# Dark trading and price discovery

Carole Comerton-Forde [a] and Tālis J. Putniņš [b] *
[a] *College of Business and Economics, Australian National University, Australia*
[b] *University of Technology Sydney* and *Stockholm School of Economics in Riga*

2 November 2012

**Abstract**

Regulators and stock exchanges around the world are concerned that growth in the share of equities volume executed without pre-trade transparency, so called 'dark trading', may harm price discovery. We empirically analyze the impact of dark trading on price discovery. We find that high levels of dark trading impede price discovery and cause prices to become less informationally efficient. Order flow that migrates to the dark is less informed than that which is left behind, but it is not entirely uninformed. Therefore, the loss of pre-trade information on migrating order flow harms price discovery. It also increases adverse selection risk, bid-ask spreads and price impact in the transparent exchange. This decreases the incentives to engage in costly information acquisition, thereby further reducing the informational efficiency of prices. We find no evidence that large block trades that occur in the dark impede price discovery.

*JEL classification:* G14

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

* Email: carole.comerton-forde@anu.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 Doug Foster, David Walsh and seminar participants at Macquarie University, Monash University, University of New South Wales, University of Queensland, University of Sydney, 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.

## 1. Introduction

Dark trading is trading that occurs without pre-trade transparency. Although the term "dark trading" is relatively new, dark trading has always been a feature of equity markets. In recent years, markets around the world have exhibited substantial growth in the level of dark trading and there have 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 31% in July 2012.[1] The US provides the most extreme example of the shift to dark trading, but similar trends have also been observed in Europe and Canada.

Dark trading offers a number of potential 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. 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

---

[1] These numbers comprise 8% (14%) in dark pools and 9% (17%) in broker-dealer internalization in July 2008 and July 2012, respectively.

[2] For example, the International Organization of Securities Commissions (IIROC) in a 2011 report by the Technical Committee states that "the development of dark pools and use of dark orders could inhibit price discovery if orders that otherwise might have been publicly displayed become dark". A 2010 joint Canadian Securities Administrators / IIROC position paper on dark pools notes that "if orders that would traditionally be sent to visible marketplaces are increasingly diverted to Dark Pools or entered as Dark Orders there could be a negative impact on the price discovery process". The European Commission in a 2010 consultation paper raises the concern that "increased use of dark pools … may ultimately affect the quality of the price discovery process on the 'lit' markets". The Australian Securities and Investment Commission in a 2011 consultation paper state that "pre-trade transparency is fundamental to price

many regulators have undertaken public consultations or proposed new regulations on dark trading.[3] However, to date, only the Canadian regulators have actually implemented new rules.[4] The extensive consultation and subsequent lack of action 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 aims to shed light on this debate by analyzing the effect of dark trading on the informational efficiency of prices. We also investigate how the process of price discovery changes in response to a greater share of volume being executed in the dark. We find that dark trading impacts the aggregate quality of price discovery and the nature of how price discovery occurs. Informational efficiency is negatively related to the share of volume executed in the dark, suggesting dark trading harms aggregate price discovery. The deterioration in informational efficiency begins to occur once dark trading exceeds approximately 10% of dollar volume, after controlling for other stock characteristics. This 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 price discovery. 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.

In contrast, we find no evidence that large block trades negotiated away from the exchange with no pre-trade transparency harm aggregate price discovery. In fact, having

---

formation". The president of TSX Markets (an operator of stock exchanges in Canada) in a 2011 letter to IIROC claims that "price discovery is of the highest quality when it is built on maximum visibility and maximum participation. Every order placed and share traded away from maximum visibility necessarily represents a deterioration from optimized price discovery". Nasdaq OMX's chief economist Frank Hatheway in a congressional testimony stated that dark pools "undermine public price discovery by shifting liquidity away from the lit markets".

[3] For example, in November 2009 the US SEC proposed rules on the "Regulation of Non-Public Trading Interest." These rules have not been implemented. 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 including pre-transparency waivers and limits on the activities of broker crossing systems. To date no new rules have been implemented.

[4] The Canadian Securities Administrator and IIROC approved new rules on 13 April 2012. These rules require that (i) priority be given to displayed orders over non-displayed orders on the same venue; (ii) dark trades for 5,000 shares or less must give the active participant at least one tick price improvement over the Canadian NBBO (one-half tick if the NBBO spread is one tick); and (iii) give IIROC the power to impose a minimum trade size on passive dark orders if they deem it to be necessary (although at present the minimum trade size has been set to zero). These rules take effect on 15 October 2012.

some block trades execute away from the lit market (up to approximately 15% of dollar volume) can be beneficial to aggregate price discovery.

As order flow migrates from lit to dark trading mechanisms the contribution of midquotes to price discovery increases compared to that of trade prices, suggesting liquidity providers become increasingly informed. Also, the speed at which dark trades reflect changes in the fundamental value increases relative to that of lit trades. This suggests there is some information contained in dark trades, and the loss of pre-trade information on this segment of volume is one of the possible reasons why dark trading harms price discovery.

In addition to the loss of pre-trade information, our results suggest a second mechanism by which dark trading harms aggregate price discovery. The order flow that migrates to the dark tends to be relatively less informed than that which is left behind in the lit exchange. This increases adverse selection risk, bid-ask spreads and price impact in the lit exchange and consequently decreases incentives for costly acquisition and analysis of information. Constraints such as restrictions on the parties allowed to trade in dark pools prevent informed traders from following the migration of relatively less informed orders. Thus, less information gathering and analysis contributes to making prices less informationally efficient, which ultimately undermines resource allocation and economic efficiency. When forced to interact with informed traders in a single market, uninformed traders implicitly contribute to price discovery by providing compensation to informed traders for the costs of becoming informed. However, when uninformed traders can to some extent avoid interacting with informed traders by using dark trading mechanisms their implicit contribution to price discovery is reduced.

Two recent theory papers, Zhu (2012) and Ye (2012), analyze the question of whether dark pools harm price discovery and arrive at opposite predictions. Zhu (2012) predicts that adding a dark pool alongside an exchange concentrates price-relevant information into the exchange and *improves* price discovery. The mechanism driving this result is that uninformed traders make relatively greater use of the dark because their execution probability in the dark is higher than that of informed traders who are more likely to cluster on either the buy or sell side of the market. Ye (2012) on the other hand predicts that a dark crossing network harms price discovery. In Ye's model informed traders submit orders in both the lit and dark markets, but reduce their aggressiveness in the lit market due to the negative externality that price impact in the lit market has on the informed trader's dark trading profits. 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.

Empirical studies of dark pools are often limited by a lack of good data and tend to focus more broadly on market quality including liquidity. Degryse et al. (2011) analyze a sample of 52 Dutch stocks between 2006-2009 and conclude that fragmentation of volume across visible order books improves global liquidity (measured by consolidating the limit order books of all lit trading venues), but dark trading has a detrimental effect. Buti et al. (2011) use data from 11 out of 32 US dark pools during 2009 and conclude that dark pool activity improves spreads, depth and short-term volatility. Our study complements Degryse et al. (2011) and Buti et al. (2011) by focusing on price discovery, another important aspect of market quality. Nimalendran and Ray (2012) examine transaction-level data from one of the 32 US dark pools and consistent with our paper find that (in relatively less liquid stocks) trading in the dark pool is associated with increased spreads and price impact on the quoting exchanges. They do not, however, examine the effects of dark trading on the aggregate quality of price discovery.

We study trading in equities on the Australian Securities Exchange (ASX). The Australian setting and the related data offer a number of advantages relative to the markets examined and data used in the existing literature. First, during our sample period, the ASX is a monopoly exchange. As a result, there is no fragmentation of displayed liquidity and all dark trades must be reported to the ASX. Therefore, our results are not influenced by competition between exchanges or the associated fragmentation in displayed liquidity. Second, we are able to measure all dark trading, both block and non-block 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. Third, our data are highly granular. Orders and trades are time-stamped to the millisecond and the ASX trading rules mean that 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 the nature of where price discovery takes place using information shares and the aggregate quality of price discovery using measures of informational efficiency. One limitation of our data is that we are not able to identify the mechanism used to execute dark trades. 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.

## 2. Institutional setting

The Australian equity market is dominated by the ASX. The ASX provides services in listing, trading, clearing and settlement. ASX has historically had a monopoly in the provision of these services for ASX-listed equities. Its monopoly in the provision of trading services ended on 31 October, 2011 when Chi-X Australia was granted a market license by the Australian government to trade ASX-listed equities. The ASX is one of the top ten equity markets in the world ranked by market capitalization. Like other major stock exchanges it is a publically listed for-profit company. There are approximately 2,200 entities listed on the ASX with a market capitalization of around AUD 1.5 trillion. There is substantial competition within the securities broking industry in Australia with around 90 brokers. However, the top 12 brokers account for approximately 80 percent of equity turnover. Most of the top brokers are large global players in the securities industry.

The ASX operates a transparent central limit order book (CLOB) in which orders are matched based on price then time priority. The *ASX Operating Rules* provide a number of exceptions that allow trades to be executed away from the CLOB with reduced pre-trade transparency. These include:

i. *Block special crossings* which must have a minimum value of $1 million and *portfolio special crossings* which must comprise a portfolio of at least ten stocks with a minimum value for each stock of at least $200,000 and a combined portfolio value of at least $5 million.[5] These trades may be negotiated away from the CLOB at any price and immediately reported to the ASX.[6]

[5] ASX Operating Rule 4810, 4811 and Procedure 4810(2) and 4810(3).
[6] ASX Operating Rule 3500.

ii. *Priority crossings* allow brokers with both sides of a trade to avoid the CLOB time priority rules. Prior to November 2009, these trades had to be executed according to the "ACE" rule, which required the broker to "appear" in the CLOB at the crossing price for 10 seconds (i.e., place an order), "create" a crossing market (i.e., ensure the spread is at the minimum tick), and then "execute" the trade at the best bid or ask price. The convention followed by brokers in most instances of priority crossings was for the broker to appear in the CLOB with an order for only one share. If their order was not executed within 10 seconds they could then execute the priority crossing. Therefore, although these orders provided some pre-trade transparency, they were not practically accessible to other participants in the market as only one share was offered. In November 2009, the requirement to wait 10 seconds before executing the priority crossing was removed, essentially allowing priority crossings to occur without any pre-trade transparency. There is no minimum size requirement for priority crossings.

In June 2010, ASX introduced a number of additional trading facilities and order types which allow orders to be executed away from the CLOB without any pre-trade transparency.[7,8] These include:

iii. *Centre Point priority crossings* which provide additional flexibility for priority crossings to be executed at the midpoint of the best bid and ask price on the CLOB. There is no requirement for the broker to appear in the CLOB before executing a *Centre Point* priority crossing. Like *priority crossings*, *Centre Point priority crossings* have no minimum size requirement.

iv. *Centre Point* which is an ASX operated dark pool that executes orders in time priority at the midpoint of the bid-ask spread on the CLOB. However, *Centre Point* orders do not interact with orders on the CLOB. There is no minimum size requirement for *Centre Point* orders.

---

[7] See *Australian Securities and Investments Commission (ASIC) Report 215* for details.

[8] At this time, ASX also introduced a second dark pool, *VolumeMatch,* which was aimed at providing liquidity for large trades over $1 million. However, only a handful of trades have been executed on *VolumeMatch* so we do not consider these trades in our analysis.

Throughout the paper we refer to *Block Special Crossings* and *Portfolio Special Crossings* as block trades. We use the term dark trades to refer to *Priority Crossings*, *Centre Point priority crossings* and *Centre Point trades*. We use the term lit trades to refer to trades executed on the CLOB.

The nature of dark and block trading on the ASX has changed in recent years.[9] Institutional investors have increased their use of algorithms to manage their trading work flow. The removal of the 10 second rule in November 2009 made it easier for brokers to systematically execute *priority crossings*. A number of large brokers developed dark pools (or adopted the dark pools operated by their overseas parents) to enable them to more systematically internalize orders or match client order flow. The first dark pool in Australia was launched by UBS in August 2005. Four additional dark pools were launched before 2010. The number of dark pools grew rapidly during 2010 and 2011, with a total of 16 dark pools operating by October 2011.[10] The development of these pools has allowed for increased use of algorithms to manage order executions. This has also meant that some orders which may previously have been executed as block trades are now broken up using algorithms and executed as dark trades. Some algorithms attempt to execute orders across both lit and dark trading mechanisms. Some algorithms dynamically adjust the volume of the order being directed to lit or dark mechanisms depending on where the order has been successfully filled earlier in the day.

There is only limited information available publicly about the nature of the order flow in the dark pools in Australia. Some dark pool operators limit the types of traders allowed in the system to institutional investors. Other 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 Trading (HFT) order

[9] This section only outlines changes made in the Australian market during our sample period of February 2008 to October 2011. There have been a number of subsequent and proposed changes after this time. Details of these changes may be found in *ASIC Consultation Papers 168* and *179*.

[10] Prior to the introduction of competition and the ASIC Market Integrity Rules (Competition in Exchange Markets) the concept of a dark pool was not formalized in the market rules. These rules use the term crossing system rather than dark pool to define the use of technology to systematically match orders away from an exchange order book. A crossing system is defined in these rules to be "any automated service provided by a Participant to its clients which matches or executes Client Orders with Orders of: (a) the Participant; or (b) other clients of the Participant, otherwise than on an Order Book." This includes both dark pools and internal matching engines that allow orders to be posted on the lit market and later withdrawn by the broker when an internal crossing becomes possible. *ASIC Consultation Paper 168* provides the names of the operators and launch dates for each of these systems.

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 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.

There are a number of important differences between the dark pools operated in Australia and those operated in the US. First, unlike the US markets where these types of venues need to be registered as an Alternative Trading System (ATS), dark pools in Australia operate under the rules of the ASX and are exempt from the requirement to obtain a market license. Second, the level of interconnectedness between dark pools is much lower. With the exception of two dark liquidity aggregators operated by two agency only brokers, orders sent to one dark pool are not 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. It is also common for parts of an order to be simultaneously posted to the lit exchange and a dark pool. Third, with the exception of block trades, dark pools may only match orders at prices corresponding to the minimum tick size on the exchange or at the midpoint of the tick size. Finally, the *ASX Operating Rules* require that all trades executed on a dark pool must be immediately reported to the exchange and disseminated to the market. During our sample period, trades executed by dark pools are reported to the exchange as priority crossings, block and portfolio crossings. Since June 2010, they have also been reported as *Centre Point priority crossings*. The trade reporting does not differentiate between trades executed in dark pools and trades executed away from the CLOB in a less automated or systematic way.

Exchange trading fees are charged as a proportion of dollar trading value. ASX reduced its trading fees on 1 July 2010. CLOB trades were charged at 0.28 bps prior to 1 July 2010 and 0.15 bps after this time. Fees for crossings were also reduced. *Block* and *Portfolio crossing* fees fell from 0.15 bps to 0.1 bps and *priority crossing* fees fell from 0.075 bps to 0.05 bps. *Centre Point priority crossings* are charged at 0.15 bps while

*Centre Point* trades are charged at 0.5 bps. With the exception of *Centre Point* trades, these fees are capped at $75 per trade.[11]

## 3. 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 approximately 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 (including lit, dark and block) 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. Therefore, we avoid issues which arise in other markets due to inconsistencies in time-stamps across different trading venues.

We restrict our sample to the ASX continuous trading hours of approximately 10:00 to 16:12.[12] 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 of volume that is dark/block, but are not included in our estimation of price discovery measures.

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. *Centre Point*

[11] Details available at: http://www.asxgroup.com.au/media/PDFs/20100603_asx_fees_rebates.pdf

[12] 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.

*priority crossings* (approximately 0.3% of trades and 0.2% of dollar volume) are unable to be classified as buyer or seller initiated and therefore do not contribute to signed volume aggregates used in estimating the informativeness of different trade types.

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 provides a time-series view of the aggregate dollar trading volume and trade frequency on ASX over the period February 2008 to October 2011. Panel A of Figure 1 shows that the combined proportion of dollar volume executed using dark and block trades has not exhibited any clear trend over the period. However, beginning in early 2010 there has been an upward trend in the proportion of dollar volume executed as dark trades, and a downward trend in the proportion of trades executed as blocks. This suggests that dark trades may have been used as a substitute for block trades as brokers increased their use of algorithms. Panel B of Figure 1 shows that there has been an extremely rapid increase in the proportion of trades executed as dark trades, rising from approximately 2% to over 10% of trades. Over the same period, the proportion of block trades remained relatively constant.

< Insert Figure 1 here >

Table 2 reports descriptive statistics on dark and block trading activity at a stock-day level. In the full sample the mean percentage of dollar volume of dark trades is 7.8%

and the proportion of dark trades is 3.2%. For block trades the corresponding mean percentages are 2.2% of total dollar volume and 0.1% of all trades. The median values for both dark and block trading are considerably lower indicating that a relatively small proportion of stock-days have very high dark and block activity. Over three-quarters of the stocks-days have no block trades and over one-quarter of the sample has no dark trades.

< Insert Table 2 here >

Table 3 illustrates the variation in dark and block trading activity by year and by size quartile. The share of volume executed in the dark increases moderately through time and is greater in larger stocks. In the largest quartile, during the first ten months of 2011 dark trades account for approximately 13% of total dollar volume (12% of total trades) compared to 10% of total dollar volume (2% of total trades) in 2008. The share of block trades is also higher in larger stocks – in the largest quartile block trades account for approximately 5% of dollar volume (0.1% of total trades). 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. There is little growth in the level of block trading over the sample period.

< Insert Table 3 here >

Figure 2 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 HFT.

< Insert Figure 2 here >

Results that are not tabulated indicate that since trading at the midquote was allowed with the introduction of *Centre Point* and *Centre Point priority crossings* 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 before the introduction of *Centre Point* 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).

## 4. 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 metrics to measure the quality of aggregate price discovery, price discovery shares for quotes vs. trades and lit trades vs. dark/block trades to measure where price discovery occurs, permanent price impacts of different types of trades to measure trade informativeness, and quoted bid-ask spreads to measure adverse selection risk.

To analyze the effects of dark trading on the price discovery characteristics we estimate the following panel regression:

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

where $y_{id}$ is one of the informational efficiency, price discovery share, trade informativeness or adverse selection risk measures for stock *i* on day *d*, $DARK_{id}$ and $BLOCK_{id}$ measure the dollar volume of dark and block trades, respectively, as a percentage of the stock-day's total dollar volume. 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. In the regression model above, $C_{jid}$ are 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 size, 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.

We estimate the regression above 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 omitted variables such as a time trend or unobservable time-invariant stock-specific characteristics. We estimate standard errors clustered by both stock and by time, as per Petersen (2009) and Thompson (2011).

We are interested in the causal effects of dark trading on price discovery. An obstacle in making such inference is the potential endogeneity of dark trading – it is plausible that a stock's informational efficiency and other price discovery characteristics influence where trades are executed. Anecdotal accounts of how traders choose where to execute trades suggest that the choice is influenced by considerations such as the amount of liquidity in the CLOB, perceived liquidity and the probability of execution the dark, whether the CLOB spread is constrained to the minimum of one tick, whether price improvement can be achieved in the dark and the degree of automation that is available for dark execution (e.g., presence of crossing systems and algorithms to manage orders across both dark and lit venues). Therefore, we would expect endogeneity concerns to be more severe in causally relating dark trading and liquidity, than relating dark trading and price discovery. Furthermore, the panel regression above controls for many variables that might influence both dark trading and price discovery, e.g., spread, whether the CLOB spread is constrained, volume, volatility, stock size and algorithmic trading.

To further examine potential endogeneity we use instrumental variables. Using a similar approach to Buti et al. (2011) and Hasbrouck and Saar (2011) we 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. Results from the two-stage instrumental variables regressions (described in more detail in Section 8) are very similar to those from the simple panel regression specified in equation (1). This suggests that after controlling for the variables mentioned above, any remaining endogeneity has little if any effect on our results. We therefore use the simpler regression specified in equation (1) for our primary specification and report instrumental variables results in robustness tests.

## 5. Informational efficiency and aggregate price discovery

In order to analyze how dark and block trading affects 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.

Positive or negative midquote autocorrelations suggest quotes deviate from a random walk process 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 $t^{\text{th}}$ 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 also 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).

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 $\sigma^2_l$ and $\sigma^2_{kl}$ 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 also 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 measure of short-term return predictability is an intraday adaptation of the Hou and Moskowitz (2005) *Delay*, i.e., the extent to which lagged market returns predict a stock's midquote returns. For each stock-day we estimate a regression of 1-minute midquote returns for stock $i$, $r_{i,t}$, on the All Ordinaries market index return, $r_{m,t}$, and ten lags (suppressing day subscripts):

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

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:

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

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.

The three intraday informational inefficiency metrics - $Autocorrelation_{Factor}$, $VarianceRatio_{Factor}$, and *Delay* - range from 0 (highly efficient) to 100 (highly inefficient). Table 4 reports panel regression estimates of the relation between the share of dark and block trading and informational efficiency. 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 efficiency, and statistically significant for almost all specifications with and without fixed effects. Therefore, the results suggest that aggregate price discovery is harmed by migration of order flow away from the lit order book to dark trading venues.

< Insert Table 4 here >

Block trading, however, does not appear to be detrimental to aggregate price discovery. 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.

The effects of dark and block trading on price discovery may be nonlinear in their share of volume. 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:

$$y_{id} = \alpha + \delta^{0-5\%} D_{id}^{0-5\%} + \delta^{5-10\%} D_{id}^{5-10\%} + \delta^{10-20\%} D_{id}^{10-20\%} + \delta^{20-30\%} D_{id}^{20-30\%} + \delta^{30-40\%} D_{id}^{30-40\%} + \delta^{>40\%} D_{id}^{>40\%}$$
$$+ \beta^{0-5\%} B_{id}^{0-5\%} + \beta^{5-10\%} B_{id}^{5-10\%} + \beta^{10-20\%} B_{id}^{10-20\%} + \beta^{20-30\%} B_{id}^{20-30\%} + \beta^{30-40\%} B_{id}^{30-40\%} + \beta^{>40\%} B_{id}^{>40\%} + \sum_{j=1}^{6} \delta_j C_{jid} + \varepsilon_{id} \quad (6)$$

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 greater than zero but less than or equal to 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 greater than 10% but less than or equal to 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.

Figure 3 plots the coefficients of the dummy variables for each of the three informational efficiency metrics.[13] Panel A suggests that low levels of dark trading are not harmful to price discovery, but as dark trading increases it eventually reaches a 'tipping point' after which it begins to have negative impact. Specifically, after

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

controlling for other stock characteristics, when dark trading accounts for 10% of total dollar volume its impact on price discovery is very close to zero, i.e., it neither harms nor benefits informational efficiency. However as dark trading increases above 10% of dollar volume informational efficiency deteriorates. For example, a large 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. 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. Although the magnitudes of these effects are not extremely large, they are nevertheless economically meaningful.

Panel B of Figure 3 suggests that the improvement in informational efficiency when large block trades are executed away from the market only hold up to a certain point. When block trading reaches approximately 15% of total dollar volume additional block trades tend to have a negative impact on informational efficiency. Block trading at 15% of dollar volume is associated with improvements in the informational efficiency measures of approximately 15% to 20% of a standard deviation. In general, small amounts of block trading away from the lit market is often good for price discovery, but as with dark trading: too much can be harmful.

< Insert Figure 3 here >

For how many stocks are the current levels of dark trading harming 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 in our sample had block trading levels in excess of the 15% ‘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, whereas approximately 12% of stocks have dark trading levels that on a typical trading day are detrimental to price discovery.

## 6. Price discovery shares

The results in the previous section suggest that after a certain point, more dark trading is associated with less informationally efficient prices. To understand why this is the case we next examine how the nature of price discovery changes in response to dark trading. We start by examining where price discovery occurs.

There are two main approaches to measuring the contributions of different markets or different sources of quotes or volumes to price discovery: Hasbrouck's (1995) information share (*IS*) and Gonzalo and Granger's (1995) common factor share (*CS*). The models are directly related (see Baillie et al., 2002) and, as pointed out by Yan and Zivot (2010), both are required together to disentangle two dimensions of market efficiency: (i) impounding of new information; and (ii) avoidance of transitory shocks. Fundamentally, both methods decompose the impact of a price innovation into permanent and temporary components. $CS_i$ measures the relative avoidance of noise trading and liquidity shocks (avoidance of "bad" volatility) in price series $i$, whereas $IS_i$ measures the relative extent to which price series $i$ incorporates more new information (more "good" volatility) and/or contains less noise trading and liquidity shocks (avoidance of "bad" volatility).

Typically, the contribution of a market or source of quotes/trades to price discovery is considered to be the extent to which it is the first to impound new information about the 'true' underlying asset value. Yan and Zivot (2010) show that this aspect of price discovery, which we refer to as "information leadership" (*IL*), can be isolated using a combination of *IS* and *CS* as follows:

$$IL_1 = \left| \frac{IS_1}{IS_2} \frac{CS_2}{CS_1} \right|, \qquad IL_2 = \left| \frac{IS_2}{IS_1} \frac{CS_1}{CS_2} \right| \tag{7}$$

Unlike *CS* and *IS*, the *IL* proposed by Yan and Zivot is not a "share" in that $IL_1$ and $IL_2$ do not sum to 1 (or even approximately 1). Rather, $IL_1$ has the range $[0,\infty)$. Values of

$IL_1$ above (below) 1 suggest that $p_1$ leads (does not lead) the process of incorporating new information about the fundamental value. In order to make the information leadership metric easier to interpret and more readily comparable to *IS* and *CS* we standardize *IL* giving rise to "information leadership shares" (*ILS*):

$$ILS_1 = \frac{IL_1}{IL_1 + IL_2}, \quad ILS_2 = \frac{IL_2}{IL_1 + IL_2} \tag{8}$$

Now, $ILS_1$ and $ILS_2$ have the range [0,1], similar to *IS* and *CS* (which we scale up), 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. The transformation from *IL* to *ILS* also helps limit the influence of outliers. In addition to *ILS* we also report results for *IS* given its widespread use in existing studies of price discovery.

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 quotes as well as prices from different types of volume (lit and dark/block) for each stock within one market (similar to Fong and Zurbruegg (2003) and Anand and Subrahmanyam (2008)). The prices of quotes, lit trades and dark/block trades for a stock are all, in the long run, 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). The estimates of $CS$ are not affected by the ordering of the price series in the VECM, however, $IS$ estimates are affected. We therefore use the approach advocated by Baillie et al. (2002) (also used by Booth et al. (2002) and Cao et al. (2009) among others) and calculate $IS$ under each of the two possible orderings and then take the simple average.

The information leadership shares reported in Table 5 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). The information share of the midquote is even larger (median $IS_{MIDQUOTE}$ of 0.85), suggesting that midquotes also contain less noise compared to trade prices, one reason being that they do not suffer from bid-ask bounce as do trade prices. Lit limit order book trades contribute more than dark and block trades to impounding new information (median $ILS_{LIT}$ of 0.85) but are also somewhat noisier than dark and block trades (because median $IS_{LIT}$ of 0.70 is below median $ILS_{LIT}$), consistent with the fact that dark trades that occur within the spread contain less noise from bid-ask bounce.

< Insert Table 5 here >

Table 6 reports regression estimates using Yan-Zivot information leadership shares ($ILS$) and Hasbrouck information shares ($IS$) of midquotes (compared to trade prices) as the dependent variables. The coefficient estimates indicate that an increasing proportion of dark dollar volume is associated with an increase in the $ILS$ and a decrease in the $IS$ of the midquote, holding other variables fixed. This suggests that as dark trading increases, the midquote becomes faster in reflecting innovations in the fundamental value compared to trade prices, but trade prices become less noisy compared to midquotes. Most coefficients for the share of block trading are not statistically significant, suggesting no clear relation between block trading and price discovery shares of the midquote relative to trade prices.

It is difficult to predict ex-ante how dark and block trading would impact *ILS* and *IS* of midquotes relative to trade prices. Goettler et al. (2009) and Menkhoff et al. (2010) point out that, among other things, 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 at the best quotes 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 can be expected to increase relative to trade prices. Therefore, the positive relation between dark trading and *ILS* of the midquote is consistent with increasingly informed liquidity providers in the lit market. This could be a direct result of order flow migration to the dark if, for example, the migrating order flow contains a high proportion of uninformed traders that had a tendency to be net liquidity suppliers. However, a change in informativeness of liquidity suppliers/consumers in the lit market could also be the result of informed or uninformed traders in the lit market changing their order submission strategies in response to volume migrating to the dark. The fact that trade prices tend to become less noisy compared to midquotes as the share of volume executed in the dark increases could also be caused by dark trades being restricted to occur at either the best quotes or at the midquote.

< Insert Table 6 here >

The coefficients of control variables suggest that the share of new information impounded by the midquote compared to trade prices decreases when the spread is constrained by the minimum tick size. The midquote's contribution to impounding new information increases relative to that of trade prices during times of high midquote volatility and wide spreads, all else equal.

Table 7 examines the contribution of lit trades to price discovery, compared to the contribution of dark and block trades. The results indicate that as the share of dark/block trading increases lit trades contribute relatively less to impounding new information, compared to dark and block trades, and as dark trading increases lit trades become relatively less noisy. These results are consistent with the notion that order flow migrating away from the limit order book contains some information about the fundamental value and therefore causes an increase in the role of dark/block trades in impounding new information relative to the role of lit trades. The loss of pre-trade information on the migrating order flow may result in less than the full information content of the trades being impounded into prices. This is one of the possible reasons why increasing levels of dark trading are associated with deteriorating price discovery.

< Insert Table 7 here >

The coefficients of control variables suggest that the share of new information impounded by lit trades compared to dark and block trades increases when the spread is narrow, in particular when it is constrained by the minimum tick size, when total dollar volume is high and during times of high midquote volatility, all else equal.

**7. Informativeness of different types of trades**

There are 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, while for informed traders a lack of transparency can help prevent information leakage. Therefore, whether relatively more or less informed orders migrate to dark venues is an empirical question, and one that has implications for how price discovery occurs and how adverse selection risk changes with the introduction of dark trading venues. The previous subsection's results on information shares suggest that the order flow that migrates away from the lit order book is relatively less informed. We now test this hypothesis in more detail.

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:

$$\begin{aligned}
x_t^{LIT} &= \mu^{LIT} + \sum_{i=1}^{60}\phi_i^r r_{t-i} + \sum_{i=1}^{60}\phi_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60}\phi_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60}\phi_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{LIT} \\
x_t^{DARK} &= \mu^{DARK} + \sum_{i=1}^{60}\theta_i^r r_{t-i} + \sum_{i=1}^{60}\theta_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60}\theta_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60}\theta_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{DARK} \\
x_t^{BLOCK} &= \mu^{BLOCK} + \sum_{i=1}^{60}\lambda_i^r r_{t-i} + \sum_{i=1}^{60}\lambda_i^{LIT} x_{t-i}^{LIT} + \sum_{i=1}^{60}\lambda_i^{DARK} x_{t-i}^{DARK} + \sum_{i=1}^{60}\lambda_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^{BLOCK} \\
r_t &= \mu^r + \sum_{i=1}^{60}\gamma_i^r r_{t-i} + \sum_{i=0}^{60}\gamma_i^{LIT} x_{t-i}^{LIT} + \sum_{i=0}^{60}\gamma_i^{DARK} x_{t-i}^{DARK} + \sum_{i=0}^{60}\gamma_i^{BLOCK} x_{t-i}^{BLOCK} + \varepsilon_t^r
\end{aligned} \tag{10}$$

where $t$ indexes 1-second intervals (individual stock and date subscripts are suppressed) and $r_t$ is the log-midquote change in the $t^{\text{th}}$ interval.

For each equation, we estimate coefficients on 60 lags of each variable. In addition to the 60 lags, midquote returns are also determined by contemporaneous order flow. 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 types of volume 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 in our permanent price impact measures we winsorise them by setting extreme positive and negative values to the 1st and 99th percentile values, for each stock and each date.

Table 8 reports permanent price impacts of lit, dark and block trades in the full sample. Consistent with previous result that lit trades impound more new information than dark and block trades, Table 8 indicates that lit trades also have higher permanent price impact (median of 2.09 bps per \$10,000, compared to medians of 0.03 and 0.01 for dark and block trades, respectively). The mean price impacts (4.46, 4.00 and 0.16 for lit, dark and block trades, respectively) suggest that on average lit trades are more informed than dark trades, which are more informed than block trades. Note that 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.

< Insert Table 8 here >

< Insert Table 9 here >

The regression results in Table 9 confirm that the order flow that migrates away from the lit order book is relatively less informed than the remaining lit trades. As dark and block trading increases, the permanent price impact of lit trades increases. The migration of relatively less informed order flow to dark venues and off-market block executions leaves behind relatively more informed order flow in the lit venue. The change in the informativeness of lit order flow increases adverse selection risks for uninformed liquidity providers in the lit limit order book. Increases in the levels of dark (block) trading are associated with decreases in the price impacts of dark (block) trades.

Consistent with the increase in adverse selection risk when dark and block trades account for a greater share of volume, Table 10 indicates that quoted spreads become wider in the lit limit order book as relatively less informed order flow migrates away. This increases the costs of trading in the lit market, which is likely to encourage further migration of order flow away from the lit market in a self-reinforcing spiral.

The finding that adverse selection risk increases as order flow migrates to the dark is consistent with the results on price discovery shares, in particular the increase in the speed at which the midquote reflects innovations in the fundamental value compared to trade prices. 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 compared to trade prices could result from increased adverse selection risk causing a higher proportion of informed traders to act as liquidity suppliers.

The magnitude of the increase in quoted spreads and price impact of lit trades as dark and block trading increases is not extremely large, but is nevertheless economically meaningful. For example, estimates using the 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 12% after controlling for other factors. This means that for the average stock spreads will increase from 129 bps to 144 bps. A more modest increase in dark trading from 10% to 12.5% is expected to increase spreads by 2.4% (an increase of 3.1 bps for the average stock). Similarly, an increase in dark trading from zero to 10% of dollar volume will increase the price impact of lit trades by 0.09 bps per $10,000 (an increase of 5% relative to the median price impact), after controlling for a range of other factors. Similarly, an increase in block trading from zero to 10% of dollar volume is expected to increase spreads by a smaller 1%, but have a larger effect on the price impact of lit trades, increasing it by around 5 bps per $10,000.

< Insert Table 10 here >

The increase in quoted spreads and in the price impact of lit order flow suggests a second mechanism, in addition to the loss of pre-trade information that might explain why dark trading harms informational efficiency. The higher trading costs and sensitivity of prices to lit order flow decrease the incentives to gather information and become informed, because the profits for doing so, at least if trading on the lit market, decrease as dark trading becomes more prevalent. Due to the selective membership in many dark pools whereby only relatively uninformed traders are allowed access (see Boni et al., 2012), informed traders that would like to follow the liquidity-motivated traders to dark venues are unable to do so. Therefore, as relatively less informed order flow migrates away from the lit limit order book traders that have the choice of gathering costly information and performing analysis will do so less intensely, leading to less information aggregation.

This explanation is supported by market microstructure theory. In the seminal models of Kyle (1985) and Glosten and Milgrom (1985) uninformed market makers break even on average, informed traders profit from trading on their information about the fundamental value, which means that on average liquidity-motivated traders in aggregate lose on their trades an amount exactly equal to the informed traders' aggregate profits. The wealth transfer from uninformed traders to informed traders occurs implicitly 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 information gathering and thus for providing price discovery. In fact, when information acquisition is costly the absence of liquidity-motivated traders (also referred to as 'noise' traders) can cause a complete breakdown of price discovery resulting in an informationally inefficient market (Grossman and Stiglitz, 1980; Black, 1986). Therefore, part of the trading costs paid by less informed traders implicitly plays an important role in facilitating price discovery by providing compensation for others to gather and impound information. Our results are consistent with the notion that the savings that less informed traders are able to make by trading with each other in restricted membership dark pools may provide them with a private gain, but this comes at the cost of less compensation being available for information gathering and therefore less aggregate price discovery. Viewing accurate stock prices as a public good due to their role in facilitating efficient allocation of resources across alternative uses, trading in dark venues is associated with a negative externality to the extent that it impedes price discovery.

The result that lit trades have higher information content as order flow migrates to dark venues suggests that if order flow migration continues the migrating order flow will become increasingly informed. If pre-trade information contributes to price discovery this implies that if the trend of order flow migration continues, it will harm aggregate price discovery at an accelerating rate. Every additional unit of order flow that migrates to the dark carries with it more valuable pre-trade information and therefore has an increasingly detrimental effect on aggregate price discovery and informational efficiency.

Is there reason to believe that order flow migration will increase in the future? Our results suggest one reason why the answer might be yes. The migration away from the lit order book has left a higher proportion of informed traders in the lit market. Many studies suggest that in equilibrium informed traders try to hide among the uninformed liquidity traders to maximize the amount they are able to trade before their information is impounded in the price. Therefore, there is reason to believe the informed traders in the lit market will attempt to follow the relatively less informed traders to the dark venues.

The extent to which they will be able to do so depends, among other things, on how well dark venues are be able to screen out informed from uninformed clientele when increasing membership.

## 8. Robustness tests

### *8.1 Potential endogeneity of dark trading*

It is possible that dark trading and characteristics of price discovery are jointly determined – dark trading may affect price discovery, *and* the nature and quality of price discovery may influence the decision of where to execute trades. This type of endogeneity could bias the results towards or against finding a negative relation between dark trading and aggregate price discovery. We examine the influence of potential endogeneity using a standard two-stage instrumental variables approach. We follow Buti et al. (2011) and Hasbrouck and Saar (2011) in selecting instrumental variables. Specifically, we 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 likely to correlate with the level of dark trading in a particular stock, 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.

We use the instrumental variables in a standard two-stage procedure. In the first stage we regress the level of dark trading, $DARK_{id}$, on its instrument and the control variables for each stock. We do the same for the level of block trading, $BLOCK_{id}$. In the second stage we estimate the main panel regression specified in equation (1) replacing $DARK_{id}$ and $BLOCK_{id}$ by their fitted values from the first stage.

< Insert Table 11 here >

Table 11 reports the results of the second-stage regressions for the informational efficiency (aggregate price discovery) dependent variables. The results are similar to those from the simple regressions, in several cases with greater magnitude and a higher level of statistical significance. This suggests that our results relating dark trading to deterioration of price discovery are not driven by traders choosing to execute in the dark when price discovery is poor. If anything, the magnitudes of the estimates in Table 11 suggest that endogeneity may work against finding significant results in our simple regressions. These results provide stronger evidence of a causal link from dark trading to a deterioration of price discovery. We also apply the instrumental variables approach to the other dependent variables (price discovery shares, price impacts of trades and quoted spreads) and find similar results as in the simple regressions.

*8.2 Other robustness tests*

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 aggregate price discovery deteriorates, the information content of lit trades increases (higher permanent price impact of lit order low), and spreads on the lit market increase.

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 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 internalization of order flow and matching of client orders. However, if dark pools make it easier to trade in the dark they may encourage growth of dark trading to levels that become harmful to price discovery. In addition to splitting our sample by calendar year, we also split it into periods before and after the November 2009 when the requirement for brokers to appear in the CLOB before executing a priority crossing was removed. The results in these two subsamples are similar, suggesting that the lack of accessibility to the order flow in addition to the lack of transparency contributes to the result.

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, 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 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 (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.

## 9. Conclusions

Our results suggest that as order flow migrates away from a lit CLOB to the dark the aggregate amount of price discovery and the nature of how price discovery occurs are impacted. Our results suggest that higher levels of dark trading are associated with decreased price discovery and less informationally efficient prices. As order flow migrates, trades take on a more important role in impounding new information, compared to order book quotes. The order flow that is first to migrate away from the lit limit order book tends to be less informed than the order flow that is left behind. Therefore, the price impact of lit trades increases, as does adverse selection risk and bid-ask spreads in the lit exchange. The increased adverse selection risk and trading costs on the lit exchange increases the incentives for more order flow to migrate away from the lit exchange, creating a self-reinforcing spiral. Informed traders left behind in the more transparent venue naturally would like to follow the less informed order flow to the dark, but their ability to do so is limited by exclusivity of dark pools whereby some dark pool operators try to limit participation to relatively less informed clientele. Instrumental variables methods suggest that our results are robust to the potential endogeneity of dark trading.

Together these results provide support for the concerns of regulators that high levels of dark trading harm price discovery. However, given that there has not been substantial growth in the level of dark trading in the Australian market this is not a new phenomenon in the market. High levels of dark trading are found to be harmful throughout the sample period. In contrast we find no evidence to suggest large block trades negotiated without pre-trade transparency harm price discovery.

There are two channels, not mutually exclusive, that might explain why dark trading harms price discovery. First, although the order flow that migrates to the dark tends to be less informed than the order flow that is left behind, it is not completely uninformative; it has higher permanent price impact per unit volume than block trades. The loss of valuable pre-trade information on this moderately informed order flow that leaves the lit market may be one of the factors that impedes price discovery. Under this mechanism a continuation of order flow migration is expected to harm price discovery at an accelerating rate because the marginal informativeness of migrating order flow is increasing.

The second channel by which dark trading may harm price discovery stems from the fact that as relatively less informed order flow migrates away from the lit market, adverse selection risk, spreads and price impact all increase in the lit market. This acts as a disincentive to engage in costly information acquisition for traders that do not have access to dark trading mechanisms, and therefore can decrease the aggregate amount of information and analysis about fundamental values. In a market where informed and uninformed traders interact, part of the trading costs paid by uninformed traders, such as the bid-ask spread; implicitly compensate informed traders for the costs of becoming informed. When relatively less informed traders are able to trade amongst themselves in dark pools they no longer make a contribution towards paying for the public good of price discovery. As prices become less informationally efficient, resource allocation and economic efficiency suffer.

From a policy perspective, informationally efficient markets are clearly desirable due to their positive effects on resource allocation and economic efficiency. For most stocks in the All Ordinaries Index, the level of dark trading has not yet reached the level where it is harmful to price discovery. However, growth in the share of trading in the dark is expected to cause an increase in the number of stocks for which the deterioration in price discovery is significant. The growth in the number of crossing systems and the increased use of algorithms to manage execution may contribute to growth in dark trading in the future. Excessive growth in the level of dark trading can be prevented with informed, carefully considered regulation that constrains the migration of order flow to the dark.

Our results show that not all non-transparent trading has the same effects on price discovery. For example, block trading does not harm price discovery. If anything, it is beneficial to the market to have large block trades negotiated away from the lit central limit order book. Therefore, regulation to constrain dark trading needs to be carefully designed to limit the migration of order flow that is beneficial to the lit market, while allowing order flow that does not positively contribute to price discovery (or may even detract from price discovery) to occur in the dark.

**References**


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 Investment Commission Report 215, Australian Equity Market Structure, November 2010.

Australian Securities and Investment Commission Consultation Paper 168, Australian Equity Market Structure: Further Proposals, October 2011.

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.

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

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.

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

Cao, C., O., Hansch, and X. Wang, 2009, The information content of an open limit-order book, *Journal of Futures Markets* 29, 16-41.

Degryse, H., F. de Jong and V. van Kervel, 2011, The impact of dark trading and visible fragmentation on market quality, Working paper.

Fong, K., and R. Zurbruegg, 2003, How much do locals contribute to the price discovery process? *Journal of Empirical Finance* 10, 305-320.

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

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

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

Hasbrouck, J., and G. Saar, 2011, Low-latency trading, Working paper.

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

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.

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

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.

Menkhoff, L., C.L. Osler, and M. Schmeling, 2010, Limit-order submission strategies under asymmetric information, *Journal of Banking and Finance* 34, 2665-2677.

Nimalendran, M., and S. Ray, 2012, Informational linkages between dark and lit trading venues, Working paper.

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

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

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.

Zhu, H., 2012, Do dark pools harm price discovery?, Working paper.

**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 (all trades that are executed with little or no pre-trade transparency, including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades, not including *Block* trades), and *Block* trades, which are large trades executed with no 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 trading descriptive statistics**

This table reports means, standard deviations and quartile points (*P25*, *Median*, *P75*) of dark and block trading variables calculated at the stock-day level. Total volume consists of *Lit* trades (trades executed in the transparent central limit order book), *Dark* trades (all trades that are executed with little or no pre-trade transparency, including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades, not including *Block* trades), and *Block* trades which are large trades executed with no pre-trade transparency.

| | Mean | Std. dev. | P25 | Median | P75 |
|---|---|---|---|---|---|
| Dark $ volume (% of total) | 7.77 | 15.86 | 0.00 | 0.43 | 7.66 |
| Dark trades (% of total) | 3.24 | 7.02 | 0.00 | 0.36 | 3.37 |
| Block $ volume (% of total) | 2.26 | 8.21 | 0.00 | 0.00 | 0.00 |
| Block trades (% of total) | 0.07 | 0.92 | 0.00 | 0.00 | 0.00 |

**Table 3**
**Dark and block trading activity by year and company size**

This table reports means (across stock-days) of *Dark* and *Block* trading activity (dollar volume and number of trades), as a percentage of total trading activity. *Dark* trades are all trades that are executed with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades), not including *Block* trades, which are large executed with no pre-trade transparency. Size is measured by market capitalization.

| | Size quartile | 2008 | 2009 | 2010 | 2011 | Pooled |
|---|---|---|---|---|---|---|
| Dark $ volume | 1 = small | 5.69 | 3.44 | 4.78 | 5.15 | 4.74 |
| | 2 | 6.63 | 5.87 | 6.94 | 6.63 | 6.52 |
| | 3 | 7.99 | 7.55 | 9.00 | 9.73 | 8.55 |
| | 4 = big | 10.15 | 10.03 | 11.57 | 12.90 | 11.13 |
| | Pooled | 7.63 | 6.81 | 8.09 | 8.63 | |
| 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 | 2.21 | 2.84 | 2.45 | 2.61 | 2.54 |
| | 4 = big | 4.64 | 5.51 | 5.07 | 5.19 | 5.11 |
| | Pooled | 2.10 | 2.42 | 2.21 | 2.32 | |
| Dark trades | 1 = small | 1.66 | 1.40 | 2.34 | 2.40 | 1.95 |
| | 2 | 0.96 | 1.44 | 2.27 | 3.35 | 2.00 |
| | 3 | 1.09 | 1.53 | 4.03 | 6.91 | 3.35 |
| | 4 = big | 1.58 | 2.57 | 6.64 | 11.97 | 5.58 |
| | Pooled | 1.32 | 1.75 | 3.83 | 6.18 | |
| Block trades | 1 = small | 0.05 | 0.02 | 0.03 | 0.05 | 0.04 |
| | 2 | 0.04 | 0.07 | 0.06 | 0.07 | 0.06 |
| | 3 | 0.07 | 0.15 | 0.11 | 0.10 | 0.11 |
| | 4 = big | 0.06 | 0.09 | 0.07 | 0.08 | 0.08 |
| | Pooled | 0.06 | 0.08 | 0.07 | 0.08 | |

**Table 4**

**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 trades that are executed in venues with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | $Autocorrelation_{Factor}$ | | | $VarianceRatio_{Factor}$ | | |
|---|---|---|---|---|---|---|
| Intercept | 2.476 | -0.073 | -0.120 | 3.255 | -0.028 | -0.070 |
| | (1.90)* | (-3.13)*** | (-1.28) | (3.30)*** | (-1.63) | (-0.91) |
| DARK | 0.032 | 0.007 | 0.032 | 0.029 | 0.010 | 0.029 |
| | (11.34)*** | (3.58)*** | (11.02)*** | (14.2)*** | (5.90)*** | (13.90)*** |
| BLOCK | -0.012 | -0.017 | -0.014 | -0.010 | -0.014 | -0.011 |
| | (-3.91)*** | (-6.89)*** | (-4.47)*** | (-4.39)*** | (-8.57)*** | (-4.96)*** |
| Market capitalization | -0.303 | 0.024 | -0.316 | -0.401 | -0.039 | -0.412 |
| | (-1.91)* | (0.11) | (-1.97)** | (-3.22)*** | (-0.24) | (-3.28)*** |
| Quoted spread | -0.367 | -0.746 | -0.359 | -0.296 | -0.124 | -0.302 |
| | (-2.45)** | (-5.23)*** | (-2.39)** | (-2.57)** | (-1.18) | (-2.62)*** |
| Constrained spread | -1.246 | -3.081 | -1.268 | -1.019 | -2.476 | -0.971 |
| | (-4.34)*** | (-14.80)*** | (-4.39)*** | (-4.44)*** | (-16.31)*** | (-4.21)*** |
| Total $ volume | 0.718 | 1.022 | 0.735 | 0.622 | 0.893 | 0.625 |
| | (9.74)*** | (25.34)*** | (9.79)*** | (10.52)*** | (32.18)*** | (10.33)*** |
| Midquote volatility | 0.049 | 0.065 | 0.050 | -0.006 | -0.005 | -0.008 |
| | (7.54)*** | (13.28)*** | (7.89)*** | (-1.32) | (-1.73)* | (-1.72)* |
| Message-to-trade | 0.010 | 0.008 | 0.010 | 0.008 | 0.006 | 0.008 |
| | (1.86)* | (1.72)* | (1.83)* | (1.86)* | (1.73)* | (1.83)* |
| $R^2$ | 0.05 | 0.03 | 0.05 | 0.07 | 0.04 | 0.07 |
| Fixed effects | None | Stock | Date | None | Stock | Date |

| | *Delay* | | |
|---|---|---|---|
| Intercept | 87.896 | 0.023 | 0.269 |
| | (30.57)*** | (0.28) | (1.50) |
| DARK | 0.044 | 0.000 | 0.047 |
| | (7.20)*** | (0.11) | (7.46)*** |
| BLOCK | -0.002 | -0.003 | -0.001 |
| | (-0.23) | (-0.61) | (-0.20) |
| Market capitalization | -1.867 | -0.535 | -1.796 |
| | (-5.76)*** | (-1.59) | (-5.57)*** |
| Quoted spread | 3.678 | 0.958 | 3.708 |
| | (10.92)*** | (3.78)*** | (11.07)*** |
| Constrained spread | 6.050 | 1.921 | 5.679 |
| | (8.86)*** | (4.80)*** | (8.54)*** |
| Total $ volume | -0.447 | 0.070 | -0.443 |
| | (-3.38)*** | (1.11) | (-3.39)*** |
| Midquote volatility | -0.123 | -0.036 | -0.101 |
| | (-8.67)*** | (-5.49)*** | (-7.28)*** |
| Message-to-trade | -0.002 | -0.004 | 0.000 |
| | (-1.11) | (-1.37) | (0.27) |
| $R^2$ | 0.16 | 0.00 | 0.16 |
| Fixed effects | None | Stock | Date |

**Table 5**

**Price discovery shares**

This table reports means, standard deviations and quartile points (*P25*, *Median*, *P75*) of price discovery shares calculated at the stock-day level. $ILS_{MIDQUOTE}$ and $IS_{MIDQUOTE}$ are the Yan-Zivot information leadership share and the Hasbrouck information share of midquotes relative to trade prices (pooling all trade types). $ILS_{LIT}$ and $IS_{LIT}$ are the Yan-Zivot information leadership share and the Hasbrouck information share of lit trade prices relative to dark and block trade prices.

| | Mean | Std. dev. | P25 | Median | P75 |
|---|---|---|---|---|---|
| $ILS_{MIDQUOTE}$ | 0.55 | 0.23 | 0.42 | 0.56 | 0.70 |
| $IS_{MIDQUOTE}$ | 0.79 | 0.21 | 0.73 | 0.85 | 0.93 |
| $ILS_{LIT}$ | 0.75 | 0.26 | 0.66 | 0.85 | 0.94 |
| $IS_{LIT}$ | 0.62 | 0.33 | 0.36 | 0.70 | 0.92 |

**Table 6**
**Effects of dark and block trading on price discovery shares of quotes vs trade prices**

This table reports regression estimates using a stock-day panel, in which the dependent variables are the Yan-Zivot information leadership share ($ILS_{MIDQUOTE}$) and the Hasbrouck information share ($IS_{MIDQUOTE}$) of midquotes relative to trade prices (scaled up by a factor of 100). *DARK* and *BLOCK* are the percentage of the stock-day's total trades that are executed in venues with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | $ILS_{MIDQUOTE}$ | | | $IS_{MIDQUOTE}$ | | |
|---|---|---|---|---|---|---|
| Intercept | 67.168 | -0.142 | 0.164 | 38.737 | 0.164 | -0.371 |
| | (21.74)*** | (-1.69)* | (0.85) | (15.53)*** | (2.14)** | (-2.14)** |
| DARK | 0.083 | 0.030 | 0.086 | -0.024 | -0.018 | -0.026 |
| | (12.26)*** | (6.28)*** | (12.88)*** | (-3.64)*** | (-4.15)*** | (-3.93)*** |
| BLOCK | 0.002 | -0.023 | 0.006 | 0.005 | 0.022 | -0.007 |
| | (0.18) | (-2.81)*** | (0.63) | (0.59) | (3.17)*** | (-0.85) |
| Market capitalization | -0.590 | -1.233 | -0.568 | -0.734 | 0.563 | -0.775 |
| | (-1.78)* | (-2.04)** | (-1.72)* | (-3.00)*** | (1.17) | (-3.19)*** |
| Quoted spread | 0.914 | 0.461 | 0.871 | 0.960 | 0.612 | 0.957 |
| | (2.35)** | (1.28) | (2.25)** | (3.31)*** | (1.67)* | (3.26)*** |
| Constrained spread | -2.841 | -1.162 | -2.976 | 11.113 | 1.316 | 11.515 |
| | (-4.07)*** | (-2.10)** | (-4.28)*** | (18.93)*** | (2.77)*** | (19.87)*** |
| Total $ volume | -1.124 | 0.357 | -1.138 | 2.915 | 1.512 | 2.892 |
| | (-5.92)*** | (2.74)*** | (-5.97)*** | (21.94)*** | (13.64)*** | (21.30)*** |
| Midquote volatility | 0.191 | 0.207 | 0.204 | -0.418 | -0.374 | -0.445 |
| | (8.59)*** | (11.30)*** | (9.07)*** | (-20.67)*** | (-21.31)*** | (-23.19)*** |
| Message-to-trade | 0.000 | 0.003 | 0.002 | 0.003 | -0.003 | 0.001 |
| | (0.02) | (0.61) | (0.41) | (1.23) | (-1.25) | (0.47) |
| $R^2$ | 0.05 | 0.01 | 0.06 | 0.19 | 0.03 | 0.20 |
| Fixed effects | None | Stock | Date | None | Stock | Date |

**Table 7**

**Effects of dark and block trading on price discovery shares of lit trade prices vs dark trade prices**

This table reports regression estimates using a stock-day panel, in which the dependent variables are the Yan-Zivot information leadership share ($ILS_{LIT}$) and the Hasbrouck information share ($IS_{LIT}$) of lit trade prices (trades executed in the transparent central limit order book) relative to dark and block trade prices (scaled up by a factor of 100). *DARK* and *BLOCK* are the percentage of the stock-day's total trades that are executed in venues with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | $ILS_{LIT}$ | | | $IS_{LIT}$ | | |
|---|---|---|---|---|---|---|
| Intercept | 59.551 | -0.347 | -5.680 | 84.903 | -0.696 | -7.977 |
| | (16.32)*** | (-2.04)** | (-32.59)*** | (14.32)*** | (-3.60)*** | (-17.89)*** |
| DARK | -0.056 | -0.043 | -0.061 | -0.018 | -0.031 | -0.020 |
| | (-7.05)*** | (-6.63)*** | (-7.78)*** | (-1.46) | (-3.62)*** | (-1.63) |
| BLOCK | -0.129 | -0.121 | -0.105 | -0.014 | -0.025 | -0.032 |
| | (-12.51)*** | (-11.88)*** | (-10.78)*** | (-0.97) | (-1.82)* | (-2.48)** |
| Market capitalization | 0.037 | 2.054 | 0.103 | -2.213 | -6.416 | -2.203 |
| | (0.14) | (2.74)*** | (0.38) | (-4.24)*** | (-5.67)*** | (-4.28)*** |
| Quoted spread | -6.261 | -3.690 | -6.131 | -15.033 | -12.274 | -15.182 |
| | (-16.42)*** | (-6.20)*** | (-15.92)*** | (-20.31)*** | (-9.18)*** | (-20.99)*** |
| Constrained spread | 5.322 | 0.104 | 5.333 | -6.644 | -6.621 | -4.839 |
| | (8.45)*** | (0.14) | (8.81)*** | (-4.99)*** | (-4.98)*** | (-3.69)*** |
| Total $ volume | 2.205 | 1.812 | 2.145 | 3.063 | 3.130 | 2.953 |
| | (13.21)*** | (10.94)*** | (12.72)*** | (9.20)*** | (10.67)*** | (8.94)*** |
| Midquote volatility | 0.218 | 0.176 | 0.204 | 0.654 | 0.448 | 0.585 |
| | (10.32)*** | (8.78)*** | (9.67)*** | (13.73)*** | (11.28)*** | (13.22)*** |
| Message-to-trade | 0.021 | -0.016 | -0.022 | 0.130 | 0.200 | -0.021 |
| | (0.81) | (-0.69) | (-1.00) | (3.11)*** | (6.11)*** | (-0.52) |
| $R^2$ | 0.13 | 0.01 | 0.13 | 0.21 | 0.03 | 0.22 |
| Fixed effects | None | Stock | Date | None | Stock | Date |

**Table 8**
**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.

| | Mean | Std. dev. | P25 | Median | P75 |
|---|---|---|---|---|---|
| $PriceImpact_{LIT}$ (bps/$10,000) | 4.46 | 8.87 | 0.74 | 2.09 | 5.27 |
| $PriceImpact_{DARK}$ (bps/$10,000) | 4.00 | 5.47 | -0.06 | 0.03 | 0.66 |
| $PriceImpact_{BLOCK}$ (bps/$10,000) | 0.16 | 3.05 | -0.02 | 0.01 | 0.12 |

**Table 9**
**Effects of dark and block trading on informativeness of trades**

This table reports regression estimates using a stock-day panel, in which the dependent variables are permanent price impacts (a measure of trade informativeness) for lit central limit order book trades ($PriceImpact_{LIT}$), dark trades ($PriceImpact_{DARK}$), and block trades ($PriceImpact_{BLOCK}$). The permanent price impacts are calculated from the cumulative impulse response functions from a vector auto-regression model and are measured in basis points per $10,000 of unanticipated volume. *DARK* and *BLOCK* are the percentage of the stock-day's total trades that are executed in venues with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | $PriceImpact_{LIT}$ | | |
|---|---|---|---|
| Intercept | 41.302 | 0.044 | 5.108 |
| | (17.64)*** | (1.39) | (13.37)*** |
| DARK | 0.036 | 0.042 | 0.032 |
| | (6.15)*** | (11.51)*** | (5.57)*** |
| BLOCK | 0.082 | 0.062 | 0.092 |
| | (12.11)*** | (12.85)*** | (15.74)*** |
| Market capitalization | 0.913 | -0.378 | 0.792 |
| | (2.62)*** | (-0.91) | (2.41)** |
| Quoted spread | -1.155 | -0.997 | -1.218 |
| | (-4.24)*** | (-2.55)** | (-4.59)*** |
| Constrained spread | -1.192 | -0.333 | -0.727 |
| | (-2.06)** | (-0.85) | (-1.21) |
| Total $ volume | -2.855 | -2.678 | -2.652 |
| | (-11.30)*** | (-16.83)*** | (-11.01)*** |
| Midquote volatility | 0.460 | 0.385 | 0.423 |
| | (13.81)*** | (13.70)*** | (13.52)*** |
| Message-to-trade | 0.031 | 0.039 | 0.003 |
| | (1.61) | (2.85)*** | (0.17) |
| $R^2$ | 0.51 | 0.26 | 0.46 |
| Fixed effects | None | Stock | Date |

| | $PriceImpact_{DARK}$ | | | $PriceImpact_{BLOCK}$ | | |
|---|---|---|---|---|---|---|
| Intercept | 59.262 | 0.115 | 4.782 | 1.262 | 0.005 | 0.188 |
| | (9.05)*** | (1.09) | (5.33)*** | (4.68)*** | (0.47) | (4.54)*** |
| DARK | -0.128 | -0.127 | -0.126 | 0.001 | 0.002 | 0.001 |
| | (-8.27)*** | (-5.87)*** | (-8.26)*** | (1.29) | (1.84)* | (0.87) |
| BLOCK | -0.021 | 0.022 | 0.020 | -0.004 | -0.004 | -0.002 |
| | (-0.45) | (0.88) | (0.60) | (-5.63)*** | (-4.30)*** | (-3.84)*** |
| Market capitalization | 0.143 | -0.546 | 0.126 | -0.042 | -0.099 | -0.045 |
| | (0.30) | (-0.46) | (0.27) | (-1.46) | (-0.79) | (-1.54) |
| Quoted spread | -3.191 | -1.868 | -3.275 | -0.076 | -0.034 | -0.080 |
| | (-4.23)*** | (-1.77)* | (-4.62)*** | (-2.15)** | (-0.29) | (-2.34)** |
| Constrained spread | 1.398 | -1.098 | 1.562 | -0.082 | -0.075 | -0.066 |
| | (0.62) | (-0.65) | (0.72) | (-1.10) | (-0.52) | (-0.91) |
| Total $ volume | -3.122 | -1.937 | -2.874 | -0.044 | -0.003 | -0.033 |
| | (-9.29)*** | (-5.11)*** | (-8.78)*** | (-1.90)* | (-0.09) | (-1.43) |
| Midquote volatility | 0.368 | 0.293 | 0.364 | 0.032 | 0.031 | 0.031 |
| | (5.44)*** | (4.49)*** | (5.29)*** | (3.61)*** | (3.39)*** | (3.47)*** |
| Message-to-trade | 0.060 | 0.025 | 0.100 | -0.009 | -0.010 | -0.003 |
| | (0.91) | (0.42) | (1.89)* | (-2.35)** | (-2.97)*** | (-1.03) |
| $R^2$ | 0.02 | 0.01 | 0.02 | 0.01 | 0.00 | 0.01 |
| Fixed effects | None | Stock | Date | None | Stock | Date |

**Table 10**
**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. *DARK* and *BLOCK* are the percentage of the stock-day's total trades that are executed in venues with little or no pre-trade transparency (including trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, respectively. *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. 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | Log quoted spread | | |
|---|---|---|---|
| Intercept | 8.180 | -4.105 | -4.105 |
| | (79.9)*** | (-94.8)*** | (-201)*** |
| DARK | 0.007 | 0.002 | 0.007 |
| | (11.8)*** | (8.33)*** | (11.58)*** |
| BLOCK | 0.005 | 0.002 | 0.005 |
| | (9.29)*** | (3.06)*** | (9.03)*** |
| Market capitalization | -0.188 | -0.501 | -0.186 |
| | (-6.81)*** | (-15.68)*** | (-6.76)*** |
| Constrained spread | 0.129 | -0.874 | 0.143 |
| | (2.22)** | (-26.9)*** | (2.44)** |
| Total $ volume | -0.251 | -0.091 | -0.253 |
| | (-15.68)*** | (-10.89)*** | (-15.6)*** |
| Midquote volatility | 0.031 | 0.012 | 0.031 |
| | (23.3)*** | (14.1)*** | (22.4)*** |
| Message-to-trade | 0.000 | 0.000 | 0.000 |
| | (1.25) | (1.14) | (1.21) |
| $R^2$ | 0.72 | 0.08 | 0.70 |
| Fixed effects | None | Stock | Date |

**Table 11**

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

This table reports estimates from a two-stage instrumental variables regression. *DARK* and *BLOCK* are the percentage of the stock-day's total trades that are executed in venues with little or no pre-trade transparency (including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades) and large negotiated off-market trades, respectively. In the first stage *DARK* and *BLOCK* are regressed on the instrumental variables *DARK_not_i* and *BLOCK_not_i*, respectively, and a set of control variables. The instrumental variables 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 a set of 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. 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 clustered both by stock and by date are reported in parentheses. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels.

| | $Autocorrelation_{Factor}$ | | | $VarianceRatio_{Factor}$ | | |
|---|---|---|---|---|---|---|
| Intercept | 0.840 | -0.055 | -0.106 | 2.034 | -0.027 | -0.077 |
| | (0.76) | (-2.85)*** | (-1.37) | (2.54)** | (-1.51) | (-1.23) |
| DARK (fitted) | 0.041 | 0.018 | 0.041 | 0.028 | 0.011 | 0.029 |
| | (17.8)*** | (9.65)*** | (17.5)*** | (17.8)*** | (7.99)*** | (17.7)*** |
| BLOCK (fitted) | -0.012 | -0.017 | -0.015 | -0.007 | -0.011 | -0.009 |
| | (-5.02)*** | (-8.75)*** | (-6.14)*** | (-3.68)*** | (-7.62)*** | (-5.39)*** |
| Market capitalization | -0.305 | -0.472 | -0.302 | -0.327 | -0.375 | -0.323 |
| | (-2.52)** | (-4.33)*** | (-2.46)** | (-3.79)*** | (-4.33)*** | (-3.70)*** |
| Quoted spread | -0.217 | -0.619 | -0.221 | -0.301 | -0.358 | -0.329 |
| | (-1.75)* | (-6.11)*** | (-1.79)* | (-3.28)*** | (-5.06)*** | (-3.62)*** |
| Constrained spread | -0.555 | -2.069 | -0.620 | -0.448 | -1.831 | -0.383 |
| | (-2.39)** | (-12.7)*** | (-2.57)** | (-2.41)** | (-15.1)*** | (-1.96)** |
| Total $ volume | 0.778 | 1.088 | 0.791 | 0.661 | 0.875 | 0.653 |
| | (14.4)*** | (31.6)*** | (14.5)*** | (16.1)*** | (37.5)*** | (15.8)*** |
| Midquote volatility | 0.033 | 0.038 | 0.033 | -0.001 | -0.002 | -0.004 |
| | (9.22)*** | (10.50)*** | (8.66)*** | (-0.48) | (-1.39) | (-1.79)* |
| Message-to-trade | 0.004 | 0.004 | 0.004 | 0.003 | 0.003 | 0.003 |
| | (1.58) | (1.51) | (1.59) | (1.60) | (1.51) | (1.61) |
| $R^2$ | 0.06 | 0.04 | 0.06 | 0.10 | 0.04 | 0.10 |
| Fixed effects | None | Stock | Date | None | Stock | Date |

**Table 11 (continued)**

| | *Delay* | | |
|---|---|---|---|
| Intercept | 91.99 | 0.031 | 0.189 |
| | (37.1)*** | (0.52) | (1.15) |
| DARK | 0.045 | 0.003 | 0.046 |
| | (8.70)*** | (1.41) | (8.86)*** |
| BLOCK | -0.005 | -0.003 | -0.008 |
| | (-0.77) | (-0.59) | (-1.38) |
| Market capitalization | -2.022 | -0.781 | -1.897 |
| | (-7.06)*** | (-4.78)*** | (-6.69)*** |
| Quoted spread | 3.150 | 1.596 | 3.314 |
| | (11.2)*** | (8.47)*** | (11.81)*** |
| Constrained spread | 6.938 | 4.040 | 6.109 |
| | (12.61)*** | (9.90)*** | (11.2)*** |
| Total $ volume | -0.625 | 0.169 | -0.589 |
| | (-5.40)*** | (3.49)*** | (-5.29)*** |
| Midquote volatility | -0.085 | -0.041 | -0.051 |
| | (-9.73)*** | (-7.88)*** | (-6.87)*** |
| Message-to-trade | 0.000 | 0.000 | 0.001 |
| | (1.27) | (0.43) | (2.07)** |
| $R^2$ | 0.17 | 0.01 | 0.17 |
| Fixed effects | None | Stock | Date |

Panel A: Dollar volume


Panel B: Number of trades

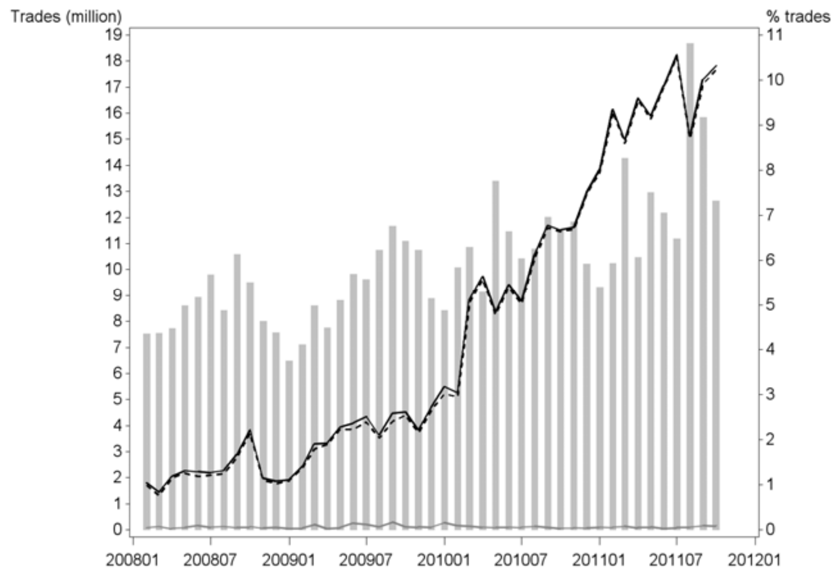



**Figure 1. Dollar volume and trades**

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 all trades that are executed with little or no pre-trade transparency, including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades, not including *Block* trades, which are large negotiated off-market trades. Panel B presents the same information but measures volume in number of trades (in millions) in place of dollars.

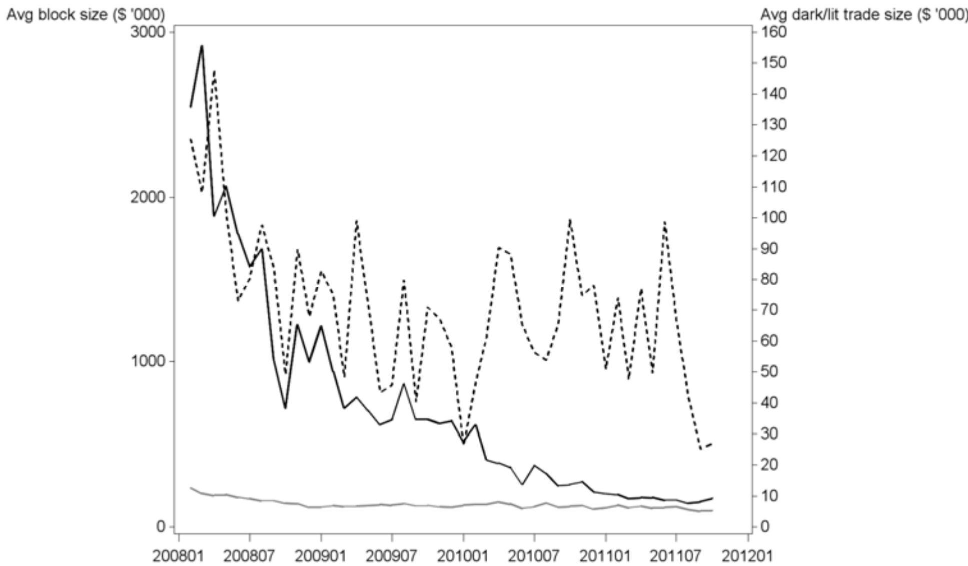
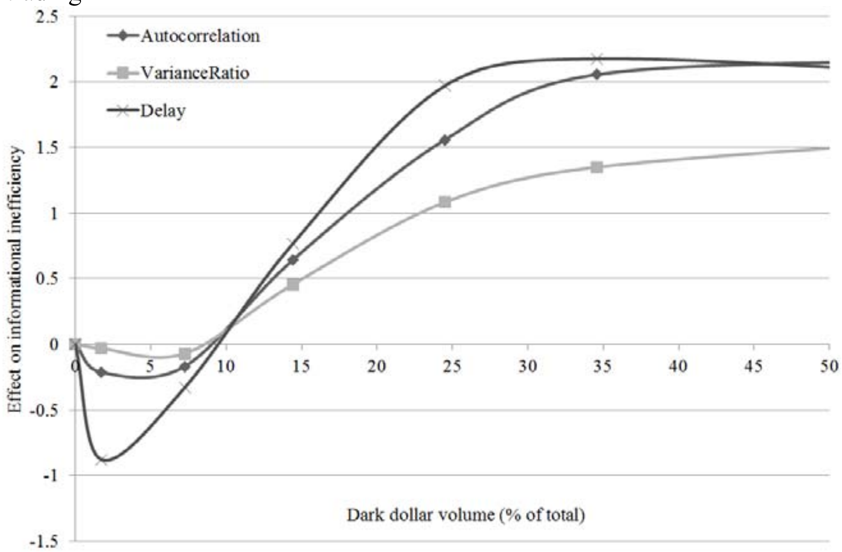



**Figure 2. Mean trade sizes**

This Figure plots the mean size (in $'000) of lit trades (solid grey line), dark trades (solid black line) and block trades (dashed black line) for our sample of stocks (All Ordinaries index constituents) during the sample period. Dark trades are all trades that are executed with little or no pre-trade transparency, including dark pool trades, broker internalization of order flow and *ASX Centre Point* trades, not including *Block* trades, which are large negotiated off-market trades.

Panel A: Dark trading




Panel B: Block trading

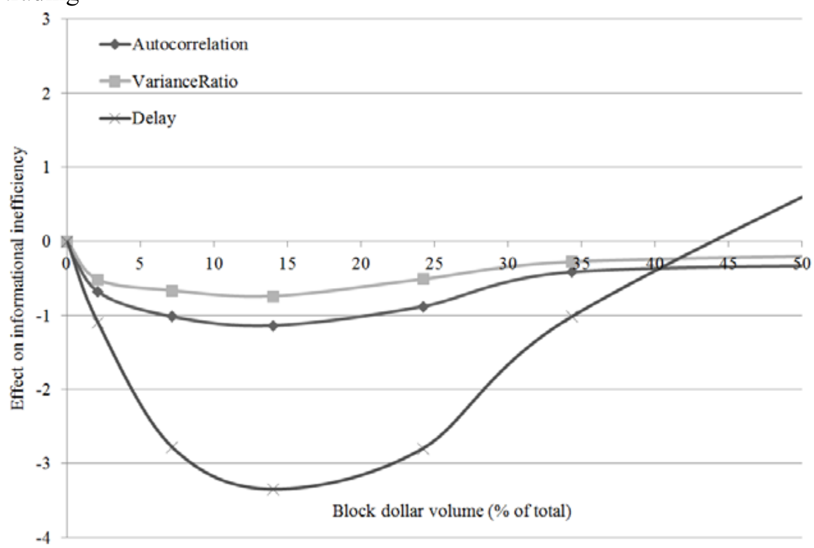



**Figure 3. Effects of dark and block trading on informational efficiency**

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 three informational inefficiency measures (larger values indicate greater informational *inefficiency*). 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.