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*

19 November 2013

Abstract

Regulators around the world are concerned that growth in dark trading may harm price discovery. We find that low levels of dark trading can be beneficial, but high levels impede price discovery and reduce the informational efficiency of prices. One reason 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. We find no evidence that block trades in the dark impede price discovery.

JEL classification: G14

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

---

∗ Email: carole.comerton-forde@unimelb.edu.au and talis.putnins@uts.edu.au. We thank the Securities Industry Research Centre of Asia-Pacific, the Australian Securities Exchange and Thomson Reuters for providing access to data used in this study. We thank Calissa Aldridge, David Brown, Doug Foster, Joel Hasbrouck, Pamela Moulton, Mark Ready, Thomas Ruf, Avanidhar Subrahmanyam, David Walsh, Geoff Warren, Mao Ye, Haoxiang Zhu, and participants at the NBER Market Microstructure Meeting, the Western Finance Association Annual Meeting, the Financial Research Network Annual Conference, and seminar participants at the Securities and Exchange Commission, Macquarie University, Monash University, University of New South Wales, University of Queensland, University of Sydney, University of Adelaide, University of South Australia, and University of Technology Sydney for helpful comments and suggestions. We also thank a number of brokers and buy-side traders that provided insights into how trading practices and processes have changed during our sample period. Comerton-Forde is an economic consultant on market structure for the Australian Securities and Investments Commission.

1. Introduction

Although 'dark' trading (trading without pre-trade transparency) has always been a feature of equity markets, in recent years, there has been substantial growth in the level of dark trading in markets around the world. There have also been changes in the manner in which dark trading takes place. New trading venues, known as dark pools, have emerged. These venues systematically match orders without providing any pre-trade transparency. Rosenblatt Securities estimate that dark trading in the US has grown from 17% of US consolidated volume in July 2008 to 38% in October 2013.[1] The US provides the most extreme example of the growth in dark trading, but similar trends have also been observed in Europe and Canada.

Dark trading offers a number of benefits to investors, particularly large investors. First, it offers additional liquidity, sometimes in the form of block liquidity. Second, it assists large investors to minimize market impact costs and reduce information leakage. This includes the ability to obtain price improvement by trading within the spread offered by lit exchanges (i.e., exchanges with pre-trade transparency). Third, it can reduce the explicit costs associated with trading as dark trade reporting is typically cheaper than trading on a lit exchange. Traditionally dark trading was managed by sales-traders, but it is now increasingly managed with technology and trading algorithms. The shift away from manual trading to more systematic matching of dark order flow has helped to fuel the growth of dark trading in some markets.

Despite these benefits, many regulators and stock exchanges have expressed concern that the migration of trading volume to venues with little or no pre-trade transparency may harm price discovery and reduce liquidity.[2] Over the last four years many regulators have undertaken public consultations or proposed new regulations on dark trading.[3] However, to date, only the Canadian and Australian regulators have

[1] These numbers comprise 8% (14.3%) in dark pools and 9% (23.4%) in broker-dealer internalization in July 2008 and April 2013, respectively.
[2] See, for example, the SEC Concept Release on Equity Market Structure (2010), the International Organization of Securities Commissions (IOSCO) Technical Committee report on Dark Liquidity Principles (2011), the Joint Canadian Securities Administrators / IIROC Position Paper 23-405 (2010) and the Australian Securities and Investments Commission Consultation Paper 168.
[3] For example, in November 2009 the US SEC proposed rules on the "Regulation of Non-Public Trading Interest." In Europe, the Committee of European Securities Regulators (CESR) undertook a review of pre-trade transparency in 2010 and provided advice to the European Commission on a number of issues

actually implemented new rules.[4] 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.

This paper sheds new light on this debate by analyzing the effect of dark trading on the process of price discovery and the outcome of this process – the informational efficiency of prices. We find that low levels dark trading can be beneficial, but high levels of dark trading are harmful to informational efficiency. The deterioration in informational efficiency begins to occur when below-block-size dark trading in a given stock exceeds approximately 10% of dollar volume, after controlling for other stock characteristics. The change in informational efficiency is economically meaningful in magnitude. 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 early and later parts of our sample period. Our main analysis uses informational efficiency metrics calculated from intra-day observations. In robustness tests we find similar results using lower frequency metrics.

In contrast, we find no evidence that large 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 taping 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

including pre-transparency waivers and limits on the activities of broker crossing systems. To date no new rules have been implemented in either the US or Europe.

[4] The Canadian Securities Administrators and IIROC implemented new rules on 15 October 2012. Under the new rules dark orders must give priority to displayed orders on the same venue, and dark trades for 5,000 shares or less must provide at least one tick price improvement over the Canadian NBBO (one-half tick if the NBBO spread is one tick). The Australian Securities and Investments Commission implemented a similar price improvement rule on 27 May 2013.

of informational efficiency at high levels of dark trading. First, our results suggest 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.

Second, our results support the hypothesis that dark trading leads to partial segmentation of informed and uninformed traders, as predicted by Zhu (2013). We find that orders that execute in the dark tend to be less informed than orders that execute 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 (2013). 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.

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.

This paper is related to a large body of theoretical literature on the effects of transparency, fragmentation, and segmentation. We review this literature in the next section and use it to formulate testable hypotheses. The most relevant theoretical studies are the recent papers by Zhu (2013) and Ye (2012), who model trading in dark pools alongside a lit market. We find support for the mechanisms in both of these studies, suggesting they each describe some of the effects of dark trading. Our characterization of how the nature of price discovery changes in response to an alternative dark trading venue should help guide further development of theory.

This paper is also related to recent empirical studies of dark trading. Degryse et al. (2013) analyze 52 Dutch stocks and conclude that fragmentation of volume across visible order books improves consolidated liquidity, but dark trading has a detrimental effect. Buti et al. (2011) use data from 11 out of 32 US dark pools and conclude that dark pool activity improves 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 (2013) examine data from one of the 32 US dark pools and find that (in relatively less liquid stocks) trading in the dark pool is associated with increased spreads and price impact on the quoting exchanges.

This study complements the existing empirical work by characterizing how dark trading affects the process of price discovery. In addition to the different focus, we are also able to overcome a number of data deficiencies and econometric concerns. To achieve this, we study equities trading on the Australian Securities Exchange (ASX), which allows a cleaner and more detailed analysis compared to what is possible in most other markets. First, during our sample period, there is no fragmentation of displayed liquidity and all dark trades must be reported to the ASX. Therefore, our results are not influenced by fragmentation in displayed liquidity. Second, we are able to precisely measure all dark trading, both block and below-block size over a long time series and broad cross-section of stocks. Therefore, our study does not suffer from biases that can arise from voluntary reporting of data or non-random subsets of dark trading. Third, our data are highly granular. Orders and trades are time-stamped to the millisecond and the time-stamps are consistent across the different trading mechanisms. The highly granular nature of our data allows us to analyze the impact of dark trading on informational efficiency and where price discovery takes place. Finally, we address the potential endogeneity of dark trading by using instrumental variables regressions with two very different classes of instrumental variables.

Dark trading in Australia accounts for an economically significant fraction of aggregate dollar volume traded, averaging around 18% over our sample period (9.4% below-block-size and 8.5% block size). Like in the US, the growth and changing nature of dark trading in recent years has attracted significant regulatory attention due to concern about its impact on market quality. In Australia, these concerns led to new rules aimed at reducing the level of dark trading, which were introduced in May 2013.

2. Theory and hypotheses

The impact of dark trading on price discovery is a complex issue because dark trading simultaneously affects several aspects of a market, including: (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. Recent work has modelled dark trading alongside trading in a lit exchange, thereby accounting for the joint effects of the various changes in market characteristics. We discuss those studies at the end of this section.

2.1 Effects of transparency

Order flow conveys information. In most models of trading and information aggregation, price discovery occurs through market makers or market participants observing incoming orders and updating their beliefs about fundamental values according to the information conveyed by the order flow. For example, in models where investors submit market orders, these orders have price impacts due to the information they convey (e.g., Kyle, 1985; Glosten and Milgrom, 1985); and in models where investors can submit market and limit orders, both order types have price impacts due to the information they convey (e.g., Roşu, 2013; Kaniel and Liu, 2006).

Dark and lit trading differ in their levels of pre-trade transparency, but not post-trade transparency.[5] 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

[5] 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.

market participants revise their beliefs about the fundamental value.[6] 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 generally never know about the order. Even if the order eventually executes and market participants observe the dark trade printed to the tape, the market will still know 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, dark trading may have a negative impact on price discovery.

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 increases the market makers' ability to discern patterns in order flow and infer whether orders are information or liquidity-motivated. This allows market makers to learn information from trades more quickly, set more efficient prices and lowers trading costs for uninformed traders. In Baruch's (2005) model, making the limit order book transparent levels the playing field and increases the ability of market participants to compete with the specialist in liquidity provision. The increased competition for liquidity provision increases the liquidity of the market and also improves price discovery because informed traders choose to trade more aggressively. In contrast, Boulatov and George (2013) model the effects of hiding versus displaying liquidity providing orders and find that

[6] Hautsch and Huang (2012), among others, find empirical evidence that limit orders have price impacts, consistent with theoretical models mentioned earlier.

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.[7]

Eom et al. (2007) argue that market quality is an increasing concave function of pre-trade transparency. Therefore, starting from an opaque market, increases in pre-trade transparency should increase market quality, but at a diminishing rate, and thus, beyond a certain point additional pre-trade transparency will no longer bring benefits and may even be harmful to market quality. An equivalent way of expressing their result is that, market quality is a *decreasing* concave function of pre-trade *opaqueness*. Dark trading reduces the overall level of pre-trade transparency (increases opaqueness) in a market, and therefore, based on the effects of transparency, we would expect high levels of dark trading to 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. This leads to our first hypothesis.

Hypothesis 1*: High levels of dark trading harm price discovery.*

Different segments of the market may benefit from different degrees of transparency. In particular, this is likely to be true for large block trades compared to smaller trades. 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 below-block-size dark trades); however, from the perspective of the upstairs market trade counterparty, the block trade has greater pre-trade transparency than a trade in a lit downstairs market. This is because in negotiating upstairs trades, brokers are able to signal the likely motivation for the trade, and thereby reduce adverse selection risks and

[7] For example, Boehmer et al. (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 et al. (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 electronic communications network suspends the display of the limit order book overall trading costs increase and price discovery deteriorates.

execution costs for large liquidity-motivated trades (Madhavan and Cheng, 1997; Bessembinder and Venkataraman, 2004).

Differences in the nature of upstairs block trading and below-block-size dark trading suggest 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 below-block size 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 to 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). This leads to our second hypothesis.

Hypothesis 2*: Upstairs block trading does not harm price discovery, and may even be beneficial to price discovery.*

In summary, the prediction based on the transparency literature is that high levels of dark trading will harm price discovery, whereas low levels of dark trading as well as moderate levels of block trading will not harm price discovery and may even be beneficial to price discovery. The issue of what constitutes a "high", "low" or "moderate" level is an empirical question. Although we have so far not considered the effects of fragmentation or segmentation of order flow, the following review of those aspects does not contradict our first two hypotheses. The effects of fragmentation and segmentation generate further testable hypotheses and some additional support for our first hypothesis.

2.2 Effects of fragmentation

Dark trading increases the number of trading venues and thus fragments trading activity. For example, with a single lit market as a benchmark, the addition of a dark trading venue results in fragmentation as long as both venues have non-zero shares of trading. More realistically, dark trading often occurs in a large number of separate dark pools and internalization engines, resulting in substantial fragmentation.

The literature identifies several ways in which fragmentation can either benefit or harm market quality, including price discovery. First, network externalities suggest there are benefits to consolidation.[8] When more traders use a particular market, the market's ability to match buyers and sellers increases. Consequently, trading costs decrease, 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 et al., 2008). Fragmentation can also decrease liquidity and harm price discovery by increasing search costs and thus decreasing competition between liquidity providers (e.g., Yin, 2005). There may, however, also be benefits from fragmentation, in the form of increased competition between trading venues, which may result in lower trading costs (e.g., Battalio, 1997; Foucault and Menkveld, 2008; Colliard and Foucault, 2012), and trading platforms that are increasingly tailored to suit the needs of different clienteles (Harris, 1993). The empirical evidence on the effects of fragmentation is mixed, mirroring the various opposing effects predicted by theory.[9]

Fragmentation can also impact markets by altering the way that traders submit orders. Ye (2012) models the change in trading behaviour that occurs when a strategic, monopolistic informed trader submits orders to both a lit market and a dark market that executes orders at the prices set in the lit market (e.g., midquote). The informed trader

---

[8] See, e.g., Mendelson, 1987; Chowdry and Nanda, 1991; Madhavan, 1995; and Hendershott and Mendelson, 2000.

[9] 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 et al. (2013) argue that fragmentation of trading across lit venues can have different impacts than fragmentation stemming from dark venues. Their empirical results for liquidity using a sample of Dutch stocks suggest 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.

knows that his trades in the lit market have price impact and therefore decrease the profits on his dark trades. Consequently, the informed trader reduces the aggressiveness of his trading in the lit market, which impedes price discovery. The incentives for the informed trader to scale back the aggressiveness of his lit trades increase in the amount of volume he executes in the dark. The informed trader is only able to execute large volumes in the dark when a large number of other traders trade in the dark. Therefore, the tendency for the informed trader to scale back the aggressiveness of his lit trades should increase with the share of volume executed in the dark. This leads to our third hypothesis.

Hypothesis 3*: Informed traders reduce the aggressiveness of their trades in the lit market when they trade in the dark, which is more likely when a larger share of volume is executed in the dark.*

Because the decrease in aggressiveness is more likely to occur at higher levels of dark trading, this mechanism provides further support for Hypothesis 1 that high levels of dark trading are detrimental to price discovery.

2.3 Effects of segmentation

Segmentation refers to the tendency for different types of traders to use different markets, either by choice or by institutional setup. In particular, the segmentation of informed and uninformed order flow has important effects on price discovery and other market characteristics.

There are several reasons why dark trading can lead to segmentation of informed and uninformed order flow. First, at any given 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 a dark crossing system than uninformed traders. This mechanism is a central feature of the model of lit and dark trading proposed by Zhu (2013). Zhu predicts that the difference in execution probabilities will lead to a higher proportion of uninformed trading in the dark, and increase adverse selection costs in the lit market. Hendershott and Mendelson (2000) model competition between a dealer network and a crossing network that executes trades at the midquote, similar to some current dark trading venues. Their model also suggests that the crossing network can increase adverse selection on the dealer network for similar reasons as in Zhu's (2013) model. When traders route orders first to the crossing network and then to the dealer network if the order does not execute on the crossing network, the dealers' adverse selection risk increases, causing dealers to set wider spreads.

A second reason why dark trading can cause segmentation of informed and uninformed order flow is that 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 et al., 2012). Third, dark trading can make it easier for brokers to internalize order flow from clients. In fact, a large proportion of dark pools largely facilitate internalization for their owner (Mittal, 2008). Internalization of uninformed order flow is more profitable for a broker than informed order flow due to the differences in adverse selection costs. Internalization of relatively uninformed order flow and the closely related practice of payment for uninformed order flow, both sometimes referred to as "cream skimming", have been documented in various settings (e.g., Chordia and Subrahmanyam, 1995; Easley et al., 1996;). Thus, internalization can also lead to disproportionately large share of uninformed trades taking place in the dark. All three of the reasons outlined above (execution probability, exclusivity, and internalization), lead to our fourth hypothesis.

Hypothesis 4*: A higher proportion of uninformed trades (than informed trades) will execute in the dark.*

If a disproportionately high share of uninformed trades execute in the dark, the increased concentration of informed traders in the lit market is likely to increase adverse selection costs and lead to an increase in spreads in the lit market, as predicted by Zhu (2013) and others. This leads to our fifth hypothesis.

Hypothesis 5*: Dark trading increases spreads in the lit market.*

The effects of segmentation on price discovery, however, are less clear. The theory on how the concentration of informed trading impacts price discovery suggests that the effects can be positive, zero or negative depending on factors such as how strategic informed investors are in the way they trade, the degree of competition among informed traders, the extent to which their information is endogenous and costly to acquire, and the nature of their information (identical signals versus unique signals).[10] Supposing informed traders are less able to trade in the dark than uninformed traders, a substantial decrease in uninformed traders in the lit market could harm price discovery by reducing the profitability of producing unique private information (e.g., Kyle, 1981, 1984, 1989). Alternatively, if all informed traders have the same piece of private information as in Zhu (2013), fewer uninformed traders in the lit market could improve price discovery. Therefore, how the concentration of informed trading in lit venues impacts price discovery is ultimately an empirical question.

2.4 Combined effects of transparency, fragmentation and segmentation

Two recent papers model a setting where dark trading co-exists with lit trading. These two papers, Zhu (2013) and Ye (2012), take different approaches to modelling the impact of dark trading and arrive at conflicting predictions about the effects of dark trading on price discovery. In Zhu's model, the tendency for informed traders to crowd on one side of the market (either buying or selling) causes higher non-execution risk in the dark for informed traders compared to uninformed traders. This leads to relatively more uninformed trading in the dark and a higher concentration of informed traders in the lit market, which Zhu's model predicts will *improve* price discovery. Earlier theoretical literature shows that the impact of a reduction in uninformed traders in a market can have

[10] For example, in sequential trade models such as Glosten and Milgrom (1985) a higher concentration of informed traders increases the rate at which prices converge to the fundamental value and therefore improves price discovery. On the other hand, in strategic trade models such as Kyle (1985) when the volume of uninformed traders decreases (which in our setting could be due to some uninformed traders migrating to the dark), an informed trader finds it optimal to trade less aggressively, and as a result the overall quality of price discovery is unaffected by the number of uninformed traders. Finally, a negative relation between the concentration of informed traders (more precisely, a decrease in uninformed trading) and price discovery could arise if information, in the form of unique noisy signals about the fundamental value, is endogenous and costly (e.g., Kyle, 1981, 1984, 1989; Admati and Pfleiderer, 1988). Intuitively, when the number of uninformed traders decreases, the market becomes thinner, which reduces the amount of profit traders can earn from trading on information. This reduces the amount of information that is acquired, which leads to less informative prices.

a positive or a negative impact on the informativeness of prices, and therefore it is not clear that segmentation will necessarily be detrimental to price discovery. Nevertheless, Zhu's (2013) model provides a useful description of how dark trading leads to segmentation and the prediction of this model are reflected in Hypotheses 4 and 5.

In contrast, Ye (2012) models a monopolistic informed trader that optimally chooses how much to trade in the lit and dark venues. The informed trader scales back his aggressiveness in the lit market so that he can earn greater profits from trading in the dark. Therefore prices are less informative when volume is spread across a lit and a dark venue. This mechanism is reflected in Hypothesis 3, and also provides one of the possible channels by which high levels of dark trading are detrimental to price discovery (Hypothesis 1).

3. Institutional setting

During the period examined in this paper (February 2008 to October 2011) the ASX was the only stock exchange operating in Australia.[11] The ASX is one of the top ten equity markets in the world ranked by market capitalization. There are approximately 2,200 companies listed on the ASX with a market capitalization of around AUD 1.5 trillion. There is substantial competition among the 90 securities brokers participating in the market. The top 12 brokers account for approximately 80% of equity turnover. Most of the top brokers are large global players.

The ASX operates a transparent central limit order book (CLOB) where orders are matched based on price-then-time priority. 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* which 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.

[11] Chi-X Australia was granted a license to offer a competing trading service in ASX-listed stocks, and commenced trading on 31 October 2011.

ii. *Below-block size dark trades* which allow brokers with both sides of a trade to avoid the CLOB time priority rules and match these orders at the prevailing best bid or ask price. There is no minimum size requirement for these trades. For brevity we refer to these trades simply as 'dark' trades.

Further details about these exceptions are provided in the Internet Appendix.[12]

Traditionally dark and block trades were executed by sales-traders; however, during our sample period more automated mechanisms for executing these trades were introduced.[13] In June 2010, the ASX launched a dark pool named *Centre Point.*[14] *Centre Point* is a separate limit order book and orders on *Centre Point* do not interact with orders on the CLOB. *Centre Point* executes orders in price-then-time priority at the midpoint of the bid-ask spread on the CLOB. Over the sample period a number of large brokers launched dark pools to enable them to more systematically execute orders away from the CLOB. The first dark pool in Australia was launched by UBS in August 2005. At the start of our sample period in February 2008 there were four dark pools operating. This grew to 16 broker-operated dark pools by the end of our sample period. Further details about the regulation of these pools and their operators are provided in the Internet Appendix.

For the period of our study there is only limited publically available information about the nature of the order flow in the dark pools in Australia.[15] Most broker-operated dark pools limit the types of traders allowed in the system to institutional investors. Very few operators permit retail investors. Some operators allow users to opt-in or -out of interacting with particular types of order flow. For example, a client may specify that they do not wish to interact with proprietary or high frequency order flow. Users may also limit the order flow that they interact with by specifying minimum execution sizes. Dark pool operators vary in the extent to which their technology allows users to specify

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

[13] This section only outlines changes made in the Australian market during our sample period. There have been subsequent changes, details of which can be found in *ASIC Consultation Papers 168* and *179* and *ASIC Report 331*.

[14] At this time, ASX also introduced a second dark pool, *VolumeMatch,* which was aimed at providing liquidity for block trades. However, only a handful of trades have been executed on *VolumeMatch* so it is now defunct.

[15] In 2012 the Australian Securities and Investments Commission (ASIC) established a dark liquidity taskforce to investigate the operations of dark pools. Details of the findings of the taskforce are available in *ASIC Report 331*.

minimum execution sizes. Some allow users to specify different execution sizes for different stocks or orders, while others allow only one minimum execution size to be specified for all order flow. The ability to opt-in or -out of interacting with particular types of order flow or to specify minimum execution sizes is not available in the CLOB or in *Centre Point.* During our sample period, orders sent to one dark pool were not typically routed to other pools. Orders typically sweep through a single dark pool before being sent to the lit exchange or they rest in the dark pool order book. There were, however, two agency-only brokers operating dark liquidity aggregator businesses.

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. On 1 July 2010, ASX reduced its trading fees for all trade types; however, the fee decrease was larger 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 which 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,[16] in Australia, retail order flow is almost exclusively executed on the ASX.[17] 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.

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

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

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. These stocks account for over 95% of the total market capitalization of all ASX-listed stocks. Our sample period extends from 1 February 2008 to 30 October 2011. The end of the sample period is chosen to avoid confounding 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 executed under the rules of the ASX and 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.[18]

One limitation of our data is that we are not able to identify the mechanism used to execute dark trades. We do not know whether it was executed manually by a sales-trader or using a system or algorithm. Therefore, we are not able to offer any insights into differences between dark trades executed in systems or using algorithms versus those executed using more manual processes.

We restrict our sample to the ASX continuous trading hours of approximately 10:00 to 16:12.[19] Trades that occur in the opening and closing auctions are included in the summations of daily volume used to calculate the proportion of trades or proportion

[18] 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 10-12% of exchange volume, with substantial variation across exchanges.

[19] 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.

of volume that is dark/block, but are not included in our estimation of informational efficiency and price discovery shares.

We are able to precisely classify lit order book 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. Buyer/seller initiated trades are defined by the direction of the trade triggering order. We classify the remaining trade types as buyer (seller) initiated if the trade price is above (below) the prevailing CLOB midquote.

Table 1 reports descriptive statistics on the characteristics of the stocks in our sample. The average stock-day has 1,050 trades per day, with a total value of $9.91 million. The median stock-day has substantially lower levels of trading activity with around 270 trades, and total value $0.7 million. Table 1 also reports that the average (median) company in the sample has a market capitalization of $2.75 billion ($422 million). The average spread of 129 bps is considerably higher than the median spread of 67 bps. On average approximately 60% of the time stocks trade at the minimum possible spread of one tick size ($0.01 for stock prices greater than $2). An average stock-day has around 4.6 quote messages for every trade.

< Insert Table 1 here >

Figure 1 Panel A provides a time-series view of dark, block and total trading activity on the ASX over our sample period, February 2008 to October 2011. The combined proportion of dollar volume executed using dark and block trades has not exhibited any clear trend over the period, accounting for approximately 18% of total dollar volume traded. However, beginning in early 2010 there has been an upward trend in the share of below-block-size dark trades, and a downward trend in the share of dark block trades. This reflects the growth in the number of dark pools, which made it easier to execute below-block-size dark trades. It also suggests that below-block-size dark trades may have been used as a substitute for block trades as brokers increased their use of algorithms.

< Insert Figure 1 here >

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

Table 2 illustrates the variation in dark and block trading activity by year and by size quartile. In Panel A stock-days are weighted by total dollar volume and therefore are comparable to the market aggregates that are often reported by regulators and the media; in Panel B stock-days are equal weighted. The dollar volume weighted averages of dark and block trading are significantly higher than the equal weighted averages, indicating that stock-days with higher trading activity tend to have higher shares of dark and block trades. Consistent with Figure 1, the share of volume executed in the dark increases moderately through time, but there is no clear trend in the level of block trading. The share of block trades is higher in larger stocks. This result is in part driven by the minimum size requirements for block trades which are relatively high for the small stocks in the sample. While equal weighted averages of dark trading are higher for larger stocks, dollar volume weighted averages are not. This suggests that the tendency for higher volume stock-days to have higher dark trading shares is stronger in smaller stocks.

< Insert Table 2 here >

Results that are not tabulated indicate that since the ASX rules were changed to allow below-block-size dark trades to occur at the midquote on 28 June 2010 approximately 11% of dark trades by dollar volume occur within the CLOB's best bid and ask quotes. Almost all remaining dark trades during this period and almost all dark trades prior to 28 June 2010 occur at the prevailing best quotes. Lit trades occur predominantly at the best quotes (96% of lit dollar volume) with a small percentage (4%) 'walking the book' and executing beyond the best quotes. Block trades, however, often occur outside of the prevailing best quotes (77% of block dollar volume).

5. Empirical approach

Our empirical approach involves: (i) estimating a range 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. The price discovery characteristics we examine consist of high-frequency informational efficiency, quoted bid-ask spreads to measure adverse selection risk, and price discovery shares for quotes vs. trades and lit trades vs. dark/block trades to measure where price discovery occurs. We also examine the informativeness of the different trade types by measuring their permanent price impacts.

In the panel regressions that relate dark and block trading to price discovery we take two different approaches: (i) one-stage OLS regression, which does not attempt to deal with the potential endogeneity of dark trading other than through inclusion of various control variables and fixed effects (stock and time); and (ii) two-stage least squares (2SLS) instrumental variables (IV) regressions, which explicitly aim to address potential endogeneity of dark trading. We estimate two forms of the IV models, each using a different set of instruments, discussed below. 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.

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. First, anecdotal accounts of how traders choose to execute trades suggest 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 the panel regressions that we use is as follows:

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

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.

In the one-stage OLS regression models, $DARK_{id}$ and $BLOCK_{id}$ measure the dollar volume of dark and block trades, respectively, as a percentage of the stock-day's total dollar volume.[20] In the 2SLS IV models $DARK_{id}$ and $BLOCK_{id}$ are replaced with fitted values from first stage regressions, following the standard IV approach. In the first stage regressions, $DARK_{id}$ is regressed on the set of instrumental variables and the control variables, for each stock. We do the same for the level of block trading, $BLOCK_{id}$. To assess the strength of our instrumental variables we conduct F-tests of the null hypothesis that the instruments do not enter the first stage regression. The results suggest that our tests do not suffer from the problem of weak instruments.[21]

We use two different sets of instrumental variables. The 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 et al. (2011). The market structure changes include the removal of the 10-second rule on 30 November 2009. Importantly, this change motivated by reasons other than price discovery made it easier to execute dark trades. We construct

[20] 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 'spikes' in dark and block volumes. The results using a log transformation are similar to the baseline specification and our conclusions hold under both measures.

[21] Bound et al. (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.

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, $DarkVenues_t^2$.[22] Together these four variables form our first set of instruments.

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.

In all specifications we estimate standard errors clustered by both stock and by time, as per Petersen (2009) and Thompson (2011). We estimate the panel regression in equation (1) without fixed effects, with stock fixed effects and with time fixed effects to examine how the different sources of variation affect our results, and control for potential

[22] Our results are robust to different combinations of these variables, including omitting the variable $DarkVenues_t^2$. In robustness tests we also include a time trend in both the first and second stages and obtain consistent results.

omitted variables such as a time trend or unobservable time-invariant stock-specific 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 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. Using high-frequency informational efficiency metrics is important to maximize the statistical power of our tests. 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 are likely to have lower statistical power and less precision.

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. Rösch et al. (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.

Both positive and negative midquote autocorrelations suggest quotes deviate from a random walk and have some short-term predictability, which is inconsistent with a highly efficient market. 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):

$$Autocorrelation_k = Corr(r_{k,t}, r_{k,t-1}) \tag{2}$$

where $r_{k,t}$ is the t^{th} 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.[23]

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 $\sigma^2_{1-periodReturn}$. 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:

$$VarianceRatio_{kl} = \left| \frac{\sigma^2_{kl}}{k\sigma^2_l} - 1 \right| \tag{3}$$

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):

[23] 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.

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \sum_{k=1}^{10} \delta_{i,k} r_{m,t-k} + \varepsilon_{it} \qquad (4)$$

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:

$$Delay = 100\left(1 - \frac{R^2_{Constrained}}{R^2_{Unconstrained}}\right) \qquad (5)$$

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.

Table 3 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, suggesting 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 3 suggest that informational efficiency is harmed by dark trading, consistent with Hypothesis 1.

< Insert Table 3 here >

Block trading, however, does not appear to be 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 suggests that trading large blocks off-exchange may even be beneficial to the efficiency of the lit market, consistent with Hypothesis 2. Previous studies suggest two possible reasons why block trades may have a different effect on the market compared to below-block size 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 4), 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 suggest larger magnitude impacts. This suggests 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. We obtain similar results in specifications that omit the control variables.

< Insert Table 4 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 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, and high levels of dark trading tend to exacerbate the inefficient overreactions and reversals.

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:

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

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.[24] 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. These results support Hypothesis 1.

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 >

Panel B of Figure 2 suggests that executing block trades away from the CLOB improves informational efficiency, but only hold 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 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

[24] 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 $DARK_{id} = 1.7\%$.

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 (2013) 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 suggests 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 were 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 had harmful levels of dark trading (>10% of dollar volume) on more than one quarter of the trading days in 2011. No stocks had block trading levels in excess of the 40% 'tipping point' for more than one quarter of the trading days in 2011. All up, these numbers suggest that block trading on a typical day is below harmful levels in all stocks, and below-block-size 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. As illustrated by Table 2, 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) suggests 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 suggests informed traders will execute a considerable share of their trades in the dark and therefore dark trades should be relatively informative. In contrast, Zhu (2013) 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:

$$
\begin{aligned}
x_t^{LIT} &= \mu^{LIT} + \sum_{i=1}^{60} \phi_i^r r_{t-i} + \sum_{i=1}^{60} \phi_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \phi_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \phi_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{LIT} \\
x_t^{DARK} &= \mu^{DARK} + \sum_{i=1}^{60} \theta_i^r r_{t-i} + \sum_{i=1}^{60} \theta_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \theta_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \theta_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{DARK} \\
x_t^{BLOCK} &= \mu^{BLOCK} + \sum_{i=1}^{60} \lambda_i^r r_{t-i} + \sum_{i=1}^{60} \lambda_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60} \lambda_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60} \lambda_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{BLOCK} \qquad (7) \\
r_t &= \mu^r + \sum_{i=1}^{60} \gamma_i^r r_{t-i} + \sum_{i=0}^{60} \gamma_i^{LIT} x_{t-i}^{LIT} + \sum_{i=0}^{60} \gamma_i^{DARK} x_{t-i}^{DARK} + \sum_{i=0}^{60} \gamma_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^r
\end{aligned}
$$

where t indexes 1-second intervals (individual stock and date subscripts are suppressed) and r_t is the log-midquote change in the t^{th} 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 5 reports permanent price impacts of lit, dark and block trades. The average permanent price impact of lit trades is slightly larger than that of dark trades (3.62 bps and 3.31 bps per $10,000) and the difference is statistically distinguishable at the 10% 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) and these differences are also statistically significant.[25] The median permanent price impact for lit 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

[25] 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 et al., 2002).

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 5 here >

These results do not rule out either of the mechanisms modeled in Ye (2012) or Zhu (2013). 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 5 suggest that 'typical' dark and block trades are less informed than 'typical' lit trades, consistent with Hypothesis 4. This finding is consistent with the predictions of Zhu (2013) 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.[26] 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.

Consistent with the predictions of Zhu (2013) and the notion that a larger share of uninformed trades will execute in the dark, Table 6 indicates that quoted spreads become wider in the lit limit order book as dark and block trading increase. These results support Hypothesis 5. The impact of dark trading on spreads in the lit market is highly

[26] An alternative explanation why dark trades tend to be less informed compared to lit trades is that dark trading venues (and even more so broker internalization of order flow) 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.

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.

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 (6)) suggest 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.

< Insert Table 6 here >

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 informative prices (e.g., Kyle, 1981, 1984, 1989; Admati and Pfleiderer, 1988).[27] As

[27] Zhu (2013) 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

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 suggest 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 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

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.

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 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 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:

$$ILS_1 = \frac{\left|\frac{IS_1}{IS_2}\frac{CS_2}{CS_1}\right|}{\left|\frac{IS_1}{IS_2}\frac{CS_2}{CS_1}\right| + \left|\frac{IS_2}{IS_1}\frac{CS_1}{CS_2}\right|}, \quad ILS_2 = \frac{\left|\frac{IS_2}{IS_1}\frac{CS_1}{CS_2}\right|}{\left|\frac{IS_1}{IS_2}\frac{CS_2}{CS_1}\right| + \left|\frac{IS_2}{IS_1}\frac{CS_1}{CS_2}\right|} \tag{8}$$

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:

$$\begin{aligned} \Delta p_{1,t} &= \alpha_1 (p_{1,t-1} - p_{2,t-1}) + \sum_{i=1}^{60} \gamma_i \Delta p_{1,t-i} + \sum_{j=1}^{60} \delta_j \Delta p_{2,t-j} + \varepsilon_{1,t} \\ \Delta p_{2,t} &= \alpha_2 (p_{1,t-1} - p_{2,t-1}) + \sum_{k=1}^{60} \phi_k \Delta p_{1,t-k} + \sum_{m=1}^{60} \varphi_m \Delta p_{2,t-m} + \varepsilon_{2,t} \end{aligned} \tag{9}$$

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 et al. (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 et al. (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 suggest 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 suggests informed traders often use limit orders and provide liquidity, consistent with other studies.[28]

Table 7 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 (Hypothesis 3). 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 7 suggest, consistent with Hypothesis 3. An increasing share of dark trading is 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

[28] For example, Bloomfield et al. (2005), Kaniel and Liu (2006), Goettler et al. (2009), Rosu (2013), and Boulatov and George (2013).

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.

The tendency for midquotes to become relatively more informative (suggesting 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.

< Insert Table 7 here >

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).[29] 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 7 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).

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

[29] 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.

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. Although the results in Table 7 suggest that the *ILS* of dark and block trades (one minus the *ILS* of lit trades) increases when dark and block trades account for a larger share of volume, the increase is considerably less than the increase in the share of volume. In the regressions of ILS_{LIT} the coefficients on *DARK* and *BLOCK* across the five specifications range from -0.41 to 0.03, whereas if the price discovery shares changed proportionally to the volume shares we would expect a coefficient around -1. Put differently, as dark and block trading increase, their contribution to price discovery increases less than proportional to their volume share. This suggests 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 suggests indirectly that the *amount* of dark trading matters for 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.

10. Conclusions

Our results suggest that high levels of below-block-size 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 (2013). 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 (either buying or selling) as in Zhu (2013) 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 support all five of the hypotheses based on theory, and suggest 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 suggests that regulatory action should consider the level of dark trading in specific stocks, rather than the aggregate market level of dark trading. Regulatory proposals such as those being considered by the European Council to cap dark trading at 4% per venue and 8% for the European Union overall may have unintended consequences.

We find no evidence to suggest large block trades negotiated without pre-trade transparency harm informational efficiency. Block trades differ from below-block-size 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, suggesting 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.

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.

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.

Battalio, R.H., 1997, Third market broker-dealers: Cost competitors or cream skimmers?, *Journal of Finance* 52, 341-352.

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.

Committee of European Securities Regulators (CESR), 2010, CESR Technical Advice to the European Commission in the Context of the MiFID Review: Equity Markets.

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

Chordia, T., and A. Subrahmanyam, 1995, Market making, the tick size, and payment-for-order flow: Theory and evidence, *Journal of Business* 68, 543–575.

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.

Easley, D., N.M. Kiefer, and M. O'Hara, 1996, Cream-skimming or profit-sharing? The curious role of purchased order flow, *Journal of Finance* 51, 811-833.

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.

Harris, L., 1993, Consolidation, fragmentation, segmentation, and regulation, *Financial Markets, Institutions and Instruments* 2, 1–28.

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.

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.

International Organization of Securities Commissions (IOSCO) Technical Committee, 2011, Report on Dark Liquidity Principles.

Joint Canadian Securities Administrators/Investment Industry Regulatory Organisation of Canada (CSA/IIROC), 2010, Position Paper 23-405, Dark Liquidity in the Canadian Market.

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.

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.

Mendelson, H., 1987, Consolidation, fragmentation, and market performance, *Journal of Financial and Quantitative Analysis* 22, 189-207.

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

Nimalendran, M., and S. Ray, 2013, Informational linkages between dark and lit trading venues, *Journal of Financial Market* (forthcoming).

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.

Petersen, M.A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435-480.

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.

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.

Thompson, S.B., 2011, Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics* 99, 1-10.

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., 2013, Do dark pools harm price discovery?, *Review of Financial Studies*, (forthcoming).

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

Dark and block trading activity by year and company size

This table reports means (across stock-days) of *Dark* and *Block* trading activity, as a percentage of total dollar volume. In Panel A stock-days are weighted by total dollar volume; in Panel B stock-days are equal weighted. *Dark* trades are trades executed without pre-trade transparency below block size. *Block* trades are large trades executed without pre-trade transparency. Stocks are sorted into quartiles based on market capitalization.

	Size quartile	2008	2009	2010	2011	Pooled
Panel A: Dollar volume weighted						
Dark $ volume	1 = small	11.13	6.83	8.69	9.48	9.11
	2	13.27	9.42	10.70	10.58	10.98
	3	11.92	11.11	12.94	13.26	12.34
	4 = big	7.99	8.33	9.94	10.54	9.20
	Pooled	8.35	8.51	10.15	10.71	9.44
Block $ volume	1 = small	5.66	2.18	4.05	4.31	4.13
	2	5.49	4.52	5.24	4.28	4.88
	3	5.27	5.52	4.67	6.08	5.34
	4 = big	8.50	10.24	8.71	7.88	8.83
	Pooled	8.22	9.82	8.34	7.66	8.51
Panel B: Equal weighted						
Dark $ volume	1 = small	5.69	3.45	4.79	5.16	4.75
	2	6.65	5.88	6.94	6.66	6.53
	3	8.00	7.56	9.01	9.73	8.56
	4 = big	10.15	10.05	11.56	12.88	11.14
	Pooled	7.64	6.82	8.09	8.66	7.79
Block $ volume	1 = small	0.55	0.23	0.32	0.46	0.38
	2	0.95	0.83	0.94	0.98	0.92
	3	2.22	2.84	2.45	2.62	2.54
	4 = big	4.64	5.51	5.08	5.19	5.12
	Pooled	2.10	2.42	2.21	2.34	2.27

Table 3
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 4

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. 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	3.200	3.571	4.632	4.797	97.439	98.956
	(2.47)**	(2.71)***	(5.58)***	(5.70)***	(37.77)***	(37.90)***
DARK	0.147	0.144	0.109	0.109	0.254	0.255
	(16.28)***	(15.85)***	(16.79)***	(16.98)***	(10.89)***	(10.84)***
BLOCK	-0.120	-0.089	-0.060	-0.049	-0.290	-0.197
	(-5.72)***	(-6.07)***	(-3.81)***	(-4.12)***	(-5.26)***	(-4.74)***
Market capitalization	-0.209	-0.215	-0.236	-0.237	-1.672	-1.678
	(-1.82)*	(-1.86)*	(-2.98)***	(-2.98)***	(-6.82)***	(-6.80)***
Quoted spread	-0.480	-0.489	-0.577	-0.583	2.370	2.309
	(-3.13)***	(-3.13)***	(-6.34)***	(-6.35)***	(8.73)***	(8.57)***
Constrained spread	-0.556	-0.510	-0.351	-0.335	7.302	7.460
	(-2.50)**	(-2.29)**	(-2.00)**	(-1.90)*	(13.48)***	(13.75)***
Total $ volume	0.594	0.565	0.459	0.447	-1.093	-1.210
	(7.40)***	(6.97)***	(9.36)***	(9.04)***	(-7.62)***	(-8.57)***
Midquote volatility	0.036	0.037	0.009	0.009	-0.044	-0.041
	(3.88)***	(3.84)***	(3.20)***	(3.20)***	(-4.63)***	(-4.65)***
Message-to-trade	0.004	0.004	0.003	0.003	0.000	0.000
	(1.57)	(1.57)	(1.58)	(1.58)	(-0.36)	(-0.48)
R^2	0.07	0.07	0.11	0.11	0.18	0.18
Estimation method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Instrumental variables	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2

Table 5
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 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	Std. dev.	P25	Median	P75
$PriceImpact_{LIT}$ (bps/$10,000)	3.62	4.94	0.69	1.91	4.74
$PriceImpact_{DARK}$ (bps/$10,000)	3.31	27.94	-0.10	0.03	1.28
$PriceImpact_{BLOCK}$ (bps/$10,000)	0.15	2.53	-0.02	0.01	0.11

Table 6

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. 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.

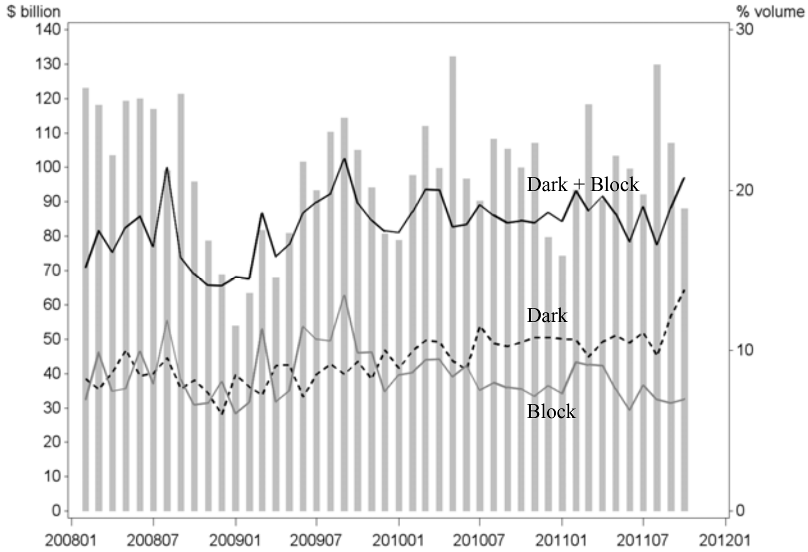
	Log quoted spread				
Intercept	8.492	-4.254	-4.273	8.537	8.557
	(77.37)***	(-102.4)***	(-211.1)***	(80.55)***	(81.71)***
DARK	0.006	0.002	0.006	0.020	0.020
	(11.05)***	(8.77)***	(11.00)***	(10.23)***	(9.75)***
BLOCK	0.004	0.002	0.004	-0.006	-0.004
	(7.52)***	(2.99)***	(7.84)***	(-1.30)	(-1.53)
Market capitalization	-0.216	-0.518	-0.210	-0.189	-0.188
	(-6.56)***	(-16.12)***	(-6.54)***	(-5.57)***	(-5.49)***
Constrained spread	-0.117	-1.022	-0.040	-0.100	-0.097
	(-1.89)*	(-24.08)***	(-0.71)	(-1.59)	(-1.55)
Total $ volume	-0.233	-0.093	-0.241	-0.257	-0.259
	(-12.68)***	(-8.92)***	(-13.53)***	(-11.96)***	(-12.29)***
Midquote volatility	0.018	0.008	0.017	0.019	0.019
	(5.52)***	(4.51)***	(4.89)***	(5.43)***	(5.43)***
Message-to-trade	0.000	0.000	0.000	0.000	0.000
	(-0.09)	(2.23)**	(0.08)	(-1.21)	(-1.23)
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 7
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. 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	70.940	70.993	36.177	-0.803	-7.101	36.100	33.830
	(21.21)***	(-1.31)	(0.57)	(18.41)***	(18.48)***	(11.57)***	(-5.65)***	(-37.52)***	(11.24)***	(10.74)***
DARK	0.085	0.029	0.090	0.293	0.292	-0.004	-0.021	-0.015	-0.016	0.030
	(13.35)***	(5.53)***	(15.11)***	(11.85)***	(11.16)***	(-0.51)	(-3.35)***	(-2.18)**	(-0.68)	(1.24)
BLOCK	-0.013	-0.035	-0.007	-0.290	-0.286	-0.129	-0.126	-0.090	-0.140	-0.406
	(-1.48)	(-4.47)***	(-0.83)	(-5.19)***	(-6.61)***	(-15.39)***	(-14.80)***	(-11.55)***	(-3.15)***	(-10.95)***
Market capitalization	-0.692	-0.533	-0.742	-0.449	-0.451	0.510	0.830	0.675	0.570	0.581
	(-2.60)***	(-1.34)	(-2.84)***	(-1.73)*	(-1.72)*	(1.84)*	(1.51)	(2.33)**	(1.99)**	(2.09)**
Quoted spread	0.444	0.458	0.217	-0.217	-0.220	-3.810	-3.020	-3.537	-3.804	-3.908
	(1.04)	(0.91)	(0.58)	(-0.44)	(-0.44)	(-10.11)***	(-4.95)***	(-9.42)***	(-10.12)***	(-10.48)***
Constrained spread	-3.689	-4.849	-2.155	-3.841	-3.840	9.017	5.513	8.203	9.200	8.731
	(-6.20)***	(-9.12)***	(-3.55)***	(-6.62)***	(-6.59)***	(12.31)***	(5.16)***	(12.50)***	(13.12)***	(12.72)***
Total $ volume	-0.643	0.603	-0.826	-0.988	-0.991	2.767	2.057	2.624	2.741	2.972
	(-3.25)***	(3.03)***	(-4.67)***	(-3.69)***	(-3.72)***	(14.25)***	(11.41)***	(13.86)***	(12.47)***	(14.26)***
Midquote volatility	0.137	0.140	0.119	0.160	0.160	0.059	0.044	0.057	0.058	0.057
	(4.53)***	(4.11)***	(4.16)***	(4.34)***	(4.33)***	(2.27)**	(2.11)**	(2.11)**	(2.23)**	(2.3)**
Message-to-trade	0.001	0.001	0.002	0.000	0.000	0.001	0.022	-0.043	0.001	0.005
	(0.65)	(0.66)	(0.99)	(0.27)	(0.27)	(0.04)	(1.19)	(-2.28)**	(0.07)	(0.25)
R^2	0.04	0.02	0.03	0.05	0.05	0.11	0.02	0.11	0.11	0.12
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

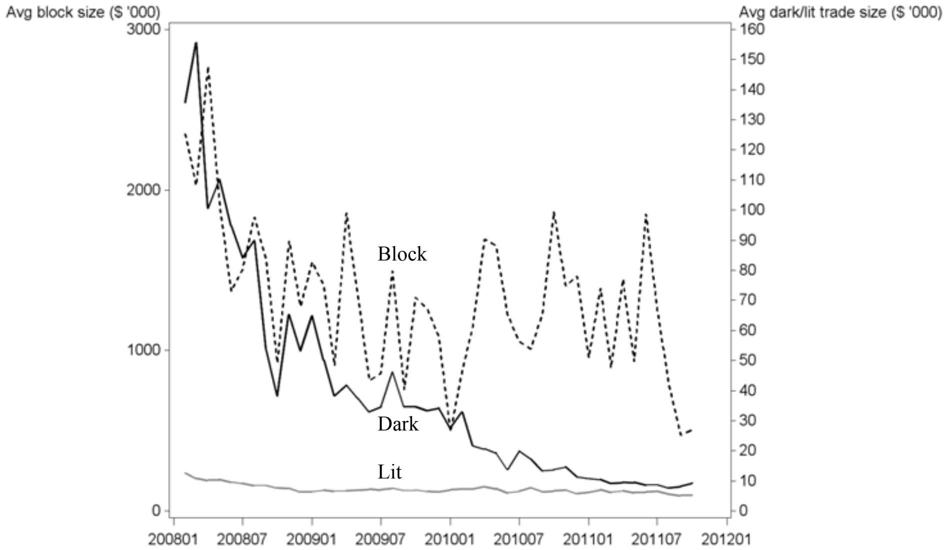


Figure 1. Dollar volume and average trade sizes

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).

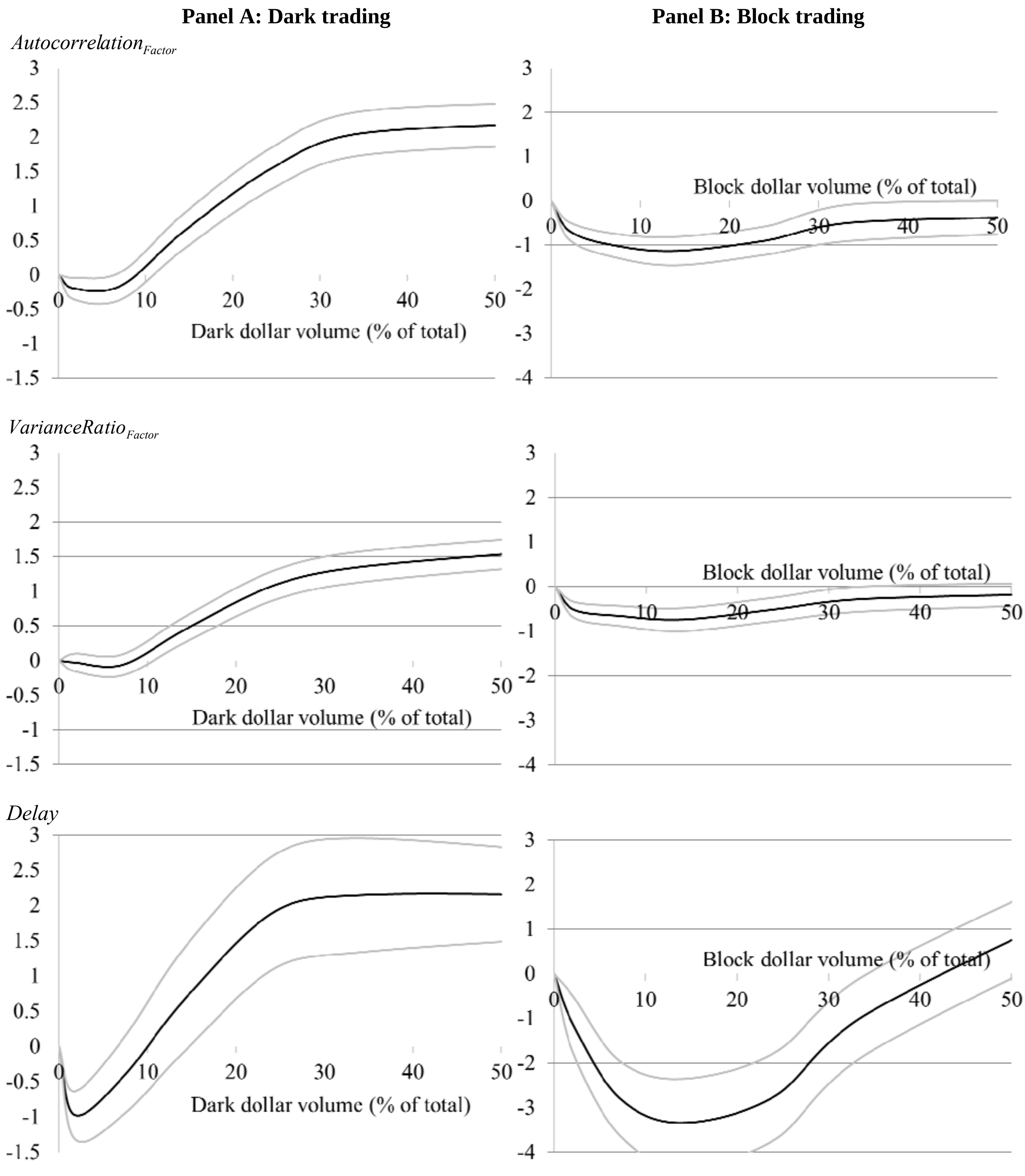


Figure 2. Effects of dark and block trading on informational efficiency

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.