in the *amount* of private information that in aggregate is held by informed traders. Thus, Zhu's result is driven by differences in the degree of competition on the same information rather than the amount of private information produced.

in an informationally inefficient market (Grossman and Stiglitz, 1980; Black, 1986). Therefore, the trading costs paid by uninformed traders play an important role in facilitating price discovery by compensating others for producing information. Our results are consistent with the notion that uninformed traders benefit from trading with each other in the dark, but their gain comes at the cost of less information production and therefore less informative prices.

8. Price discovery shares

In this section we analyze how and where information enters the market, and how this process is impacted by dark and block trading. This provides some more insights about the mechanisms underpinning our earlier results; in particular, whether informed traders scale back their aggressiveness in the lit market (as predicted by Ye (2012)) making quotes relatively more informative than trade prices, and to what extent dark and block trade prices contribute to price discovery.

Two traditional approaches to measuring the contributions of different markets or types of order flow to price discovery are Hasbrouck's (1995) information share (*IS*) and Gonzalo and Granger's (1995) common factor share (*CS*). Fundamentally, both methods decompose price innovations into permanent and temporary components. As pointed out by Yan and Zivot (2010) and Putniņš (2013), both metrics measure (with different weights) a combination of two dimensions of market efficiency: (i) timeliness in impounding of new information; and (ii) avoidance of transitory shocks. For the purpose of identifying where information enters the market, we are interested in measuring the first component: the extent to which a price or order flow type it is the first to impound new information about the 'true' underlying asset value. Of the two traditional measures Hasbrouck's IS_i comes closer to identifying the leader in impounding new information, but is also influenced by the relative amount of noise in the price channels (Putniņš, 2013). To isolate the relative speed at which information is impounded by a price series from its relative level of noise Putniņš (2013) extends the analytic results of Yan and Zivot (2010) and defines the "information leadership share" (*ILS*) as:

$$ILS_1 = \frac{\begin{vmatrix} IS_1 & CS_2 \\ IS_2 & CS_1 \end{vmatrix}}{\begin{vmatrix} IS_1 & CS_2 \\ IS_2 & CS_1 \end{vmatrix} + \begin{vmatrix} IS_2 & CS_1 \\ IS_1 & CS_2 \end{vmatrix}}, \qquad ILS_2 = \frac{\begin{vmatrix} IS_2 & CS_1 \\ IS_1 & CS_2 \end{vmatrix}}{\begin{vmatrix} IS_1 & CS_2 \\ IS_2 & CS_1 \end{vmatrix} + \begin{vmatrix} IS_2 & CS_1 \\ IS_1 & CS_2 \end{vmatrix}} \qquad (4)$$

The ILS_1 and ILS_2 each have the range [0,1] (and they sum to 1), similar to IS and CS, with values above (below) 0.5 indicating the price series impounds new information faster (slower) than the other price series and thereby leads (does not lead) the process of price discovery. Using simulations, Putniņš (2013) shows that ILS is robust to differences in noise levels and therefore correctly attributes price discovery in a wider range of settings. Therefore, we report results using ILS (results using IS are available upon request).

Estimation of the information share metrics relies on price series being co-integrated. In studies of cross-listed stocks the law of one price keeps the two prices of the stock within certain arbitrage limits and therefore ensures co-integration. In this paper, we study the contribution of limit order book quotes as well as prices of different types of trades (lit and dark/block) for each stock within one market (similar to Anand and Subrahmanyam (2008)). The limit order book quotes, lit trade prices and dark/block trade prices for a stock are all linked to the fundamental value of the stock and are therefore co-integrated.

Following Hasbrouck (1995), we estimate the following vector error correction model (VECM) for each stock-day using 1-second intervals, t:

$$\Delta p_{1,t} = \alpha_1 (p_{1,t-1} - p_{2,t-1}) + \sum_{i=1}^{60} \gamma_i \Delta p_{1,t-i} + \sum_{j=1}^{60} \delta_j \Delta p_{2,t-j} + \varepsilon_{1,t}$$

$$\Delta p_{2,t} = \alpha_2 (p_{1,t-1} - p_{2,t-1}) + \sum_{k=1}^{60} \phi_k \Delta p_{1,t-k} + \sum_{m=1}^{60} \varphi_m \Delta p_{2,t-m} + \varepsilon_{2,t} \qquad (5)$$

where $p_{1,t}$ and $p_{2,t}$ are the last available log prices of price series 1 and 2, respectively.

We estimate two different versions of the VECM above. In the first, the two price series are: (i) midquotes, calculated from the prevailing best bid and ask prices; and (ii) trade prices, using the last available trade price irrespective of the trade type. This version allows us to analyze the contribution to price discovery made by the best quotes (pre-trade information), compared to trade prices (post-trade information). In the second

version the two price series are: (i) lit trade prices; and (ii) dark/block trade prices. This version allows us to analyze the relative contribution of post-trade information about lit trades compared to dark and block trades. We calculate IS_1, IS_2 and CS_1, CS_2 from the error correction parameters and variance-covariance of the error terms, following Baillie, Booth, Tse and Zabotina (2002), and ILS_1 and ILS_2 following Putniņš (2013).

We examine the information leadership share of midquotes compared to trade prices because this metric tells us about the extent to which liquidity providers (the limit order traders that set the best quotes in the market) are informed compared to liquidity demanders (traders that initiate trades by submitting market orders). Goettler, Parlour and Rajan (2009) point out that the informativeness of the best quotes relative to the informativeness of trade prices depends on the order submission strategies of informed and uninformed traders. For example, if informed traders tend to demand liquidity and trade with market orders and uninformed traders tend to be liquidity providers then trade prices will convey relatively more information about the fundamental value than quotes. If informed traders begin supplying liquidity as well as consuming it, the informativeness of quotes will increase relative to that of trade prices.

The information leadership shares for the full sample show that in the median stock-day the midquote has a slightly larger contribution to impounding new information about the underlying fundamental value compared to trade prices (median $ILS_{MIDQUOTE}$ of 0.56). This indicates informed traders often use limit orders and provide liquidity, consistent with other studies.[23]

Table 6 reports how dark trading impacts the information leadership share of midquotes, using regressions similar to those used in our earlier analysis. Recall that Ye (2012) predicts dark trading will harm price discovery because informed traders will scale back the aggressiveness of their trading in the lit market to avoid overly impacting the profits earned by their dark trades. In a limit order market setting, this would involve informed traders increasing their use of limit orders and reducing their use of market orders, thereby increasing the informativeness of midquotes compared to trade prices. This is indeed what the results in Table 6 show. An increasing share of dark trading is

---

[23] For example, Bloomfield, O'Hara and Saar (2005), Kaniel and Liu (2006), Goettler et al. (2009), Roşu (2013), and Boulatov and George (2013).

associated with an increase in the *ILS* of the midquote, holding other variables fixed. This is true in the single-stage OLS specifications as well as the 2SLS IV specifications. The effect of block trading is opposite, consistent with the fact that block trade prices are not mechanically derived from the prices in the lit market unlike the prices of many dark trades, and block trades tend to be less informed.

< Insert Table 6 here >

The tendency for midquotes to become relatively more informative (suggesting that informed traders increasingly supply liquidity in the lit market) as the share of dark trading increases is also consistent with our earlier results on how dark trading impacts adverse selection. As Rindi (2008) and others point out, informed traders are particularly effective liquidity suppliers when adverse selection risks are high because of their informational advantage. Therefore, the increasing informativeness of midquotes could result from increased adverse selection risk in the lit market (as a disproportionately large share of uninformed trades execute in the dark) causing a higher proportion of informed traders to act as liquidity suppliers.

Turning to the price discovery shares of lit and dark/block trades, on average, lit limit order book trades contribute substantially more than dark and block trades to impounding new information (median ILS_{LIT} of 0.84).[24] This is because lit trades account for a larger share of volume than dark and block trades but also because typically lit trades are more informative. The results in Table 6 indicate that as the share of block trading increases, lit trades contribute relatively less to impounding new information, compared to dark and block trades. The impact of dark trading is less clear, because the coefficient on the variable *DARK* is only statistically distinguishable from zero in two of the five regression specifications (in those specifications the coefficient is negative, similar to the effect of block trades).

---

[24] The median of ILS_{LIT} does not take into consideration the stock-days on which we are unable to compute the information shares of lit and dark/block trade prices (days when there are zero or very few dark/block trades) and therefore understate the contribution of lit trades to price discovery overall.

As the share of volume executed in the dark (or executed as blocks) increases, the price discovery share of lit trades should decrease somewhat mechanistically.[25] If the price discovery shares changed proportionally to the volume shares we would expect the coefficients on *DARK* and *BLOCK* to be around -1, indicating that an *x*% increase in the share of dark or block volume leads to an *x*% decrease in the share of price discovery attributable to lit trades. However, in the regressions of ILS_{LIT} (Table 6), the coefficients on *DARK* and *BLOCK* across the five specifications range from -0.41 to 0.03. Therefore, as dark and block trading increase, their contribution to price discovery increases less than proportional to their volume share. This indicates dark and block trades on average contain less private information than lit trades consistent with our previous results, and/or the market is less able to infer and incorporate the private information of dark trades.

9. Robustness and subsample tests

In this section we detail a range of additional robustness tests. First, we examine whether our results hold for stocks of different sizes. We estimate our full set of analyses separately for large and small stocks (defined as market capitalization above and below the median, respectively). We find that our key results hold for both subsamples, in particular, as dark trading increases informational efficiency deteriorates and spreads on the lit market increase. Large stocks also have a substantially lower proportion of stock-day observations with zero dark trading than small stocks. The consistency of results across the two groups therefore also provides some evidence of the robustness of the results to the proportion of zero dark trading observations included in the sample.

The results are similar in both the first and second halves of the sample period (2008-2009 and 2010-2011). This indicates that the potentially harmful effects of dark trading are not a new phenomenon. Given the changes in how dark trading takes place (increasing automation during the sample period due to an increasing number of dark pools) this result provides indirect evidence that the *amount* of dark trading matters for

---

[25] Price discovery occurs through two channels: (i) public information entering the market and causing a revision in quotes; and (ii) private information being impounded into prices via trades (Hasbrouck, 1991). In a market where price discovery occurs predominantly through public information and trades are largely liquidity-motivated, the price discovery shares of different volume types should increase roughly in proportion to their share of volume (Anand and Subrahmanyam, 2008). This is because when a trade type accounts for a greater share of volume it will more often by chance be the first to reflect innovations in the quotes and thus changes in the fundamental value.

price discovery rather than the way in which dark trading takes place. Dark pools as such are not necessarily any more harmful than manual dark trading. However, if dark pools make it easier to trade in the dark they may encourage growth of dark trading to levels that are harmful to price discovery.

We also find that high levels of dark trading are associated with a deterioration of informational efficiency when we replace the high-frequency measures with lower-frequency measures. We measure midquote autocorrelations, variance ratios and the delay in incorporating market-wide information for each stock-month using daily data. We calculate autocorrelations using 1-day, 2-day and 3-day returns, variance ratios that compare the variance of 1-day returns with those of 2-day, 3-day and 4-day returns, and the extent to which daily midquote returns can be predicted using 10 lags of daily market returns. Across all of these lower frequency measures we find negative and statistically significant relations between dark trading and informational efficiency after controlling for other variables, consistent with the results from the high-frequency measures.

We examine alternative measures of dark/block trading activity, using number of trades instead of dollar volume, as well as log-transforms of the dark and block trading shares, and find similar results. Changes to the number of lags used in the VAR, VECM and return predictability regressions do not have a substantial impact on our results. We also estimate the VAR and VECM models at lower frequency using 10-second intervals in place of 1-second intervals, allowing the lags to span a ten times larger window of past observations, and find qualitatively similar results.

Estimation of the VAR requires at least one lit trade, dark trade and block trade, as well as changes in the midquote. This requirement is met and the VAR is successfully estimated for approximately 81,000 stock-days (from a total of approximately 408,000). Because many stock-days do not contain block trades we also estimate a simpler version of the VAR in which we pool dark and block trades into a single volume category. This allows greater coverage across the sample (approximately 223,000 stock-days). We also estimate a version of VAR in which we sign trades as buyer/seller initiated using only information that is readily available to market participants: trades with price above (below) the prevailing midquote are classified as buyer (seller) initiated and trades at the midquote are discarded. Our main results are robust to these alternative specifications.

32

10. Conclusions

Our results provide evidence that high levels of dark trading harm price discovery and lead to less informationally efficient prices. Order flow that executes in the dark tends to be less informed than the trades that execute in the lit market. Therefore, by disproportionately reducing the number of uninformed trades in the lit market, dark trading increases adverse selection risk and the lit market's bid-ask spreads, consistent with the theoretical predictions of Zhu (2014). The increased adverse selection risk and trading costs in the lit market increases the incentives for order flow to migrate away from the lit market, potentially leading to a self-reinforcing spiral. As dark trading increases, order book quotes take on a more important role in impounding new information compared to trade prices, consistent with informed traders scaling back the aggressiveness with which they submit lit orders (as predicted by Ye (2012)) and thus liquidity providers in the lit market becoming increasingly informed. Informed traders in the lit market naturally would like to trade with the less informed order flow in the dark, but their ability to do so is limited by lower execution probability in the dark due to their tendency to cluster on one side of the market as in Zhu (2014) and due to exclusivity of some dark pools where the operators limit participation to relatively uninformed clientele (e.g., Boni et al., 2012). Our results indicate that the impact of dark trading occurs through the joint effects of reduced transparency, increased fragmentation and segmentation of order flow.

Together, the results provide support for the concerns of regulators that high levels of dark trading harm informational efficiency and price discovery. High levels of dark trading are found to be harmful throughout the sample period, in large and in small stocks. This does not, however, mean that dark trading in general or even in equilibrium is harmful. Our results indicate that low levels of dark trading do not have a negative impact on informational efficiency and may even be beneficial. For most stocks in our sample, the level of dark trading on a typical day is below harmful levels. This result has important policy implications. It provides evidence that regulatory action should consider the level of dark trading in specific stocks, rather than the aggregate market level of dark trading. Regulatory changes such as those proposed in MiFID II to cap dark

trading at 4% per venue and 8% for the European Union overall may have unintended consequences.

We find no evidence that block trades negotiated without pre-trade transparency harm informational efficiency. Block trades differ from dark trades in that upstairs brokers' unique role as 'information repositories' allows block trades to tap into additional liquidity that would not otherwise be expressed in the limit order book, thereby expanding aggregate liquidity (Grossman, 1990; Bessembinder and Venkataraman, 2004). Furthermore, block trades are largely uninformed, but due to their size they would cause significant temporary price distortions if submitted to the limit order book. In the upstairs market, a block broker can reduce adverse selection risk for the trade's counterparty by signaling the motivation for the trade, thereby reducing price impact and avoiding the temporary price distortions that would occur in the limit order book (Bessembinder and Venkataraman, 2004). Again, this result has important policy implications, providing evidence that regulation of dark trading needs to be carefully designed to account for the fact that not all dark trading has the same effects on price discovery.

Appendix A: Informational efficiency measures

Our informational efficiency metrics follow the existing empirical literature. They measure the extent to which prices deviate from a random walk and/or are predictable using past information. In a perfectly informationally efficient, frictionless market, prices at all times equal the fundamental value (the expected value of the stock given all available information). Prices change only due to the arrival of new information. Because new information is unpredictable by definition, prices follow a martingale. Thus price changes should not be predictable using past information (such as lagged stock or market returns), returns should have zero autocorrelation, and (with the additional assumption that innovations in the log fundamental are *i.i.d.* Gaussian) the variance of returns should increase linearly with the return horizon.

Many different imperfections and sources of inefficiency in real markets cause deviations from the characteristics of perfectly efficient markets. Trade prices have negative autocorrelation due to the mechanical effect of bid-ask bounce (e.g., Roll, 1984). This liquidity-related effect is not, however, present in midquotes. Midquotes can (and do) deviate from the characteristics expected under perfectly efficient markets, and importantly, all of the mechanisms that causes such deviations imply some degree of informational inefficiency. For example, inventory control models such as Ho and Stoll (1981) show that inventory management by liquidity providers causes negative autocorrelation in midquote returns. Risk-averse liquidity providers that hold positive inventory will adjust quotes upward so that the midquote is greater than the fundamental value to attract sellers and reduce their inventory position. The opposite occurs when liquidity providers are short. The negative autocorrelation in midquotes (and ability to predict returns using past order flow) is associated with midquote deviations from fundamental values, i.e., informational inefficiency. As another example, consider price formation models such as the sequential trade model of Glosten and Milgrom (1985) and the Kyle (1985) model with repeated auctions. In both models, prices are initially efficient with respect to public information and gradually approach strong-form efficiency as private information is revealed through the course of trading. Convergence of prices toward the full-information values (expected fundamental value given all public *and* private information) causes positive short-run autocorrelation in ex-post midquote

returns. Importantly, the autocorrelation occurs as a result of prices being less that fully informationally efficient; once prices reflect all public and private information returns no longer display autocorrelation, and if private information were revealed instantaneously rather than gradually, returns would also have zero autocorrelation. Finally, both under- and over-reaction to information, as well as delayed reaction to information (e.g., non-synchronous trading, stale prices) cause midquotes to deviate from the characteristics expected under perfect efficiency. In summary, the many different mechanisms that cause midquotes to deviate from the characteristics expected under a perfectly efficient market imply some degree of informational efficiency.

Both positive and negative midquote autocorrelations imply less than perfect informational efficiency. We calculate first-order return autocorrelations for each stock-day, at various intraday frequencies, $k \in \{10\,\text{sec},\ 30\,\text{sec},\ 60\,\text{sec}\}$, similar to Hendershott and Jones (2005):

$$Autocorrelation_k = Corr(r_{k,t}, r_{k,t-1}) \qquad\qquad \text{(A.1)}$$

where $r_{k,t}$ is the tth midquote return of length k for a stock-day (stock-day subscripts are suppressed). Taking the absolute value of the autocorrelation gives a measure of informational efficiency that captures both under- and over-reaction of returns to information, with larger values indicating greater inefficiency. We compute a combined autocorrelation measure, $Autocorrelation_{Factor}$, by taking the first principal component of the absolute autocorrelations at the three frequencies, and then scaling the measure so that it ranges from 0 (highly efficient) to 100 (highly inefficient). Using the first principal component is a way of summarizing the results across informational efficiency metrics calculated at different frequencies and can help reduce error in the individual proxies.[26]

If a stock's price follows a random walk, the variance of its returns is a linear function of the measurement frequency, i.e., $\sigma^2_{k-periodReturn}$ is k times larger than

---

[26] We expect the informational efficiency metrics calculated at different frequencies to correlate with the underlying latent variable (informational inefficiency) but each will also contain some measurement error. Therefore, the metrics at different frequencies will have some common variance arising from the variance in informational inefficiency. The first principal component is the linear combination of the different frequency metrics that explains the maximal amount of common variance and thus should be closely related to the underlying latent variable, while containing less noise if the measurement errors are less than perfectly correlated across the different frequencies.

$\sigma^2_{1-periodReturn}$. The variance ratio exploits this property to measure inefficiency as a price series' deviation from the characteristics that would be expected under a random walk (e.g., Lo and MacKinlay, 1988). We calculate three variance ratios for each stock-day at different intra-day frequencies:

$$VarianceRatio_{kl} = \left| \frac{\sigma^2_{kl}}{k\sigma^2_l} - 1 \right| \tag{A.2}$$

where σ^2_l and σ^2_{kl} are the variances of l-second and kl-second midquote returns for a given stock-day. We use the (l, kl) combinations: (1-sec, 10-sec), (10-sec, 60-sec), (1-min, 5-min). We compute a combined variance ratio, $VarianceRatio_{Factor}$, by taking the first principal component of the three variance ratios, and then scaling the measure so that it ranges from 0 (highly efficient) to 100 (highly inefficient).

Our third measure of informational efficiency is an intraday adaptation of the Hou and Moskowitz (2005) *Delay*, i.e., the extent to which lagged market returns predict a stock's midquote returns. For each stock-day we estimate a regression of 1-minute midquote returns for stock i, $r_{i,t}$, on the All Ordinaries market index return, $r_{m,t}$, and ten lags (suppressing day subscripts):

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \sum_{k=1}^{10} \delta_{i,k} r_{m,t-k} + \varepsilon_{it} \tag{A.3}$$

We save the R^2 from the above unconstrained regression, $R^2_{Unconstrained}$, and re-estimate the regression constraining the coefficients on lagged market returns to zero, $\delta_{i,k} = 0, \forall k$, again saving the R^2, $R^2_{Constrained}$. *Delay* is then calculated as:

$$Delay = 100 \left(1 - \frac{R^2_{Constrained}}{R^2_{Unconstrained}} \right) \tag{A.4}$$

and takes values between 0 and 100. The larger this measure, the more variation in stock returns is explained by lagged market returns, which implies more sluggish incorporation of market-wide information into the stock's price, and, therefore, lower informational efficiency.

References

Admati, A.R., and P. Plfeiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 1, 3-40.

Anand, A., and A. Subrahmanyam, 2008, Information and the intermediary: Are market intermediaries informed traders in electronic markets?, *Journal of Financial and Quantitative Analysis* 43, 1–28.

Anderson, R.M., K.S. Eom, S.B. Hahn, and J.-H. Park, 2013, Autocorrelation and partial price adjustment, *Journal of Empirical Finance* 24, 78-93.

Australian Securities and Investments Commission, Report 331, Dark liquidity and high frequency trading, March 2013.

Australian Securities and Investments Commission Consultation Paper 168, Australian equity market structure: Further proposals, October 2011.

Australian Securities and Investments Commission Consultation Paper 179, Australian market structure: Draft market integrity rules and guidance, June 2012.

Baillie, R.T., G.G. Booth, G., Y. Tse, and T. Zabotina, 2002, Price discovery and common factor models, *Journal of Financial Markets* 5, 309-321.

Baruch, S., 2005, Who benefits from an open limit order book?, *Journal of Business* 78, 1267–1306.

Bessembinder, H., and K. Venkataraman, 2004, Does an electronic stock exchange need an upstairs market?, *Journal of Financial Economics* 73, 3-36.

Black, F., 1986, Noise, *Journal of Finance* 41, 529-543.

Bloomfield, R., M. O'Hara, and G. Saar, 2005, The "make or take" decision in an electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165-199.

Boehmer, E., G. Saar, and L. Yu, 2005, Lifting the veil: An analysis of pre-trade transparency at the NYSE, *Journal of Finance* 60, 783–815.

Boni, L., D.C. Brown, and J.C. Leach, 2012, Dark pool exclusivity matters, Working paper.

Booth, G.G., J. Lin, T. Martikainen, and Y. Tse, 2002, Trading and pricing in upstairs and downstairs stock markets, *Review of Financial Studies* 15, 1111-1135.

Boulatov, A., and T.J. George, 2013, Hidden and displayed liquidity in securities markets with informed liquidity providers, *Review of Financial Studies* 26, 2095-2137.

Bound, J., D.A. Jaeger, and R.M. Baker, 1995, Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak, *Journal of the American Statistical Association* 90, 443-450.

Buti, S., B. Rindi, and I.M. Werner, 2011, Diving into dark pools, Working paper.

Buti, S., B. Rindi, and I.M. Werner, 2013, Dark pool trading strategies, market quality and welfare, Working paper.

Chordia, T., R. Roll, and A. Subrahmanyam, 2008, Liquidity and market efficiency, *Journal of Financial Economics* 87, 249-268.

Chowdhry, B., and V. Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies* 4, 483–511.

Colliard, J.-E., and T. Foucault, 2012, Trading fees and efficiency in limit order markets, *Review of Financial Studies* 25, 3389-3421.

Degryse, H., F. de Jong and V. van Kervel, 2013, The impact of dark trading and visible fragmentation on market quality, Working paper.

Eom, K.S., J. Ok, and J. Park, 2007, Pre-trade transparency and market quality, *Journal of Financial Markets* 10, 319–341.

Foucault, T., and A.J. Menkveld, 2008, Competition for order flow and smart order routing systems, *Journal of Finance* 63, 19–58.

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

Goettler, R.L., C.A. Parlour, and U. Rajan, 2009, Informed traders and limit order markets, *Journal of Financial Economics* 93, 67-87.

Gonzalo, J., and C. Granger, 1995, Estimation of common long-memory components in cointegrated systems, *Journal of Business and Economic Statistics* 13, 27-35.

Grossman, S., 1992, The informational role of upstairs and downstairs markets, *Journal of Business* 65, 509–529.

Grossman, S., and Stiglitz, J., 1980, On the impossibility of informationally efficient markets, *American Economic Review* 10, 393-408.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179-207.

Hasbrouck, J., 1995, One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50, 1175-1199.

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

Hautsch, N., and R. Huang, 2012, The market impact of a limit order, *Journal of Economic Dynamics and Control*, 36, 501–522.

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

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

Hendershott, T., and H. Mendelson, 2000, Crossing networks and dealer markets: Competition and performance, *Journal of Finance* 55, 2071- 2115.

Ho, T.S.Y., and H.R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics* 9, 47-73.

Hou, K., and T.J. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981-1020.

Kaniel, R., and H. Liu, 2006, So what orders do informed traders use?, *Journal of Business* 79, 1867-1913.

Kyle, A.S., 1981, An equilibrium model of speculation and hedging, Doctoral dissertation (University of Chicago).

Kyle, A.S., 1984, Market structure, information, futures markets, and price formation, in *International agricultural trade: Advanced readings in price formation, market structure, and price instability*, ed. by Gary G. Storey, Andrew Schmitz, and Alexander H. Sarris. Boulder and London: Westview Press, 45-64.

Kyle, A.S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Kyle, A.S., 1989, Informed speculation with imperfect competition, *Review of Economic Studies* 56, 317-356.

Kwan, A., T.H. McInish, and R.Masulis, 2014, Trading rules, competition for order flow and market fragmentation, *Journal of Financial Economics* (forthcoming).

Lo, A., and C. MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41-66.

Madhavan, A., 1995, Consolidation, fragmentation, and the disclosure of trading information, *Review of Financial Studies* 8, 579-603.

Madhavan, A., and M. Cheng, 1997, In search of liquidity: Block trades in the upstairs and downstairs markets, *Review of Financial Studies* 10, 175-203.

Madhavan, A., D. Porter, and D. Weaver, 2005, Should securities markets be transparent?, *Journal of Financial Markets* 8, 266–288.

Nimalendran, M., and S. Ray, 2014, Informational linkages between dark and lit trading venues, *Journal of Financial Market* 17, 230-261.

O'Hara, M., and M. Ye, 2011, Is market fragmentation harming market quality?, *Journal of Financial Economics* 100, 459-474.

Pagano, M., and A.A. Röell, 1996, Transparency and liquidity: A comparison of auction and dealer markets with informed trading, *Journal of Finance* 51, 579–611.

Putniņš, T.J., 2013, What do price discovery metrics really measure?, *Journal of Empirical Finance* 23, 68-83.

Ready, M.J., 2013, Determinants of volume in dark pools, Working paper.

Rindi, B., 2008, Informed traders as liquidity providers: Anonymity, liquidity and price formation, *Review of Finance* 12, 497-532.

Roll, R., 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 39, 1127-1139.

Rösch, D.M., A. Subrahmanyam, and M.A. van Dijk, 2013, An empirical analysis of co-movements in high- and low-frequency measures of market efficiency, Working paper.

Roşu, I., 2013, Liquidity and information in order driven markets, Working paper.

Securities and Exchange Commission (SEC), 2010, Concept Release on Equity Market Structure, Release # 34-61358, File # S7-02-10.

Seppi, D., 1990, Equilibrium block trading and asymmetric information, *Journal of Finance* 45, 73–94.

Tuttle, L., 2013, Alternative Trading Systems: Description of ATS Trading in National Market System Stocks, Division of Economic and Risk Analysis, Securities and Exchange Commission.

Yan, B., and E. Zivot, 2010, A structural analysis of price discovery measures, *Journal of Financial Markets* 13, 1-19.

Ye, M., 2012, Price manipulation, price discovery and transaction costs in the crossing network, Working paper.

Yin, X., 2005, A comparison of centralized and fragmented markets with costly search, *Journal of Finance* 60, 1567–1590.

Zhu, H., 2014, Do dark pools harm price discovery?, *Review of Financial Studies* 27, 747-789.

Table 1
Descriptive statistics

This table reports means, standard deviations and quartile points (*P25*, *Median*, *P75*) of variables calculated at the stock-day level. Total volume consists of *Lit* trades (trades executed in the transparent central limit order book), *Dark* trades (trades executed without pre-trade transparency below block size), and *Block* trades (large trades executed without pre-trade transparency). *Constrained spread* measures the proportion of trading day for which the stock's spread is constrained to one tick size. *Midquote volatility* is the standard deviation of 1-minute midquote returns. *Message-to-trade* is the ratio of number of order messages (including order entry, amendment and cancellation) to the number of trades.

	Mean	Std. dev.	P25	Median	P75
Volumes and trades					
Total $ volume ($ mil)	9.91	38.75	0.12	0.70	4.49
Total trades (count)	1,050	1,959	41	268	1,267
Stock characteristics					
Market capitalization ($ million)	2,749	9,482	193	422	1,553
Quoted spread (bps)	129	172	32	67	158
Constrained spread	0.60	0.36	0.29	0.71	0.93
Midquote volatility (bps)	16.61	14.64	8.01	12.77	20.60
Message-to-trade (ratio)	4.58	34.48	2.56	3.56	4.90

Table 2
Effects of dark and block trading on aggregate price discovery (market efficiency)

This table reports regression estimates using a stock-day panel, in which the dependent variables are estimates of market informational inefficiency, which range from 0 (perfect efficiency) to 100 (complete inefficiency). $Autocorrelation_{Factor}$ and $VarianceRatio_{Factor}$ are the first principle components of absolute autocorrelations of midquote returns and variance ratios at different intraday frequencies. *Delay* measures intraday midquote return predictability using lagged market returns. *DARK* and *BLOCK* are the percentage of the stock-day's total dollar volume executed without pre-trade transparency below block size and at block size, respectively. *Market capitalization*, *Quoted spread* (time-weighted average of the stock-day's limit order book proportional quoted spread) and *Total $ volume* (comprising dark, block and lit limit order book volume) are in logs. *Constrained spread* measures the proportion of trading day for which the stock's spread is constrained to one tick size. *Midquote volatility* is the standard deviation of 1-minute midquote returns. *Message-to-trade* is the ratio of number of quote messages (including order entry, amendment and cancellation) to the number of trades. The regression model is estimated for each dependent variable without fixed effects, with stock fixed effects and with date fixed effects. R^2 estimates exclude the variance explained by the fixed effects. Standard errors are clustered both by stock and by date and t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

	$Autocorrelation_{Factor}$			$VarianceRatio_{Factor}$			*Delay*		
Intercept	0.251	-0.046	-0.091	1.980	-0.024	-0.068	92.905	0.029	0.171
	(0.22)	(-2.33)**	(-1.18)	(2.49)**	(-1.35)	(-1.09)	(36.02)***	(0.47)	(1.04)
DARK	0.042	0.018	0.041	0.029	0.012	0.030	0.048	0.004	0.048
	(16.84)***	(8.63)***	(16.57)***	(17.75)***	(8.07)***	(17.67)***	(8.72)***	(1.83)*	(8.95)***
BLOCK	-0.013	-0.018	-0.015	-0.006	-0.011	-0.009	-0.002	-0.001	-0.006
	(-5.15)***	(-8.53)***	(-6.19)***	(-3.55)***	(-7.50)***	(-5.13)***	(-0.26)	(-0.27)	(-1.02)
Market capitalization	-0.334	-0.571	-0.333	-0.328	-0.396	-0.321	-1.923	-0.633	-1.830
	(-2.74)***	(-4.79)***	(-2.70)***	(-3.82)***	(-4.54)***	(-3.71)***	(-6.79)***	(-3.85)***	(-6.55)***
Quoted spread	-0.120	-0.542	-0.130	-0.290	-0.391	-0.326	3.028	1.697	3.267
	(-0.90)	(-4.18)***	(-0.99)	(-3.22)***	(-5.51)***	(-3.66)***	(10.19)***	(8.64)***	(11.79)***
Constrained spread	-0.553	-2.267	-0.584	-0.397	-1.950	-0.316	7.442	4.597	6.563
	(-2.28)**	(-12.82)***	(-2.32)**	(-2.05)**	(-14.55)***	(-1.55)	(12.71)***	(9.82)***	(11.43)***
Total $ volume	0.817	1.144	0.826	0.661	0.880	0.647	-0.754	0.115	-0.674
	(13.80)***	(23.45)***	(14.10)***	(16.39)***	(37.78)***	(15.91)***	(-5.98)***	(2.22)**	(-6.08)***
Midquote volatility	0.023	0.027	0.023	-0.002	-0.003	-0.004	-0.067	-0.034	-0.041
	(3.75)***	(3.53)***	(3.43)***	(-0.95)	(-1.94)*	(-2.20)**	(-4.91)***	(-5.53)***	(-4.64)***
Message-to-trade	0.004	0.004	0.004	0.004	0.003	0.004	0.000	0.000	0.001
	(1.59)	(1.52)	(1.60)	(1.60)	(1.51)	(1.61)	(1.31)	(0.41)	(2.17)**
R^2	0.06	0.04	0.06	0.10	0.04	0.10	0.17	0.01	0.18
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Fixed effects	None	Stock	Date	None	Stock	Date	None	Stock	Date

Table 3
Instrumental variable regression results for aggregate price discovery (market efficiency)

This table reports estimates from a two-stage least squares regression (2SLS), using two different sets of instrumental variables. *DARK* and *BLOCK* are the percentage of the stock-day's total dollar volume executed without pre-trade transparency below block size and at block size, respectively. In the first stage *DARK* and *BLOCK* are regressed on the instrumental variables and control variables. The first set of instrumental variables comprise a dummy variable for the removal of the 10-second rule that restricted dark trading ($D_t^{NO\_10SEC\_RULE}$), a dummy variable for a change in exchange fees and the introduction of *Centre Point* ($D_t^{NEW\_FEES}$), and the number of dark pools in operation, *DarkVenues*$_t$ as well as its square, *DarkVenues*$_t^2$. The second set of instrumental variables ($DARK_i^{NOT}$ and $BLOCK_i^{NOT}$) are the average of *DARK* and *BLOCK*, respectively, on the same day for all other stocks in the relevant size (market capitalization) quartile. In the second stage we regress each of the dependent variables on fitted values of *DARK* and *BLOCK* from the first-stage regressions, and control variables. The dependent variables are estimates of market informational inefficiency, which range from 0 (perfect efficiency) to 100 (complete inefficiency). *Autocorrelation*$_{Factor}$ and *VarianceRatio*$_{Factor}$ are the first principle components of absolute autocorrelations of midquote returns and variance ratios at different intraday frequencies. *Delay* measures intraday midquote return predictability using lagged market returns. *Market capitalization*, *Quoted spread* (time-weighted average of the stock-day's limit order book proportional quoted spread) and *Total $ volume* (comprising dark, block and lit limit order book volume) are in logs. *Constrained spread* measures the proportion of trading day for which the stock's spread is constrained to one tick size. *Midquote volatility* is the standard deviation of 1-minute midquote returns. *Message-to-trade* is the ratio of number of quote messages (including order entry, amendment and cancellation) to the number of trades. *Time* is a linear time trend starting at 0 and incrementing by 1 for every date in our sample. $Y_{OtherStocks}$ is the average of the dependent variable (*Autocorrelation*$_{Factor}$, *VarianceRatio*$_{Factor}$, or *Delay*), on the same day for all other stocks in the relevant size quartile. Standard errors are clustered both by stock and by date and t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

	$Autocorrelation_{Factor}$		$VarianceRatio_{Factor}$		$Delay$	
Intercept	3.052	0.803	4.984	1.949	97.233	49.310
	(2.41)**	(0.62)	(6.06)***	(2.30)**	(37.57)***	(8.12)***
DARK	0.142	0.136	0.106	0.099	0.247	0.262
	(15.91)***	(16.48)***	(16.47)***	(18.28)***	(10.75)***	(10.04)***
BLOCK	-0.104	-0.083	-0.045	-0.043	-0.273	-0.210
	(-5.17)***	(-6.61)***	(-2.96)***	(-4.63)***	(-5.14)***	(-5.47)***
Market capitalization	-0.215	-0.534	-0.243	-0.616	-1.680	-0.515
	(-1.87)*	(-4.15)***	(-3.05)***	(-6.65)***	(-6.82)***	(-1.71)*
Quoted spread	-0.463	-0.430	-0.581	-0.500	2.393	2.476
	(-3.05)***	(-2.82)***	(-6.49)***	(-5.28)***	(8.92)***	(9.27)***
Constrained spread	-0.576	-0.649	-0.270	-0.538	7.279	7.086
	(-2.52)**	(-2.94)***	(-1.48)	(-3.12)***	(12.95)***	(13.26)***
Total $ volume	0.596	0.582	0.446	0.466	-1.087	-1.072
	(7.55)***	(7.52)***	(9.20)***	(9.71)***	(-7.63)***	(-7.36)***
Midquote volatility	0.037	0.033	0.008	0.006	-0.044	-0.029
	(3.81)***	(3.77)***	(2.97)***	(2.50)**	(-4.51)***	(-4.36)***
Message-to-trade	0.004	0.004	0.003	0.003	0.000	0.000
	(1.57)	(1.57)	(1.57)	(1.56)	(-0.39)	(-0.39)
Time	0.000		0.000		0.000	
	(1.05)		(-2.02)**		(0.47)	
$Y_{OtherStocks}$		0.501		0.648		0.456
		(15.4)***		(18.56)***		(10.94)***
R^2	0.07	0.08	0.11	0.13	0.18	0.20
Estimation method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Instrumental variables	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2

Table 4
Informativeness of trade types

This table reports means, standard deviations and quartile points (*P25*, *Median*, *P75*) of trade informativeness variables calculated at the stock-day level. $PriceImpact_{LIT}$, $PriceImpact_{DARK}$, and $PriceImpact_{BLOCK}$ are the permanent price impacts (bps/$10,000) of lit, dark and block volume calculated from the cumulative impulse response functions from a vector auto-regression model. *Lit* trades are trades executed in the transparent central limit order book, *Dark* trades are trade executed without pre-trade transparency below block size), and *Block* trades are large trades executed without pre-trade transparency.

	Mean (equal weighted)	Mean ($volume weighted)	Std. dev.	P25	Median	P75
$PriceImpact_{LIT}$	3.62	0.65	4.94	0.69	1.91	4.74
$PriceImpact_{DARK}$	3.31	0.41	27.94	-0.10	0.03	1.28
$PriceImpact_{BLOCK}$	0.15	0.02	2.53	-0.02	0.01	0.11

Table 5
Effects of dark and block trading on the bid-ask spread

This table reports regression estimates using a stock-day panel, in which the dependent variable is the log time-weighted average proportional quoted bid-ask spread in the central limit order book. The key independent variables, *DARK* and *BLOCK* are the percentage of the stock-day's total dollar volume executed without pre-trade transparency below block size and at block size, respectively. We report five models: (i) one-stage OLS, (ii) one-stage OLS with stock fixed effects; (iii) one-stage OLS with date fixed effects; (iv) two-stage least squares (2SLS) using the first set of instruments (a dummy variable for the removal of the 10-second rule that restricted dark trading, a dummy variable for a change in exchange fees and the introduction of *Centre Point*, and the number of dark pools in operation, as well as its square); and (v) 2SLS using the second set of instruments (the average of *DARK* and *BLOCK* on the same day for all other stocks in the relevant size quartile). In the first stage of the 2SLS models we regress *DARK* and *BLOCK* on the instrumental variables and control variables, and in the second stage we regress each of the dependent variables on fitted values of *DARK* and *BLOCK* from the first-stage regressions, and control variables. *Market capitalization* and *Total $ volume* (comprising dark, block and lit limit order book volume) are in logs. *Constrained spread* measures the proportion of trading day for which the stock's spread is constrained to one tick size. *Midquote volatility* is the standard deviation of 1-minute midquote returns. *Message-to-trade* is the ratio of number of quote messages (including order entry, amendment and cancellation) to the number of trades. *Time* is a linear time trend starting at 0 and incrementing by 1 for every date in our sample. $Y_{OtherStocks}$ is the average of the dependent variable ($Autocorrelation_{Factor}$, $VarianceRatio_{Factor}$, or *Delay*), on the same day for all other stocks in the relevant size quartile. R^2 estimates exclude the variance explained by the fixed effects. Standard errors are clustered both by stock and by date and t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

|	Log quoted spread				
Intercept	8.492	-4.254	-4.273	8.651	7.589
	(77.37)***	(-102.4)***	(-211.1)***	(72.91)***	(25.51)***
DARK	0.006	0.002	0.006	0.019	0.021
	(11.05)***	(8.77)***	(11.00)***	(10.47)***	(10.49)***
BLOCK	0.004	0.002	0.004	-0.005	-0.005
	(7.52)***	(2.99)***	(7.84)***	(-1.10)	(-1.82)*
Market capitalization	-0.216	-0.518	-0.210	-0.189	-0.130
	(-6.56)***	(-16.12)***	(-6.54)***	(-5.66)***	(-3.51)***
Constrained spread	-0.117	-1.022	-0.040	-0.066	-0.057
	(-1.89)*	(-24.08)***	(-0.71)	(-1.11)	(-0.91)
Total $ volume	-0.233	-0.093	-0.241	-0.261	-0.260
	(-12.68)***	(-8.92)***	(-13.53)***	(-12.57)***	(-12.70)***
Midquote volatility	0.018	0.008	0.017	0.019	0.019
	(5.52)***	(4.51)***	(4.89)***	(5.23)***	(5.38)***
Message-to-trade	0.000	0.000	0.000	0.000	0.000
	(-0.09)	(2.23)**	(0.08)	(-0.87)	(-1.10)
Time				0.000	
				(-3.69)***	
$Y_{OtherStocks}$					0.136
					(3.18)***
R^2	0.74	0.15	0.68	0.75	0.75
Estimation method	OLS	OLS	OLS	2SLS	2SLS
Fixed effects	None	Stock	Date	None	None
Instrumental variables	None	None	None	Set 1	Set 2

Table 6
Effects of dark and block trading on price discovery shares

This table reports regression estimates using a stock-day panel, in which the dependent variables are the information leadership share ($ILS_{MIDQUOTE}$) of midquotes relative to trade prices, and the information leadership share of lit trade prices (trades executed in the transparent central limit order book) relative to dark and block trade prices (ILS_{LIT}). Both $ILS_{MIDQUOTE}$ and ILS_{LIT} are scaled up by a factor of 100. The key independent variables, *DARK* and *BLOCK* are the percentage of the stock-day's total dollar volume executed without pre-trade transparency below block size and at block size, respectively. For each dependent variable we report five models: (i) one-stage OLS, (ii) one-stage OLS with stock fixed effects; (iii) one-stage OLS with date fixed effects; (iv) two-stage least squares (2SLS) using the first set of instruments (a dummy variable for the removal of the 10-second rule that restricted dark trading, a dummy variable for a change in exchange fees and the introduction of *Centre Point*, and the number of dark pools in operation, as well as its square); and (v) 2SLS using the second set of instruments (the average of *DARK* and *BLOCK* on the same day for all other stocks in the relevant size quartile). In the first stage of the 2SLS models we regress *DARK* and *BLOCK* on the instrumental variables and control variables, and in the second stage we regress each of the dependent variables on fitted values of *DARK* and *BLOCK* from the first-stage regressions, and control variables. *Market capitalization*, *Quoted spread* (time-weighted average of the stock-day's limit order book proportional quoted spread) and *Total $ volume* (comprising dark, block and lit limit order book volume) are in logs. *Constrained spread* measures the proportion of trading day for which the stock's spread is constrained to one tick size. *Midquote volatility* is the standard deviation of 1-minute midquote returns. *Message-to-trade* is the ratio of number of quote messages (including order entry, amendment and cancellation) to the number of trades. *Time* is a linear time trend starting at 0 and incrementing by 1 for every date in our sample. $Y_{OtherStocks}$ is the average of the dependent variable ($Autocorrelation_{Factor}$, $VarianceRatio_{Factor}$, or *Delay*), on the same day for all other stocks in the relevant size quartile. R^2 estimates exclude the variance explained by the fixed effects. Standard errors are clustered both by stock and by date and t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

	$ILS_{MIDQUOTE}$					ILS_{LIT}				
Intercept	66.321	-0.099	0.092	79.291	37.347	36.177	-0.803	-7.101	29.953	14.468
	(21.21)***	(-1.31)	(0.57)	(24.35)***	(8.07)***	(11.57)***	(-5.65)***	(-37.52)***	(8.29)***	(4.09)***
DARK	0.085	0.029	0.090	0.311	0.321	-0.004	-0.021	-0.015	-0.023	-0.024
	(13.35)***	(5.53)***	(15.11)***	(13.83)***	(12.78)***	(-0.51)	(-3.35)***	(-2.18)**	(-0.99)	(-0.99)
BLOCK	-0.013	-0.035	-0.007	-0.285	-0.220	-0.129	-0.126	-0.090	-0.131	-0.287
	(-1.48)	(-4.47)***	(-0.83)	(-5.56)***	(-6.21)***	(-15.39)***	(-14.80)***	(-11.55)***	(-2.99)***	(-9.65)***
Market capitalization	-0.692	-0.533	-0.742	-0.479	0.416	0.510	0.830	0.675	0.643	-0.652
	(-2.60)***	(-1.34)	(-2.84)***	(-1.95)*	(1.64)	(1.84)*	(1.51)	(2.33)**	(2.16)**	(-2.31)**
Quoted spread	0.444	0.458	0.217	-0.545	-0.548	-3.810	-3.020	-3.537	-3.489	-3.417
	(1.04)	(0.91)	(0.58)	(-1.29)	(-1.20)	(-10.11)***	(-4.95)***	(-9.42)***	(-8.37)***	(-9.38)***
Constrained spread	-3.689	-4.849	-2.155	-2.090	-2.114	9.017	5.513	8.203	7.711	7.511
	(-6.20)***	(-9.12)***	(-3.55)***	(-3.70)***	(-3.65)***	(12.31)***	(5.16)***	(12.50)***	(11.16)***	(10.46)***
Total $ volume	-0.643	0.603	-0.826	-1.239	-1.252	2.767	2.057	2.624	2.941	2.940
	(-3.25)***	(3.03)***	(-4.67)***	(-5.31)***	(-5.01)***	(14.25)***	(11.41)***	(13.86)***	(12.56)***	(14.48)***
Midquote volatility	0.137	0.140	0.119	0.134	0.147	0.059	0.044	0.057	0.072	0.057
	(4.53)***	(4.11)***	(4.16)***	(4.26)***	(4.28)***	(2.27)**	(2.11)**	(2.11)**	(2.21)**	(2.25)**
Message-to-trade	0.001	0.001	0.002	0.002	0.001	0.001	0.022	-0.043	-0.054	-0.030
	(0.65)	(0.66)	(0.99)	(0.83)	(0.66)	(0.04)	(1.19)	(-2.28)**	(-2.58)***	(-1.56)
Time				-0.009					0.005	
				(-15.19)***					(5.53)***	
$Y_{OtherStocks}$					0.571					0.376
					(19.23)***					(10.08)***
R^2	0.04	0.02	0.03	0.05	0.06	0.11	0.02	0.11	0.12	0.13
Estimation method	OLS	OLS	OLS	2SLS	2SLS	OLS	OLS	OLS	2SLS	2SLS
Fixed effects	None	Stock	Date	None	None	None	Stock	Date	None	None
Instrumental variables	None	None	None	Set 1	Set 2	None	None	None	Set 1	Set 2

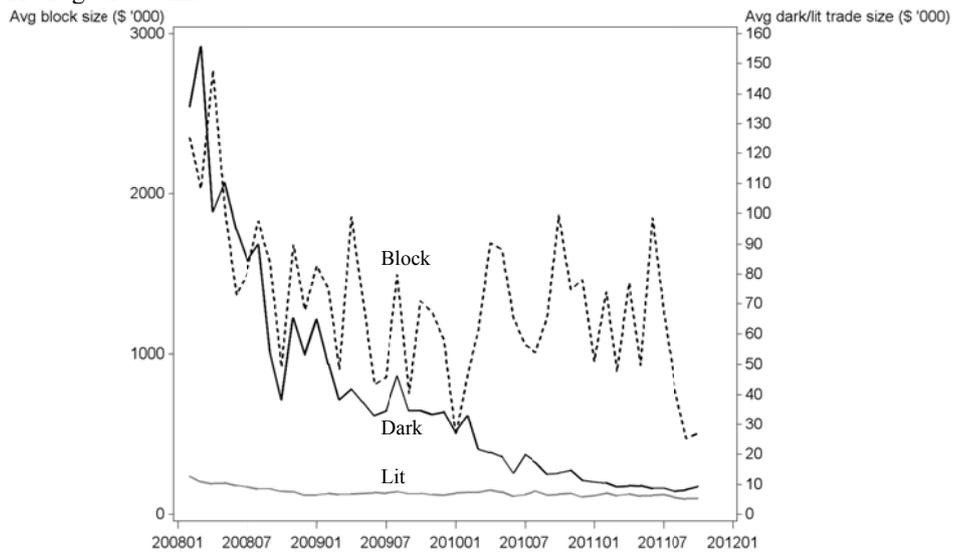Panel A: Dollar volume
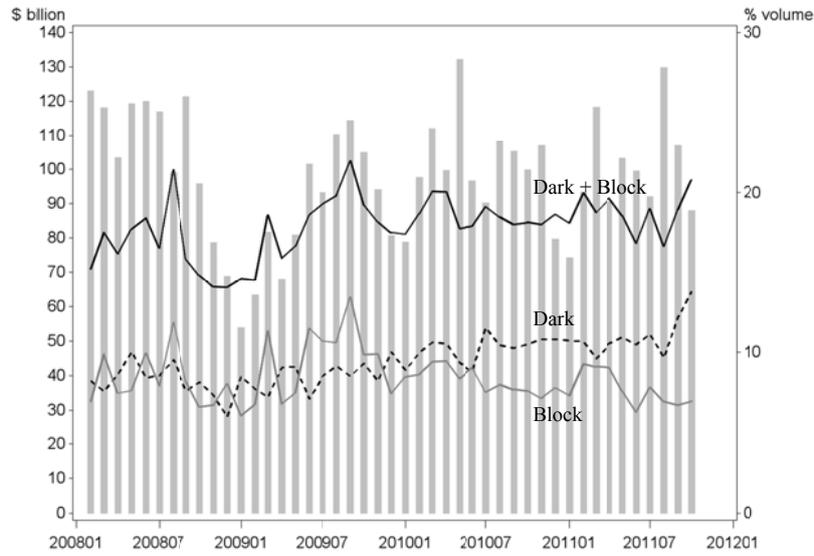
Panel B: Average trade sizes

Figure 1. Dollar volume and average trade sizes

Panel A plots the total dollar volume (in $ billion per month, grey bars) for our sample of stocks (All Ordinaries index constituents) during the sample period. The solid grey and dashed black lines indicate the dollar volume of *Block* and *Dark* trades, respectively, as a percentage of total dollar volume. The solid black line plots the sum of *Block* and *Dark* dollar volume as a percentage of total dollar volume. *Dark* trades are trades executed without pre-trade transparency below block size and *Block* trades are large trades executed without pre-trade transparency. Panel B plots the mean size (in $'000) of *Lit* trades (solid grey line, right hands side scale), *Dark* trades (solid black line, right hands side scale) and *Block* trades (dashed black line, left hands side scale).
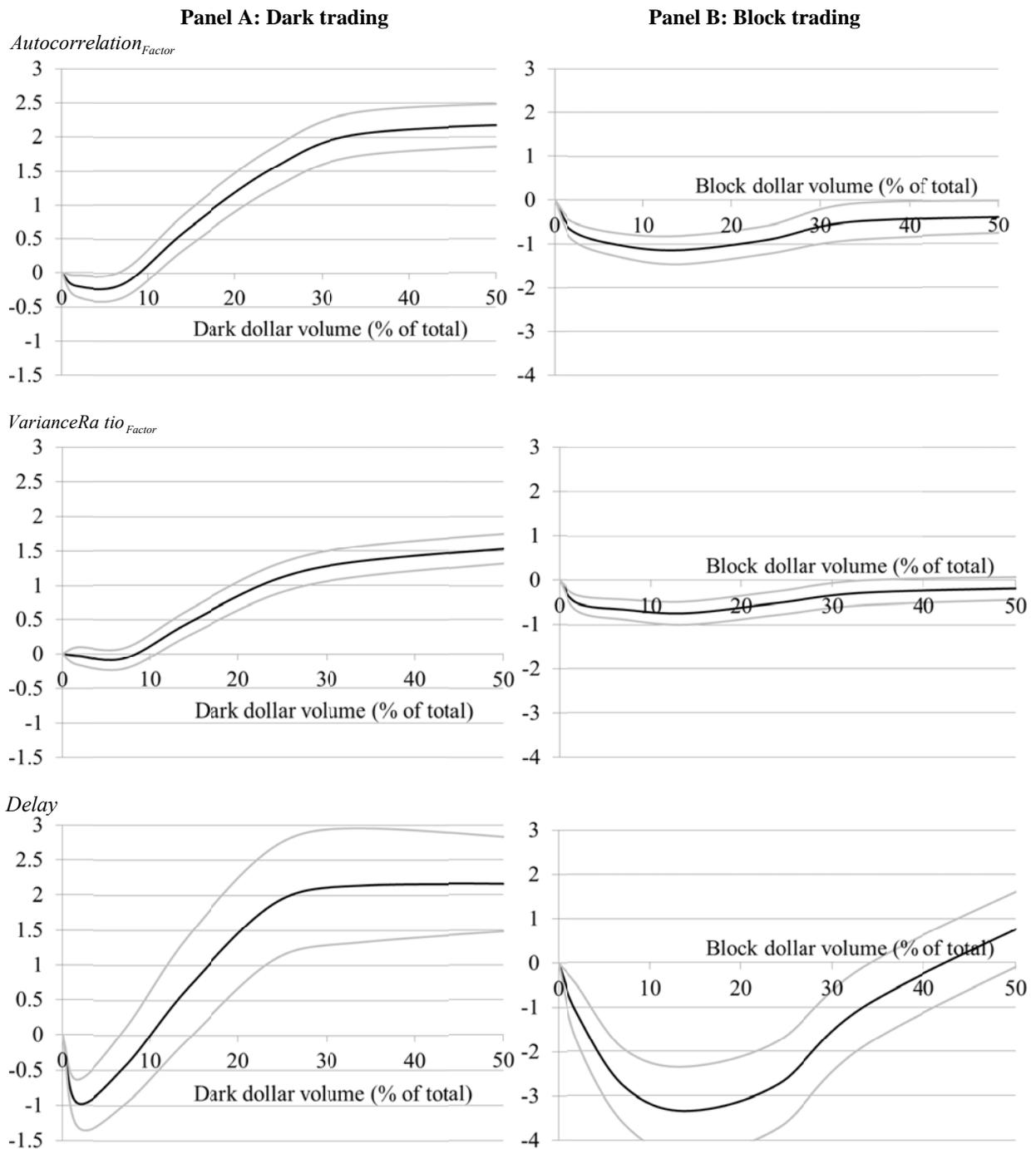
49

Panel A: Dark trading **Panel B: Block trading**

Figure 2. Effects of dark and block trading on informational efficiency

This figure plots the estimated effects of dark trading (Panel A) and block trading (Panel B) (measured as a percentage of total dollar volume, on the horizontal axis) on three informational inefficiency measures (larger values indicate greater informational *inefficiency*, on the vertical axis). The dark lines plot point estimates and the light lines plot error bounds defined by +/- two standard errors. The estimated effects of dark/block trading are obtained from stock-day panel regressions in which the dependent variables are the informational inefficiency measures and the independent variables comprise a set of dummy variables covering various ranges of dark and block trading (0-5%, 5-10%, 10-20%, 20-30%, 30-40%, >40%) and a set of control variables.