Diving Into Dark Pools[1]

Sabrina Buti

Rotman School of Management
University of Toronto

Barbara Rindi

Bocconi University
IGIER and Carefin

and

Ingrid M. Werner

Fisher College of Business
The Ohio State University

Date: September 7, 2011

[1] We are grateful to SIFMA for assisting us in collecting the Dark Pool share volume data that forms the basis for this study. We also want to thank Jamie Selway for making this study possible, Kewei Hou for helpful comments on the empirical design, and participants at the Hong Kong University of Science and Technology Symposium and the Notre Dame Conference on Current Topics in Financial Regulation for comments.

ABSTRACT

This paper examines unique data on dark pool activity for a large cross-section of US stocks in 2009. Dark pool activity is concentrated in liquid stocks. NASDAQ (AMEX) stocks have significantly higher (lower) dark pool activity than NYSE stocks controlling for liquidity. For a given stock, dark pool activity is significantly higher on days with high share volume, high depth, low intraday volatility, low order imbalances relative to share volume, and low absolute returns. Results show that increased dark pool activity improves market quality measures such as spreads, depth, and volatility. The relationship between dark pool activity and measures of price-efficiency is more complex.

1. INTRODUCTION

There are several reasons for why institutional traders may want to avoid displaying their orders in the continuous limit order market. Order display invites imitation, potentially reducing the alpha of the underlying investment strategy. Displayed orders also invite front running and quote matching by broker-dealers as well as by opportunistic traders, resulting in higher trading costs. Further, traditional order display is associated with direct broker involvement, generating significant commission costs. Institutional traders worry about counterparty risk, the risk of trading against informed order flow especially order flow from proprietary trading desks. Institutional sized orders also face another problem; average trade and order sizes have fallen dramatically in recent years, making it virtually impossible to trade in size in the continuous limit order market.

It is therefore not surprising that there is a growing demand for trading venues that make it possible for institutions to keep their orders secret, offer low commission rates, maximizes the chances of trading with other institutions (as naturals), and allow institutions to trade in size at the mid-quote. Such non-displayed pools of liquidity have been present in US equity markets for a very long time. Examples include reserve and hidden orders within exchanges' and Electronic Communication Networks' (ECNs) trading systems, floor broker orders and specialist capital on floor-based exchanges, working orders handled by agency brokers or broker-dealers, dealer capital and stand-alone as well as broker and exchange/ECN operated crossing networks.[2] More recently, non-displayed liquidity pools such as Internalization Pools and Ping Destinations have been added to the list. Nowadays opaque sources of liquidity are often grouped under a single label (with unfortunate nefarious connotations): Dark Pools.

In broad brush terms, dark pools are characterized by limited or no pre-trade transparency, anonymity, and derivative (almost exclusively mid-quote) pricing. However, they differ in terms of whether or not they attract order flow through Indications of Interests (IOIs)/advertising and whether or not they allow interaction with proprietary and black box order flow.[3] It is difficult to accurately measure the amount of volume that is actually matched through Dark Pools but estimates range from 8-9% of share volume.[4]

[2] Sofianos (2007).
[3] See Mittal (2009) for a discussion of Dark Pool characteristics.
[4] Rosenblatt Securities, Inc. started tabulating monthly share volume for Dark Pools of Liquidity in its Trading Talk publication in March 2008 and TABB Group started its Liquidity Matrix publication in April 2007. Efforts to track volume in these venues are problematic due to a lack of uniform Dark Pool reporting standards.

The SEC has recently openly criticized the impact of dark pools on the price discovery process. In May, 2009, James Brigagliano, SEC's Division of Trading and Markets, said dark pools could impair price discovery by drawing valuable order flow away from the public quoting markets. "To the extent that desirable order flow is diverted from the public markets, it potentially could adversely affect the execution quality of those market participants who display their orders in the public markets," he said. He added that anything that "significantly detracts from the incentives to display liquidity in the public markets could decrease that liquidity and, in turn, harm price discovery and worsen short-term volatility."[5]

SEC Chairman Mary Schapiro announced on June 19th, 2009, that the SEC is "taking a serious look at what regulatory actions may be warranted" and that she has asked SEC staff to review ways to "best bring light" to dark pools.[6] In testimony before the House Committee on Financial Services' Subcommittee on Capital Markets, Insurance and Government-Sponsored Enterprises, on July 14, 2009, Chairman Schapiro further stated: "We have heard concerns that dark pools may lead to lack of transparency, may result in the development of significant private markets that exclude public investors (through the use of 'indications-of-interest' that function similar to public quotes except with implicit pricing), and may potentially impair the public price discovery function if they divert a significant amount of marketable order flow away from the more traditional and transparent markets."

In its recent *Concept Release on Equity Market Structure* (SEC, 2010), the SEC raises concerns about the consequences of a rising dark pool market share on public order execution quality and price discovery. In Congressional testimony, Dr. Hatheway (Nasdaq OMX) speaks to this issue and argues that when stocks experience "dark" trading in excess of 40 percent of total volume, execution quality begins to deteriorate. Weaver (2011) studies broader measures of market fragmentation and also argues that dark trading is associated with a reduction in market quality. In contrast O'Hara and Ye (2011) find that fragmentation of trading generally reduces transactions costs and increases execution speed. These contradictory results are not surprising as the researchers rely on very imprecise proxies for dark trading. The O'Hara and Ye (2011) study focuses on the effect of fragmentation on market quality during 2008 and uses TRF volume as a proxy without even netting out fully transparent venues such as BATS and DirectEdge. The same strategy is used by Weaver (2011), but his sample is more recent, from

[5] Chapman, Peter, SEC Worried About Dark Pools, Traders Magazine, July 2009.
[6] David Scheer and Jesse Westbrook, SEC May Force More Disclosure About 'Dark Pools,' Schapiro Says, Bloomberg.com, June 19th, 2009.

October 2009. The Nasdaq OMX study uses TRF volume minus BATS and DirectEdge as a proxy for dark pools, but this still includes internalized order flow.

To better inform the regulatory debate, we use more granular data to empirically assess the relationship between dark pools on market quality and price discovery. Specifically, the Securities Industry and Financial Market Association (SIFMA) solicited daily stock-level dark pool share volume data for the 2009 calendar year from all their members operating dark pools. The reporting was completely voluntary, and in the end SIFMA collected data on daily single-counted share volume from eleven dark pools on our behalf. The data is anonymous, and no attempt to study the data by individual dark pools will be made.

This study will focus on answering three questions:

1. How does dark pool market share vary across stocks and time?
2. Is Dark Pool volume associated with lower market quality?
3. Is Dark Pool volume associated with impaired price efficiency?

There is very limited empirical evidence on dark pool activity in the cross-section and the time-series. A few studies have focused on crossing networks. Gresse (2006) finds that crossing networks have a very limited market share and do not have a detrimental effect on the liquidity of the continuous market. Conrad, Johnson, and Wahal (2003) find that institutional orders executed in crossing networks have significantly lower realized execution costs and that traders use the continuous market to trade their exhaust. Naes and Odegaard (2006) find that institutional orders sent first to crossing networks and then to the continuous market obtain lower realized execution costs for the crossed component, but not necessarily for the entire order. Fong, Madhavan, and Swan (2004) find no evidence of a liquidity drain away from the continuous market when traders can trade in a crossing network. The only empirical study that we are aware of that takes a more comprehensive look at dark pools is by Ready (2010). He studies monthly volume by stock in three dark pools: Liquidnet, POSIT, and Pipeline during June 2005-September 2007. He finds that the market share of these dark pools is less than one percent of consolidated volume, and that dark pool volume is concentrated in liquid stocks (low spreads, high share volume). Two more recent papers by Brandes and Domowitz (2010) and Buchanan et al (2011) study dark pool trading in Europe and find that increased participation of dark pools is beneficial for price discovery and that it enhances price discovery process.

Our sample has several advantages compared to the Ready (2010) sample: it covers more dark pools, includes daily share volume data, and is more recent. Nevertheless, several caveats apply. First of all, the SIFMA dark pool data covers only those eleven dark pools that voluntarily responded to the data request. According to the SECs 2010 Concept Release on Market Structure, there are approximately 32 active Dark Pools during our sample period. Hence, our sample of eleven dark pools captures only roughly 1/3rd of Dark Pools operating in the US equity market. Second, to our knowledge there is no publicly available data on Dark Pools which makes it difficult to check the SIFMA data for accuracy. To gauge the coverage of our data, we compare it to monthly data reported by Rosenblatt, Inc. However, we note that this source is based on a combination of self-reported data and Rosenblatt estimates. Third, while our data permits a study of both time-series and cross-sectional variation in dark pool activity for the SIFMA sample of dark pools, we have no way of knowing if these eleven dark pools represent the same fraction of dark pool activity over stocks and over time. Therefore, we cannot claim that the variation in dark pool activity within our sample is representative of the entire population of dark pools. These caveats should all be kept in mind when drawing conclusions based on the SIFMA data.

We describe our sample construction in Section 2, and provide a univariate analysis of dark pool activity in Section 3. Descriptive statistics for our explanatory variables are in Section 4. Our analysis of the dark pool activity in the cross-section and in the time series is in Section 5. In Section 6, we study the relationship between dark pool activity and measures of market quality. Section 7 explores the relationship between dark pool activity and price efficiency. Section 8 concludes.

2. SAMPLE CONSTRUCTION

We first benchmark the raw SIFMA data against the monthly total share volume in dark pools as reported by Rosenblatt, Inc. in their monthly *Let There Be Light* publication. Figure 1 shows that the SIFMA data mirrors the monthly time series variation in the Rosenblatt share volume pretty closely. Figure 2 shows that Dark Pool share volume as reported in the SIFMA (Rosenblatt) data represents 3.65 (7.74) percent of consolidated volume in January, and 6.10 (10.15) percent of consolidated volume in December. Finally, Figure 3 shows that the SIFMA data covers roughly half of the Rosenblatt share volume. Specifically, the market share of the dark pools submitting data for our study increases from 47% in January to 60% in December.

The raw SIFMA data covers 10,178 unique securities and the coverage by individual Dark Pools ranges from a low of 5,646 to a high of 8,251 securities. In order to merge the SIFMA data with data from TAQ, CRSP, etc., we screen the data following standard practice as summarized in Table 1. We first exclude 1,525 ticker symbols with suffixes (e.g., preferred, warrants, non-voting, etc) and the ticker symbols with a fifth character (unless also in CRSP as A, B, or K). Second, we exclude 4,035 stocks that are not common stocks (SHRCD 10 or 11) covered by CRSP. As we need to merge CRSP with the SIFMA data, we also exclude 87 stocks with missing ticker symbols in CRSP and 49 stocks with duplicate stock identifiers (permno or cusip) for the same ticker symbol. Our SIFMA sample consists of 4,482 stocks.

We also create subsamples that are similar to the samples used by Weaver (2011) and O'Hara and Ye (2011) to benchmark our data against previous samples. Weaver (2011) excludes stocks with price above $1,000 and O'Hara and Ye (2011) exclude stocks with price below $5.00 and with less than 1,000 shares average daily volume. The discussion of these samples and replication of methodologies pursued by previous authors is in the Appendix.

3. UNIVARIATE STATISTICS

To examine the cross-sectional distribution of dark pool activity, we compute dark pool volume (DPVOL) as the number of shares per stock per day (single-counted) that execute in one of our eleven dark pools. We also compute the fraction of daily consolidated share volume (VOL) as reported in CRSP that was executed in one of the dark pools as 100*DPVOL/VOL for every stock in our sample. This variable will be labeled RELDP. Further, we count the number of different dark pools that are active in a stock on a given day and call this variable COUNTDP. To get a better sense of the degree of competition among dark pools, we compute the inverse Herfindahl index (IHERF) based daily stock-level dark pool market shares. Recall that if the market shares are evenly distributed across dark pools, IHERF will be equal to COUNTDP. IHERF will be lower than COUNTDP the more concentrated dark pool trading activity is for a given stock day.

We report the overall results in Table 2, Panel A. Dark pool volume represents on average 4.51 percent of consolidated volume. Dark pool activity is skewed as the median is lower, at 3.05 percent. On average almost half the SIFMA reporting dark pools (5.27) are active in a stock on any given day. However, dark pool activity is concentrated based on the inverse of the Herfindahl Index (IHERF=2.43).

Previous research has found significant differences across NASDAQ and NYSE when it comes to fragmentation. Specifically, fragmentation has been found to be higher for small stocks on NASDAQ by both O'Hara and Ye (2011) and Weaver (2011). To examine the extent to which the SIFMA sample has a similar pattern, we compare dark pool activity across primary listing venues in Table 2, Panels B, C, and D. Dark pool activity is much lower on AMEX/ARCA with 1.87 percent of consolidated volume on average. In fact, the median stock day in this subsample has no dark pool activity. NASDAQ dark pool activity is 4.32 percent and NYSE dark pool activity is 5.49 percent of consolidated volume on average. Again, the distributions are skewed, particularly on NASDAQ. This is not surprising as the NASDAQ sample includes many low priced stocks. Recall from Table 1 that 2,254 stocks in the overall sample have a price below $5.00 and these are mostly listed on NASDAQ. The median NYSE stock has as many as nine out of eleven active dark pools trading on any given day. However, note that dark pool activity is more concentrated based on the Inverse Herfindahl Index (IHERF=3.35).

Finally, we subsample based on market capitalization to show how dark pool activity varies with firm size. We sort stocks on market capitalization based on the number of shares outstanding multiplied by the closing price from CRSP. SMALL capitalization stocks have market capitalization less than $50 million, MEDIUM capitalization stocks have market capitalization between $50 million and $1 billion, and LARGE capitalization stocks have market capitalization above $1 billion.

In Table 2, Panel E, we find that there is relatively limited dark pool activity for SMALL capitalization stocks, 1.82 percent of share volume and only 0.97 active dark pools on average. By contrast, the MEDIUM capitalization category in Panel F has more dark pool activity on average, 5.11 percent, but there is also much more variation across stocks and days. Moreover, there appears to be more specialization for this group of stocks judging by the distribution of COUNTDP and the IHERF. Panel G of Table 2 shows that dark pool activity is highest for the LARGE capitalization stocks with an average RELDP of 5.74 percent. For LARGE capitalization stocks, 75 percent of the stock days have nine or more active dark pools. In other words, dark pools appear to compete intensively for this group of stocks. However, the dark pool share volume is much more concentrated based on the Inverse Herfindahl index (median IHERF=3.70).

4. DESCRIPTIVE STATISTICS

Our first goal is to examine both what factors explain the cross-sectional distribution and the time-series evolution of dark pool activity. To do so, we gather additional information for our sample stocks from CRSP and from TAQ. We get daily market capitalization, share volume, closing stock price, intraday price range (defined as (high-low)/high) based on quotes, stock returns and market (S&P 500) returns from CRSP. We compute daily time-weighted quoted and share-weighted effective spreads, bid depth at the National Best Bid Offer (NBBO), (bid) depth, (buy) order imbalances (defined as the absolute value of (buys-sells)/share volume where buys are classified using a modified Lee and Ready (1991) algorithm),[7] and the standard deviation of mid-quote returns from TAQ. Table 3 provides the descriptive statistics for our SIFMA sample.

We have over one million stock-day observations in the SIFMA sample. The average firm in our sample has a $2.6 billion market capitalization. The average stock in our sample has a price of 40 dollars and trades 1.6 million shares per day. The average quoted depth is 124 shares, and the average quoted spread is 175 basis points or 13 cents. The average effective spread is 41 basis points or 2.6 cents.

5. DETERMINANTS OF DARK POOL ACTIVITY

To better understand how dark pool activity varies with market characteristics, we first sort stocks every day into quintiles based on a particular market characteristic. We then compute the daily average dark pool activity, RELDP, and the average number of active dark pools, COUNTDP, for each quintile portfolio. This gives us 252 daily observations of means for RELDP and COUNTDP for each quintile. We test whether RELDP or COUNTDP are higher for the fifth quintile (High) than for the first quintile (Low) portfolio based on a particular market characteristic using a time-series t-test of the difference in means. The market characteristics are market capitalization, volume, price, intraday range, absolute return, spreads, depth, order imbalance, and the standard deviation of mid-quote returns. The results are in Table 4.

Panel A of Table 4 shows that dark pool activity is higher for the fifth quintile than for the first quintile based on firm size, volume, and price. Dark pool activity is significantly higher for the low spread portfolio than for the high spread portfolio. By comparison, the differences across quintile portfolios based on depth are small and insignificant. Dark pool activity is significantly lower for the high volatility than for the low volatility portfolio. The results also show that dark pool activity is significantly lower on days with high order imbalances relative to share volume. This makes sense as the likelihood of getting an order executed in a dark pools should be lower when the market is one-sided, i.e., when there is significant buying or selling pressure in the market. Finally, dark pool activity is significantly

[7] We classify trades as buys (sells) if the execution price is above (below) the mid-quote in effect at the time of the trade, and use a tick-test to classify trades that execute at the mid-quote.

lower for stock days with large absolute returns than for stock days with small absolute returns. This is consistent with the result that dark pool activity is significantly lower for high volatility portfolios.

We report the corresponding results for the number of active dark pools, COUNTDP, by quintile portfolios in Panel B of Table 4. The results are very similar to those reported for RELDP. Specifically, more dark pools are active for: large firms, stocks with high share volume, high price, low spreads, and low volatility. Fewer dark pools are active for stock-days with large relative order imbalances and on days with large absolute returns.

Our next step is to examine the cross-sectional and time series variation in dark pool activity. We explore the cross-sectional variation using monthly Fama-Macbeth cross-sectional regressions with RELDP on the left hand side and a number of stock and market characteristics on the right hand side. The average monthly estimated coefficients are reported in Table 5. The t-statistics are based on the Newey-West adjusted standard errors.

In our first specification (1), we control for listing exchange by including a dummy variable for Nasdaq-listing and one for AMEX/ARCA –listing. We also control for the logarithm of market capitalization. The results show that dark pool activity is increasing in market capitalization and is higher (lower) for Nasdaq (AMEX) stocks than for NYSE stocks after controlling for market capitalization. In specification (2) we replace market capitalization with share volume and price, and the results show that dark pool activity is increasing in share volume and price. In other words, more liquid stocks have more dark pool activity. We add the quoted spread in cents and the bid depth in specification (3) as added measures of liquidity and find that stocks with narrower quoted spreads holding listing exchange, share volume, and price constant have higher dark pool activity. More depth is also associated with more dark pool volume, but the effect is not significant. We replace quoted spread in cents and price with quoted spread in basis points in specification (4) and find that this variable is highly statistically significant. Stocks with narrower basis point spreads have more dark pool activity, controlling for listing exchange and share volume. Note also that with this measure of spreads, the coefficient on bid depth is statistically significant and positive. Finally, in specification (5) we drop share volume and include the relative order imbalance in percent of share volume and volatility as measured by the intraday range divided by the high as reported by CRSP. We find that dark pool volume decreases significantly in relative order imbalances and volatility. For robustness, we also rerun specification (5) on the O'Hara and Ye (2011) sample and report the results in column (6). Our conclusions are generally robust to applying their sample screens (excluding low priced and low volume stocks), but for this sample the associations between dark pool activity and bid depth and between dark pool activity and volatility are not statistically significant.

In sum, the multivariate Fama-Macbeth regression analysis shows that dark pool activity is significantly higher (lower) for NASDAQ (AMEX) stocks than for NYSE stocks all else equal. Liquid stocks have more dark pool activity as predicted by Buti, Rindi, and Werner (2010). Stocks with higher price have more dark pool activity than low-price stocks. We also find that dark pool activity is higher for stocks with narrow quoted spreads and high inside bid depth. These results confirm that dark pools are more active the higher the degree of competition in the limit order book as predicted by Buti, Rindi, and Werner (2010). Dark pool activity is higher for stocks with low intraday volatility as measured by the intraday range. Finally, dark pool activity is significantly higher for stocks with low average order imbalances relative to share volume. As mentioned in the discussion of the univariate results, this makes sense as the likelihood of getting an order executed in a dark pool should be lower when the market tends to be one-sided.

We explore the time-series variation in dark pool activity in Table 6. We de-mean all variables to take stock fixed effects into account and cluster standard errors by firm and day. Specification (1) shows that dark pool activity is significantly higher on days with higher than average share volume, narrower quoted spreads and higher bid depth. The results are very similar when we add order imbalances relative to share volume in specification (2). The new variable has a statistically significant and negative coefficient which means that dark pool activity is low on days with unusually large order imbalances relative to share volume. As mentioned above, this is natural as it is more difficult to obtain an execution in a dark pool when the market is one-sided. We add both intraday range and the absolute return in specification (3) and the coefficients are both significant and negative. In other words, dark pool activity is significantly lower on unusually volatile days and on days with unusually large market moves. Note that with these additional variables included in the panel regressions, the sign of the coefficient on quoted spread flips – the coefficient is now positive and significant. The most likely explanation for this sign reversal is that the quoted spread and volatility tend to be positively related for a particular stock. Moreover, days with unusually large amounts of uncertainty tend to be days with unusually wide spreads for a particular stock. Finally we include lagged dark pool activity and lagged absolute returns in specification (4). The results show that unusually large lagged dark pool activity is associated with unusually large contemporaneous dark pool activity, i.e., dark pool activity is auto-correlated. This result is consistent with Buti, Rindi, and Werner (2010) who predict that dark pools generate a liquidity externality effect. Furthermore, we find that large lagged absolute returns are associated with lower dark pool activity. For robustness, we rerun this specification for the O'Hara and Ye (2011) sample and the results are in column (5). Our conclusions are robust to applying their sample screen (excluding low price and low volume stocks), but note that the association between the quoted spread and dark pool activity is not statistically significant for the O'Hara and Ye (2011) sample.

In sum, the time-series analysis shows that after controlling for stock fixed effects, days with unusually high share volume, unusually high bid depth, unusually low degree of one-sided order flow, and unusually low volatility tend to have higher dark pool activity. These results make sense as it is more likely that dark pool orders execute when trading interest is high and two-sided (balance between buyers and sellers). The relationship between dark pool activity and quoted spreads is more complex. As described in Buti, Rindi, and Werner (2010), wider quoted spreads makes it relatively more attractive to send an order to a dark pool that would execute at the mid-quote instead of sending a marketable order to the limit order book and incur the spread. At the same time, a wider spread makes it more attractive for a patient trader to submit a limit order to the book. In equilibrium, Buti, Rindi, and Werner (2010) show that the latter effect dominates so that an unusually wide spread is predicted to discourage dark pool order submission. This theoretical prediction is consistent with the result that an unusually wide spread is associated with unusually low dark pool activity (specifications (1) and (2)). Similarly, as explained by Buti, Rindi, and Werner (2010), higher limit order bid depth reduces the incentives for an institution to submit a limit order relative to submitting an order to a dark pool. The reason is that the limit order would have to compete with the orders already in the limit order book, reducing the probability of the order getting filled without offering price improvement. Finally, during periods of unusually high volatility traders are all else equal more likely to forgo the uncertain executions associated with dark pools and instead rely on marketable orders to gain immediacy. However, controlling for volatility (specifications (3) and (4)), an unusually wide spread is associated with more dark pool activity (i.e., a substitution away from marketable orders to dark pool orders).

Having analyzed the cross-sectional and time-series patterns of dark pool activity as captured by the SIFMA sample, we now move on to examining the relationship between dark pool activity and market quality and price efficiency, respectively.

6. DARK POOLS AND MARKET QUALITY

A central question is whether there are any detrimental effects of dark pool activity on public market quality. This question is challenging to answer as causality is notoriously difficult to prove. In our case, this is particularly complicated as dark pool activity and market quality measures are jointly determined. For example, the theoretical model developed in Buti, Rindi, and Werner (2010) predicts that dark pool market share is higher when limit order depth is high, when limit order spreads are narrow, and when the tick size is larger. In other words, strategic traders decide whether to submit an order to a dark pool or to the public limit order book based on observing the depth and the spread. Therefore, we cannot simply run a regression of contemporaneous market quality measures on dark pool activity and interpret the coefficients as evidence of a causal relationship.

To deal with the inherent endogeneity of dark pool activity and market quality, we need to find good instruments for dark pool activity and market quality respectively. In a recent paper studying the impact of low latency trading on market quality, Hasbrouck and Saar (2011) propose using low latency trading in other stocks during the same time period as an instrument for low latency trading in a particular stock. We follow their suggestion and use dark pool trading for other stocks (*not i*) on day *t* as an instrument for dark pool trading in stock *i*. We refine their instrument slightly by requiring that the other stocks (*not i*) be listed on the same exchange as stock *i,* that their market capitalization is in the same size-grouping (LARGE, MEDIUM, SMALL) as stock *i,* and that they are in the same two-digit SIC code. The idea is that we have observed that there are systematic differences between exchanges and across size grouping in dark pool trading. The matching on SIC code serves to control for industry effects. We use the same logic in creating instruments for each of our market quality measures: the time-weighted percent and cent quoted spread, the share-weighted percent and cent effective spread, the (log of) time-weighted bid-depth, (log of) share volume, the standard deviation of mid-quote returns, and the intraday range divided by the intraday high. [8]

We estimate a two-equation simultaneous model for dark pool activity (RELDP) and market quality measures (MQMs) using Two Stage Least Squares (2SLS). Specifically, we estimate the following system of equations:

$$MQM_{i,t} = a_1\, RELDP_{i,t} + a_2\, MQMnot_{i,t} + e_{1,t}$$

$$RELDP_{i,t} = b_1\, MQM_{i,t} + b_2\, RELDPnot_{i,t} + e_{2,t}$$

[8] Hasbrouck and Saar (2011) were able to use the spreads for other markets quoting the same security in their analysis of low latency orders on the NASDAQ. We unfortunately do not know in which market dark pool trades are executed so we cannot follow their strategy.

As instruments for $RELDP_{i,t}$, we use $RELDPnot_{i,t}$ which is the average dark pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code. Note that we exclude stock *i*. Similarly, as an instrument for $MQM_{i,t}$, we use $MQMnot_{i,t}$, which is the average market quality measure for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code. We again exclude stock *i*.

Since both $RELDP_{i,t}$ and $MQM_{i,t}$ are endogenous in this system, the 2SLS estimation of the first equation involves replacing $RELDP_{i,t}$ with the fitted value from a regression of $RELDP_{i,t}$ on $MQM_{i,t}$ and $RELDPnot_{i,t}$. Similarly, the estimation of the second equation involves replacing $MQM_{i,t}$ by the fitted value from a regression of $MQM_{i,t}$ on $RELDP_{i,t}$ *and* $MQMnot_{i,t}$. We estimate the above system of equations for all stocks and days in a panel. To control for stock fixed effects, we demean all variables by deducting the in-sample average and divide the demeaned variables by their in-sample standard deviation. As a result, the estimated coefficients can be interpreted as the response to a one standard deviation shock.

Table 7 reports the results from estimating the simultaneous equation model for the SIFMA sample in Panel A. We are primarily interested in the a_1 and b_1 coefficients: a_1 measures the effect of dark pool activity on market quality and b_1 measures the effect of market quality on dark pool activity. The estimated a_1 coefficients show that dark pool is significantly negatively related to quoted and effective spreads. For example, a one standard deviation increase in dark pool activity is associated with a 0.119 (0.357) standard deviation decrease in the quoted (effective) percent spread. Further, we find that dark pool activity is significantly positively related to bid-depth and significantly negatively related to volatility. A one standard deviation increase in dark pool activity is associated with a 0.092 standard deviation increase in bid-depth and a 0.283 standard deviation reduction in the intraday range. Interestingly, dark pool activity is significantly *negatively* related to consolidated share volume: a one standard deviation increase in dark pool activity is associated with a 0.233 standard deviation decrease in share volume. Similarly, the estimated b_1 coefficients show that poorer market quality (wider spreads, lower depth, more volatility) is significantly negatively related to dark pool activity as predicted by Buti, Rindi, and Werner (2010). The coefficients (a2 and b2) on our instruments are positive and highly significant across the board. In other words, they appear to be good instruments.

We repeat the analysis for NYSE-listed stocks in Panel B and NASDAQ-listed stocks in Panel C of Table 7. The results are very similar to the results for the overall sample. However, the magnitude of the effect of dark pool activity on quoted spreads is several times larger for NYSE-listed than NASDAQ-listed stocks (e.g., the coefficients for percent quoted spreads are -0.179 for NYSE-listed and -0.050 for NASDAQ-listed). The magnitude of the effect of dark pool activity on share volume is also much larger for NYSE-listed than for NASDAQ-listed stocks (-0.306 for NYSE-listed and -0.126 for NASDAQ-listed stocks). Interestingly, the effect of dark pool activity on volatility is similar for NYSE-listed and NASDAQ-listed stocks.

To investigate whether Table 7 masks systematic differences across stocks with different liquidity, we repeat the simultaneous equation analysis for stocks sorted by size-grouping in Table 8. Recall that the size-groupings are: SMALL with a market capitalization less than $50 million, MEDIUM with a market capitalization between $50 million and $1 billion, and LARGE with a market capitalization of $1 billion and above. The results in Table 8 show that dark pool activity is associated with better market quality for all size groupings. In fact, the positive effect of dark pool activity on market quality is generally stronger for SMALL stocks in Panel A than for MEDIUM and LARGE stocks in Panels B and C. For example, a one standard deviation increase in dark pool activity is associated with a 0.793 (0.780) standard deviation decrease in quoted (effective) percent spreads and a 0.380 standard deviation increase in bid-depth for SMALL caps. The corresponding numbers for LARGE caps are 0.077 (0.333) and 0.072. The magnitude of the effect of dark pool activity on volatility is also much larger for SMALL caps than LARGE caps: a one standard deviation increase in dark pool activity for a SMALL cap results in a 0.843 standard deviation reduction in the intraday range. The corresponding number for a LARGE cap is 0.215. Finally, note that more dark pool activity is associated with significantly *higher* share volume for SMALL caps, but significantly lower share volume for MEDIUM and LARGE caps.

For robustness, we estimate the simultaneous equation system stock-by-stock. The results of this estimation are summarized in Table 9. We report the median estimated coefficient and the p-values from a rank test which tests whether the coefficients are different from zero. The results are weaker than in Table 7, Panel A, but the conclusions are the same: dark pool activity is associated with better market quality as measured by spreads, bid-depth, and volatility. However, as noted above, dark pool activity appears to be associated with lower consolidated share volume.

7. DARK POOLS AND PRICE EFFICIENCY

In the previous section, we showed that increased dark pool activity leads to an improvement in measures of market quality such as spreads, depth, and volatility. However, it is possible that dark pools could harm other aspects of market quality such as price efficiency. To study the relationship between dark pool activity and the efficiency of market prices, we rely on three standard measures of price efficiency: short-term volatility, return autocorrelations, and the variance ratio. Short-term volatility is here the variance of mid-quote log returns measured over 15-minute and 30-minute intervals.[9] Short-term volatility can be viewed as a measure of trading frictions, and a market with lower volatility is viewed as more efficient in this context. Return autocorrelations are simply the first order autocorrelation of the 15-minute log returns. In an efficient market, returns should be uncorrelated since prices should follow a random walk. In other words, markets with return autocorrelations close to zero are considered more efficient in that price changes are less predictable. Finally, the variance ratio (see Lo and MacKinlay (1988)) is defined as the absolute value of the ratio of the variance of the 30-minute log returns divided by the two times the variance of the 15-minute log returns. The closer this number is to one, the more prices behave like a random walk and hence the more efficient is the market.

There are alternative measures of price efficiency in the literature. For example, Hasbrouck (1993) suggests using a decomposition approach to measure price efficiency. His approach uses signed order flow to distinguish the noise variance component (related to frictions and hence inefficiency) from the information-based variance component. As emphasized by O'Hara and Ye (2011), this approach is less appropriate for studying today's fragmented trading environment. The approach also requires the researcher to classify trades as buyer and seller initiated, which is increasingly difficult to do in a reliable fashion. Therefore, we follow O'Hara and Ye (2011) and concentrate on our three simple price efficiency measures.

We divide each trading day into 26 15-minute intervals starting at 9:30am. We first calculate both the log 15-minute returns and the (overlapping) log 30-minute returns. The short-term volatility is defined as the standard deviation of the 15-minute (30-minute) log returns for each stock and month (year).[10] As mentioned above, a market with lower short-term volatility is considered to be more efficient. Return autocorrelations are estimated monthly for each stock based on the 15-minute log returns. A market with return autocorrelations closer to zero is considered to be more efficient. Significantly positive return autocorrelations (continuations) suggests that prices under-react, while negative return autocorrelations (reversals) suggest that prices over-react.

[9] These measures complement our previously calculated volatility measures: the intraday range (low frequency measure) and the standard deviations of mid-quote returns (ultra high-frequency measure).
[10] We exclude the overnight return.

Slightly more work is required to conduct the variance ratio test. The idea behind this test is that if prices follow random walks the variance of a 30-minute log return should be twice as large as the variance of a 15-minute log return. Following Lo and MacKinlay (1988), we correct for the bias induced by using overlapping returns. We also correct for the bias in estimating the variance of returns before computing the monthly variance ratio. Specifically, we first compute the mean 15-minute return as $\hat{\mu} = \frac{1}{nq}\sum_{k=1}^{nq}(lnP_k - lnP_{k-1})$, where P_k is the mid-quote at the end of interval *k* and *nq+1* is the number of mid-quote observations in the sample. The estimator of the 15-minute log return variance is given by $\bar{\sigma}_1^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(lnP_k - lnP_{k-1} - \hat{\mu})^2$ and the estimator for half of the 30-minute log return variance is given by $\bar{\sigma}_2^2(q) = \frac{1}{m}\sum_{k=2}^{nq}\left(lnP_k - lnP_{k-q} - q\hat{\mu}\right)^2$ where $m = q(nq - q + 1)\left(1 - \frac{q}{nq}\right)$ and in our case, *q*=2. We define the variance ratio statistic as: $\overline{VR}(q) = \frac{\bar{\sigma}_2^2(q)}{\bar{\sigma}_1^2} - 1$.

Lo and MacKinlay (1988) show that a transformation of the variance ratio asymptotically follows a standard normal distribution:

$$ratio \;\equiv \theta(q) \equiv \sqrt{nq}\left(\overline{VR}(q)\right)\left(\frac{2(2q-1)(q-1)}{3q}\right)^{-\frac{1}{2}} \sim N(0,1)$$

where *nq* is the number of observations and *q* is the number of periods in the longer-horizon return, in our case *q*=2. We compute the variance ratio for each stock in our sample for each month (year). A variance ratio close to zero indicates that the market is efficient. If the variance ratio is significantly positive, the 30-minute variance is higher than twice the 15-minute variance, which suggests that the market price under-reacts. By contrast, if the variance ratio is significantly negative, the 30-minute variance is lower than twice the 15-minute variance, which suggests that the market price over-reacts. In other words, the market displays “excess” short-term volatility.

We first report descriptive statistics for our price efficiency measures in Table 10. Recall that we calculate (12) monthly observations for each stock of short-term volatility, the variance ratio, and estimate autocorrelation based on daily mid-quote return data. The average standard deviation of 15-minute mid-quote returns is lower than the average standard deviation of 30-minute mid-quote returns, but not by a factor of two. Indeed, the variance ratio is negative suggesting that stocks on average over-react to information. In other words, the 15-minute return volatility is too high relative to the 30-minute return volatility. A similar conclusion can be drawn from the negative average autocorrelation of daily closing mid-quote returns. Under the null-hypothesis of market efficiency, both the variance ratio and the return-autocorrelation should be zero. In other words, either a positive or a negative deviation from zero implies that the market is inefficient. Therefore, we henceforth define our second and third price-efficiency measures as the absolute value of the variance ratio and the absolute return-autocorrelation respectively.

In Table 11, we explore how dark pool activity varies with our measures of price efficiency. We sort stocks into price-efficiency quintiles and then examine dark pool activity (RELDP) on average for each quintile in Panel A, and how many dark pools (COUNTDP) are active for each quintile on average in Panel B. Stocks with higher short-term volatility have significantly lower dark pool activity and fewer active dark pools than those with lower short-term volatility. Similarly, stocks with higher absolute variance ratios and larger absolute return- autocorrelations have significantly less dark pool activity and a significantly lower number of active dark pools. The differences in the High-Low columns are highly statistically significant. In other words, there is more dark pool activity for more efficient stocks.

The third question our paper seeks to answer is the effect of dark pool activity on price efficiency. From Table 11, we know that dark pools are more active in stocks with more efficient prices. We would like to answer how an unusual amount of dark pool activity relates to price efficiency, taking the potential joint determination of dark pool activity and price efficiency into account. We therefore again estimate a simultaneous equation model using 2SLS with dark pool activity and our four price efficiency measures: standard deviation of 15-minute returns, standard deviation of 30-minute returns, the absolute variance ratio, and the absolute return-autocorrelations. We follow the same strategy as in the previous section and use the dark pool activity for other stocks listed on the same exchange, from the same size grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code as an instrument for dark pool activity in stock *i*. Similarly, we use the average price efficiency measures for other stocks as instrument for price efficiency measures for stock *i*. The results are reported in Table 12 for the SIMFA Sample in Panel A, and for NYSE-listed stocks and NASDAQ-listed stocks separately in Panels B and C.

As mentioned before, we are mostly interested in a_1 and b_1. The estimates of a_1 and b_1 are negative and statistically significant for three out of our four price-efficiency measures in all three Panels. The results show that more dark pool activity is associated with lower short-term volatility and that lower short-term volatility is associated with more dark pool activity. Similarly, more dark pool activity is associated with lower absolute return-autocorrelations. However, note that the reduction in the standard deviation of 15-minute returns is not large enough relative to the reduction in the standard deviation of 30-minute returns for the absolute variance ratio to fall. Therefore, the results show that more dark pool activity is associated with a higher absolute variance ratio. Analysis of the signed variance ratio (unreported) suggests that this result is driven by an increase in short-term overreaction associated with more dark pool activity. The estimated coefficients for instruments a_2 and b_2 are positive and highly significant in all regressions suggesting that our instruments work well.

To check if there are differences across stocks by liquidity, we repeat the exercise for stocks grouped by market capitalization in Table 13. The results for SMALL stocks are in Panel A, and for MEDIUM and LARGE stocks in Panels B and C respectively. Overall, the conclusion from Table 14 holds for all subsamples by market capitalization. However, the results show that the effect of dark pool activity on the absolute variance ratio is much smaller for LARGE stocks than for MEDIUM and SMALL stocks. For example, when dark pool activity increases by one standard deviation the absolute variance ratio increases by 0.241 standard deviations for SMALL caps, by 0.206 standard deviations for MEDIUM caps, but only by 0.101 standard deviations for LARGE cap stocks. The absolute value of the autocorrelation of 15-minute returns decreases significantly for SMALL caps (-0.318) and MEDIUM caps (-0.158) while the effect on LARGE caps is very limited (-0.026) and insignificant. Overall, the effects of dark pool activity on price efficiency appear to be relatively limited for LARGE caps but larger and more significant for MEDIUM caps.

Finally, we estimate the simultaneous system of equations stock-by-stock and report the results in Table 14. Based on the stock-by-stock analysis, we conclude that both short-term volatility measures decline significantly in dark pool volume. However, as before, the absolute variance ratio increases significantly suggesting that more dark pool activity. By contrast, there is no significant effect of dark pool activity on the absolute return-autocorrelations. Hence, the stock-by-stock results generally support the conclusion from the panel regressions, but the results are as expected weaker statistically.[11]

In sum, our results show that increased dark pool activity improves price efficiency as measured by short-term intraday volatility and absolute return-autocorrelations. However, it also appears that increased dark pool activity contributes to higher absolute variance ratios. Based on unreported results, the link between dark pool activity and the absolute variance ratio appears to be primarily driven by an increase in short-term overreaction (more negative variance ratio). We plan to investigate these results further in the near future.

[11] There are only 12 monthly observations per firm, rendering the sample for each 2SLS estimation small.

8. SUMMARY AND CONCLUSIONS

In this paper, we study dark pool trading activity for a large cross-section of stocks based on a unique self-reported sample of daily dark pool share volume during 2009. The sample was collected by SIFMA and covers eleven out of roughly 32 dark pools active in the US equity markets during our sample period. We find that our SIFMA sample represents roughly 50 to 60 percent of dark pool volume as reported by Rosenblatt Securities Inc. The market share of reported dark pools increase over the sample period, from slightly below 4 percent of consolidated share volume in January to above 6 percent of consolidated share volume in December. Moreover, we note that SIFMA sample dark pools report activity in over 10,000 distinct securities. For individual dark pools, this figure ranges from a low of 5,646 to a high of 8,251 securities. In other words, the dark pools in our sample are active for a very large cross-section of stocks.

The average daily market share of our SIFMA dark pools based on the benchmark sample of common stocks also in CRSP with non-zero share volume is 4.5 percent of share volume. If we exclude low price and low liquidity stocks, dark pool activity increases to 5.3 percent of share volume on average. While we do not have data on all dark pools, we surmise that the overall market share of dark pools is roughly twice as large based on the overall market share of our SIFMA reporting dark pools.

We examine whether dark pools specialize by computing the number of different dark pools active on the typical stock-day as well as the inverse of the Herfindahl index which measures market concentration. The average stock-day in our SIFMA screened sample has five active dark pools, and the market-share equivalent number of dark pools is 2.4. If we screen out low-price stocks, the median stock-day has 8 active dark pools with a market share equivalent number of dark pools of 3.1. In other words, there is significant competition among dark pools for institutional order flow.

We study dark pool activity separately for stocks based on the primary listing exchange and based on market capitalization. Generally, we find that dark pool activity is higher for the NYSE (5.5 percent of share volume) than for Nasdaq (4.3 percent of share volume). There are also more active dark pools for a typical NYSE stock-day (9) than for a Nasdaq stock-day (4). SIFMA dark pool activity is strongly increasing in market capitalization, with a market share of 1.8 percent for firms below $50 million, 5.1 percent for firms between $50 million and $1 billion, and 5.7 percent for firms with market capitalization above $1 billion. For firms above $1 billion, a typical stock-day has ten active SIFMA dark pools with a market share equivalent number of dark pools of 3.7.

In a preliminary analysis of dark pool activity and market quality, we sort stocks into quintiles by dark pool activity and test for differences in market quality measures between the group of high dark pool and the group of low dark pool activity. We find strong evidence that stocks with high dark pool activity are significantly larger, more liquid stocks with higher average price. Stocks with higher dark pool activity are also associated with lower quoted and effective spreads, lower intraday volatility, and lower measures of absolute buy-sell imbalances relative to share volume. From this analysis, we cannot conclude that dark pool activity causes higher market quality as we have not yet controlled for characteristics that are likely to affect market quality such as market capitalization and price.

Taken together, our univariate results confirm aggregate market statistics from for example Rosenblatt Securities Inc. indicating that dark pool activity is a significant component of equity trading in US markets. Moreover, our results show that dark pool activity is concentrated in stocks with higher market capitalization and higher price. Statistics commonly referred to in the regulatory debate do not address this cross-sectional variation in dark pool activity so we have no benchmarks to compare our study to in this regard.

A unique feature of the SIFMA sample is that it permits us to examine the cross-sectional and time-series variation in dark pool activity at a more granular level. Our cross-sectional analysis shows that dark pool activity is increasing in average share volume and price, but is decreasing in average quoted and effective spreads, average intraday volatility, average absolute order imbalances relative to share volume. We also find that dark pool activity is higher (lower) for NASDAQ (AMEX) stocks controlling for size, share volume, and stock price. In the time-series, we find that dark pool activity is significantly higher on days with unusually high share volume, unusually narrow quoted spreads, unusually high depth at the inside, and unusually low intraday volatility for a particular stock. We find that dark pool activity is lower on days with more imbalanced order flow and larger absolute return for a particular stock. In other words, holding the stock constant, dark pool activity is lower when the market is one-sided.

Given that dark pool activity is not only significant on average, but also displays significant cross-sectional and time-series variation, it is clearly important to understand how dark pool activity is related to measures of market quality and market efficiency. We investigate this important question using a simultaneous equation system to account for the fact that market quality and dark pool activity are jointly determined. Our results show that more dark pool activity is associated with better market quality: narrower spreads, more depth, and lower volatility. These results are robust to sub-sampling by listing exchange, market capitalization, and to using stock-by-stock instead of panel regression estimation. By contrast, we find that more dark pool activity is generally associated with lower share volume, suggesting that dark pool trading has a crowding-out effect overall and for both NYSE-listed and

NASDAQ-listed stocks. This effect is notably absent for SMALL cap stocks where dark pool activity is instead associated with higher share volume.

Finally, we use the same simultaneous equation system estimation to investigate the effect of dark pool activity on measures of price efficiency. Our results show that more dark pool activity is associated with lower short-term volatility and lower absolute return-autocorrelations across the board. Moreover, based on these price-efficiency measures, the beneficial effects of dark pool activity are largest for smaller capitalization stocks. The absolute variance ratio instead suggest that higher dark pool activity is associated with less efficient prices, and that the relationship is stronger for small capitalization stocks than for large and medium caps. We conclude that a further investigation into the relationship between dark pool activity and price efficiency based on alternative measures of price efficiency is warranted.

In future work, we also plan to conduct an analysis based on Rule 605 market quality data which has the advantage of being measured at the order level as opposed to the execution point which is the case for TAQ data.

References

Buti, S., Rindi, B., and I.M. Werner, 2010, Dark Pool Trading Strategies, Fisher College of Business working paper, June.

Brandes, Y. and I. Domowitz, 2010, Alternative Trading Systems in Europe, Trading Performance by European Venues Post MiFID, Working Paper, ITG, Inc., May.

Buchanan, M., J. Tse, D. Vincent, X. Lin, and A. Kumar, 2011, Measuring Dark Pools' Impact, Portfolio Strategy, Credit Suisse Europe, January.

Concept Release on Equity Market Structure, Securities and Exchange Commission, 17CFR PART 242 [Release No. 34-61358; File No. S7-02-10].

Conrad, J., Johnson, K., and S. Wahal, 2003, Institutional Trading and Alternative Trading Systems, *Journal of Financial Economics* 70, 99-134.

Davies, R.J., and S.S. Kim, 2008, Using Matched Samples to Test for Differences in Trade Execution Costs, *Journal of Financial Markets*.

Fong, K., Madhavan, A., and P. Swan, 2004, Upstairs, Downstairs: Does the Upstairs Market Hurt the Downstairs? Working Paper, University of New South Whales.

Gawronski, J., and J. Schack, 2010, Let There be Light, Rosenblatt's Monthly Dark Liquidity Tracker, Special Issue: 2009 In Review, Rosenblatt Securities Inc., January.

Gresse, C., 2006, The Effect of Crossing-Network Trading on Dealer Market's Bid-Ask Spreads, *European Financial Management* 12, 143-160.

Hasbrouk , J., 1993, Assessing the Quality of a Security Market: A New Approach to Transaction- Cost Measurement, *Review of Financial Studies* 6, 191-212.

Hasbrouck, J., and G. Saar, 2011, Low Latency Trading, Working Paper, New York University.

Hatheway, F., 2009, Testimony for the Senate Banking, Housing, and Urban Affairs Committee: The Securities Insurance and Investment Subcommittee, Wednesday, October 28, 2009.

Lo, A., and A. C. MacKinlay, 1988, Stock Market Prices Do Not Follow Random Walks: Evidence from a simple Specification Test, *Review of Financial Studies*, 1, 41-66.

Lee, S., 2009, Dark pools 2009: Not So Dark Anymore, Aite Group White paper, September.

Lee, C., and M. Ready, 2009, Inferring Trade Direction from Intraday Data, *Journal of Finance* 46, 733-746.

Mittal, H., 2008, Are You Playing in a Toxic Dark Pool? A Guide to Preventing Information Leakage, ITG, June.

Naes, R., and B. Odegaard, 2006, Equity Trading by Institutional Investors: To Cross or Not to Cross? *Journal of Financial Markets* 9, 79-99.

O'Hara, M., and M. Ye, 2011, Is Market Fragmentation Harming Market Quality? Cornell University *Journal of Financial Economics* 100, 459-474.

Ready, M., 2010, Determinants of Volume in Dark Pools, working paper, University of Wisconsin, November.

Sofianos, G., 2007, Dark Pools and Algorithmic Trading, Chapter 6 in *Algorithmic Trading: A Buy-Side Handbook*.

Weaver, D., 2011, Off-exchange Reporting and Market Quality in a Fragmented Market Structure, Rutgers Business School Working Paper, May.

APPENDIX

A. BENCHMARK SAMPLES

Appendix Table 2 compares the SIFMA sample to samples used in papers studying dark pools (fragmentation). The Weaver (2011) sample excludes a handful of very high priced stocks, but the distributions of dark pool activity looks very similar. By contrast, O'Hara and Ye (2011) exclude low priced stocks and the dark pool activity is considerably higher for this subsample (5.27 percent).

Appendix Table 3 provides descriptive statistics for the two reference subsamples in Panels A and B. The main difference between the SIFMA sample and the Weaver (2011) sample is the average price: $39.78 in the SIFMA sample compared to $15.38 in the Weaver sample. Using the screens employed by O'Hara and Ye (2011) we lose almost half the observations, and the average price is now $74.67 as a result of excluding all stocks with prices below $5.00 (but including high-priced stocks). This sample also has substantially lower spreads and much lower volatility of mid-quote returns than the other two subsamples.

B. REPLICATING WEAVER (2011)

We follow Weaver (2011) in estimating a cross-sectional regression of average daily market quality measures on average daily dark pool activity controlling for factors that the previous literature has been shown to affect market quality measures in the cross-section. In addition to the market quality measures already discussed in previous sections, Weaver (2011) adds the Amihud measure (the average absolute daily return divided by share volume from CRSP) as a measure of price impact. The results are reported in Appendix Table 4. In order to compare our results to those of Weaver (2011), we conduct this analysis for several sub-samples: the SIFMA sample (Panel A); the Weaver (2011) sample (Panel B); the O'Hara and Ye (2011) sample; and for Nasdaq and NYSE stocks separately (Panels C and D).

Weaver (2011) runs regressions of market quality statistics on controls and a quadratic function in his measure of fragmentation, percent TRF volume. We examine the effect of different specifications of the functional form of the relationship between our measure of dark pool activity, RELDP, and market quality in Figure 4. On the left hand side, we have the quoted spread in percent and on the right hand side an intercept, share volume, price, and the standard deviation of mid-quote returns. We estimate this cross-sectional regression with: a linear RELDP; a quadratic function of RELDP; and a third-order polynomial in RELDP. The linear specification suggests a negative relationship between dark pool activity and quoted spreads. The quadratic specification instead suggests that quoted spreads decline initially as dark pool activity increases, but that beyond RELDP of about 8 percent, a higher amount of dark pool activity is associated with higher quoted spreads. By contrast, the third-order polynomial instead shows that there is no significant effect of increases in dark pool activity beyond a RELDP of roughly 10 percent. The reason for the different results is the data. As shown in Figure 5, the bulk of the data has a RELDP below 10 percent. Beyond this point, the estimates become unstable and unreliable. Based on this graphical illustration, we estimate third-order polynomials instead of quadratic functions of RELDP in our cross-sectional analysis. However, the conclusions are robust to both linear and quadratic specifications.[12]

Focusing on the effect of RELDP on market quality measures, we see in Panel A that virtually all market quality measures are declining significantly in RELDP. Moreover, both the positive quadratic and the negative third order RELDP terms are statistically significant. In other words, we find that the relationship between RELDP and market quality measures is non-linear as described in Figure 4. For completeness, we repeat this analysis for the different sub-samples and find very similar results. Statistically and in terms of magnitude, the relationship between dark pool activity and market quality is stronger for the SIFMA and Weaver samples than for the O'Hara and Ye samples, and it is also stronger for NASDAQ stocks than for NYSE stocks. Recall that Weaver (2011) excludes a handful of high priced stocks, while O'Hara and Ye (2011) instead exclude all stocks with price below $5.00 and average daily volume below 1,000 shares (but include high priced stocks). When high priced stocks like Berkshire Hathaway are included, price is an extremely important explanatory variable for the cross-section of cent spreads and we obtain unrealistically high t-statistics as a result. We conclude that in the cross-section, a higher amount of dark pool activity is associated with lower quoted and effective spreads, lower price impacts, and lower short-term volatility. In other words, more dark pool activity is generally associated with higher market quality.

Note that our results are different from those found by Weaver (2011) even when we select a sample very similar to the one he analyzes (Panel B). Recall that he finds that fragmentation is detrimental for market quality. His data is from October 2009, while ours is January through December 2009. We do not, however, believe that the difference in sample periods can explain the results. Rather, the most likely explanation for the different results is that the data he uses is TRF volume which includes not just dark pools but also internalized trades and trades using DirectEdge and BATS.

[12] These results are available from the authors on request.

Consequently, while fragmentation may be detrimental more generally based on the Weaver (2011) study, there is no evidence in Table 9 that more dark pool activity, and specifically more dark pool activity in the venues that participated in the SIFMA sample, is associated with poorer market quality.[13]

C. REPLICATING O'HARA AND YE (2011)

To benchmark our sample against the one studied by O'Hara and Ye (2011), we conduct a matched sample analysis similar to theirs for our SIFMA sample. We start with our version of the O'Hara and Ye (2011) sample, but constrain it further by only including stocks with either NASDAQ or the NYSE as their primary listing exchange and requiring at least 90 days of trading in both the first and the second half of 2009.[14] This leaves us with 954 NASDAQ stocks and 919 NYSE stocks. We collect the price and market capitalization for each firm in our sample as of January 2, 2009, and sort stocks separately for each exchange by size. Ten stocks are sampled randomly from each size decile to select 100 stocks from each exchange. For each stock *i* in our stratified random sample, we then search all stocks with the same primary listing exchange to find the stock *j* that minimizes the sum of the absolute percentage difference in size and the absolute percentage difference in price (Davies and Kim (2008)):

$$\mathrm{D_{i,j}} = \left|\mathrm{Size_i}/\mathrm{Size_j} - 1\right| + \left|\mathrm{P_i}/\mathrm{P_j} - 1\right|.$$

The matching is done by the exchange of primary listing to control for exchange-specific effects.

We proceed to compute the average daily dark pool activity (RELDP) based on our SIFMA data for both stocks in each matched pair, *i,j*, over January – June, 2009. The stock in the pair with the lowest dark pool activity is designated as a Low RELDP stock while the one with the highest dark pool activity is designated as the High RELDP stock in the pair. Having controlled for factors that should affect market quality statistics such as size, price, and primary listing exchange, we then conduct tests for differences in market quality over the July – December, 2009, between our High and Low RELDP samples. The results are in Appendix Table 5.

The table reports the overall results based on a battery of market quality statistics referred to previously in this paper. We conduct both a paired t-test for difference in means and a Wilcoxon signed

[13] It is tempting to conclude that it is internalization that produces the Weaver (2011) result that increased fragmentation is detrimental for market quality. However, as we do not believe that this analysis proves a causal relationship between fragmentation and market quality, such a conclusion is definitely premature.
[14] Our results are not sensitive to this requirement, but it avoids manually dealing with a few cases of stocks that are either IPOs or delisted in the initial sample.

rank test for differences in medians for each market quality statistic. Note that we have daily data for between 90-128 days for each stock for this test, so we have more power than previous researchers using monthly Rule 605 data.[15] The results show that stocks with high amounts of dark pool activity (RELDP) have significantly lower quoted and effective average percentage spreads. The differences are small, 3.51 basis points for the quoted spread and 0.53 basis points for effective half spreads. Similarly, stocks with high amounts of dark pool activity have lower average quoted and effective cent spreads. Stocks with high dark pool activity have significantly higher depth both at the bid and the offer than their matched stocks. In terms of volatility, there is no significant difference in our High-Low measures, but the standard deviation of mid-quote returns is significantly lower for stocks with high dark pool activity. Finally, our various imbalance measures suggests that stocks with high dark pool activity have larger absolute depth and order imbalances, but significantly lower order imbalances relative to share volume.

To verify that our aggregate results in Appendix Table 5 does not mask differences across large and small stocks, we repeat the analysis based on Large and Small capitalization stocks separately in Appendix Table 6. Note that our definition of Large is the top three deciles, while Small are the lowest three deciles in the size distribution. It is clear from the table that average quoted and effective percent spreads are significantly lower for stocks with high dark pool activity whether they are Large caps or Small caps. The magnitude of the differences is large for small caps, on the order of 7.6 basis points or 3.64 cents for the quoted spread. The picture is a bit more mixed for cent spreads. For Small caps stocks, cent spreads are also uniformly significantly lower for stocks with high dark pool activity. However, there is some evidence based on the Wilcoxon test of medians that cent spreads for Large caps may be somewhat higher. Note though that the magnitude of the difference is miniscule, on the order of 0.04 to 0.01 cents depending on the test. For both groups, stocks with higher dark pool activity have significantly more depth. Interestingly, the evidence suggests that stocks with higher dark pool activity have more intraday variation in prices based on the High-Low measures, but not based on the standard deviation of mid-quote returns. The imbalance measures are consistent with the overall sample, absolute depth and order imbalances are higher but relative order imbalances tend to be lower for stocks with higher dark pool activity.

---

[15] Conducting the paired t-tests based on differences in daily equally-weighted means for the high and low portfolios does not change the conclusions from Table 10.

We conclude that there is no evidence in Appendix Tables 5 and 6 that stocks with high dark pool activity have worse market quality than stocks with low dark pool activity, controlling for size, price, and market of primary listing. We also do not find any evidence that small capitalization stocks with higher dark pool activity have worse market quality. The same is true for large capitalization stocks. One caveat should be kept in mind. While we are eliminating look-ahead bias by matching stocks based on dark pool activity for January – June, 2009, and then estimating market quality over July – December, 2009, this methodology has the drawback that the distance between the matching and the market quality measures is significant. In our case, this should bias the results against finding that dark pool activity is associated with higher market quality if in reality the relationship was the opposite. The reason is that the activity in our SIFMA reporting dark pools, as well as the overall dark pool activity based on Rosenblatt Securities Inc., was rising significantly (Figure 2.). In other words, if more dark pool activity were to be associated with worse market quality, this would weaken our results. Yet, our results are highly significant.

D. PRICE EFFICIENCY FOR WEAVER (2011) AND O'HARA AND YE (2011) SUBSAMPLES

We analyze the relationship between dark pool activity and price efficiency following the methodology in previous literature in Appendix Tables 8 and 9. Appendix Table 8 follows Weaver (2011) and conducts cross-sectional regressions for the SIFMA and the O'Hara and Ye (2011) samples in Panels A and B respectively. We suppress the Weaver (2011) sample as the results are virtually identical to those for the SIFMA sample. Short-term volatility decreases in dark pool activity for the SIFMA sample in Panel A, but short-term volatility increases in dark pool activity for the O'Hara and Ye (2011) sample in Panel B. The variance ratio is not significantly increasing in dark pool activity for the SIFMA sample in Panel A, but does increase significantly for the O'Hara and Ye (2011) sample in Panel B. By contrast, the autocorrelation of returns increases significantly in dark pool activity for both subsamples. In other words, the evidence suggests that dark pool activity tends to be associated with improved market efficiency.

Appendix Table 9 replicates the analysis of dark pool activity for a set of matched stocks following the O'Hara and Ye (2011) methodology. Panel A shows the overall results, and Panel B shows the results for LARGE and SMALL cap stocks respectively. The results in both Panels show that stocks with high dark pool activity have significantly higher variance ratio and autocorrelation of returns. In other words, they display less evidence of short-term overreaction and prices are hence more efficient.

TABLE 1. SAMPLE CONSTRUCTION

Sample	Securities
SIFMA data	**10,178**
Exclude symbol.XX and NDQ 5th	1,525
	8,653
CRSP SHRCD=10, 11	4,035
	4,618
CRSP missing symbol	87
	4,531
Duplicate permno/cusip	49
SIFMA Sample	**4,482**
Stocks with Price > $1,000	3
Weaver Sample	**4,479**
Stocks with Price < $5.00	2,254
Stocks with volume < 1,000 shares/day	23
O'Hara and Ye Sample	**2,205**

TABLE 2. UNIVARIATE DARK POOL ACTIVITY

The table reports univariate statistics based on daily stock level data. DPVOL is SIFMA reported daily Dark Pool single-counted share volume per stock in thousands. RELDP is 100 times DPVOL divided by daily consolidated volume as reported in CRSP. COUNTDP is the number of active Dark Pools per day per stock. IHERF is the inverse of the Herfindahl Index for Dark Pools. The SIFMA sample is defined in Table 1. SMALL includes stocks with market capitalization less than $50 million, MEDIUM includes stocks with market capitalization between $50 million and $1 billion, and LARGE includes stocks with market capitalization of $1 billion and more.

A. SIFMA SAMPLE	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	85.08	794.03	0.21	5.55	37.90
RELDP	4.51	5.74	0.65	3.05	6.22
COUNTDP	5.27	3.97	1	5	9
IHERF	2.43	1.74	1.00	2.40	3.70
B. AMEX/ARCA (8%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	3.60	20.22	0.00	0.00	0.95
RELDP	1.87	5.19	0.00	0.00	1.58
COUNTDP	1.34	2.11	0	0	2
IHERF	0.84	1.14	0	0	1.34
C. NASDAQ (60%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	41.74	221.90	0.10	2.12	15.90
RELDP	4.32	6.19	0.15	2.45	5.87
COUNTDP	4.27	3.7	1	4	7
IHERF	2.13	1.7	1	2	3.34
D. NYSE (32%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	184.08	1,355.02	8.55	35.15	125.07
RELDP	5.49	4.64	2.51	4.48	7.16
COUNTDP	8.02	3.05	6	9	11
IHERF	3.35	1.46	2.33	3.32	4.34
E. SMALL (23%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	1.83	12.79	0.00	0.00	0.40
RELDP	1.82	5.47	0.00	0.00	1.24
COUNTDP	0.97	1.6	0	0	1
IHERF	0.68	0.98	0	0	1
F. MEDIUM (51%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	22.67	86.72	0.80	5.00	18.91
RELDP	5.11	6.23	1.27	3.31	6.68
COUNTDP	5.16	3.26	2	5	8
IHERF	2.56	1.53	1.47	2.5	3.6
G. LARGE (26%)	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	283.23	1,541.20	27.52	82.04	228.46
RELDP	5.74	3.87	3.14	4.98	7.40
COUNTDP	9.34	2.15	9	10	11
IHERF	3.73	1.34	2.79	3.7	4.63

TABLE 3. DESCRIPTIVE STATISTICS

The table reports descriptive statistics based on daily stock level data for the SIFMA sample. Market capitalization is in billion dollars, share volume is in million shares, and bid depth is in shares. Relative order imbalance is defined as the absolute value of buys - sells in percent of consolidated share volume., where buys (sells) are classified based on a modified Lee and Ready (1991) algorithm. Percent spreads are in basis points. (High-Low)/High measures and returns are multiplied by 100. Standard Deviation of midquote returns are

Daily Measures	**Observations**	**Mean**	**StDev**
Market Capitalization (CRSP)	1,011,760	2.59	12.45
Share Volume (CRSP)	1,011,760	1.59	13.26
Price (CRSP)	1,011,760	39.78	1485.09
(High-Low)/High (CRSP, quotes)	1,011,760	5.88	5.21
Absolute Return (CRSP)	1,011,643	0.27	6.47
Quoted Spread Basis Points (TAQ)	1,010,740	174.98	352.28
Quoted Spread Cents (TAQ)	1,010,740	13.23	246.70
Effective Spread Basis Points (TAQ)	1,001,464	41.25	127.95
Effective Spread Cents (TAQ)	1,001,464	2.61	1.08
Bid Depth (TAQ)	1,010,740	124.08	750.81
Relative Order Imbalance in Percent (TAQ)	1,001,464	20.42	23.92
StDev Midquote Returns (TAQ)	1,010,388	26.59	60.02

TABLE 4. DARK POOL ACTIVITY AND CONTEMPORANEOUS STOCK AND MARKET QUALITY

The table reports average dark pool activity for quintile portfolios sorted daily by market characteristics based on stock level data for the SIMFA sample. RELDP is 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. COUNTDP is the number of active Dark Pools per day per stock. Market quality statistics are described in Table 3. The table reports time-series averages of daily means based on daily sorts of stocks into quintile market characteristics portfolios. The t-tests are based on the time series of the difference in the daily average dark pool activity between the High and Low portfolios.

A. RELDP

Daily Measures	Low	2	3	4	High	High-Low	t-statistic
Market Capitalization (CRSP)	1.642	4.013	5.473	5.789	5.673	4.031	53.97
Share Volume (CRSP)	1.872	3.820	5.399	5.791	5.547	3.675	52.21
Price (CRSP)	2.311	4.430	5.115	5.571	5.164	2.853	38.15
(High-Low)/High (CRSP, quotes)	4.294	5.319	5.123	4.632	3.221	-1.073	-15.16
Absolute Return (CRSP)	4.702	5.040	4.875	4.537	3.438	-1.264	-17.92
Quoted Spread Percent (TAQ)	5.760	5.986	5.460	3.769	1.625	-4.134	-53.08
Quoted Spread Cents (TAQ)	4.818	5.348	5.302	5.542	2.590	-2.229	-31.89
Effective Spread Percent (TAQ)	5.161	6.063	5.688	4.051	1.848	-3.313	-42.58
Effective Spread Cents (TAQ)	4.415	5.726	5.268	4.592	2.809	-1.606	-22.22
Bid Depth (TAQ)	4.461	4.441	4.659	4.581	4.459	-0.002	-0.02
Relative Order Imbalance in Percent (TAQ)	5.231	5.334	5.161	4.504	2.551	-2.710	-36.81
StDev Midquote Returns (TAQ)	5.646	5.839	5.382	3.949	1.792	-3.854	-49.92

B. COUNTDP

Daily Measures	Low	2	3	4	High	High-Low	t-statistic
Market Capitalization (CRSP)	0.878	2.778	5.472	7.581	9.647	8.769	250.70
Share Volume (CRSP)	0.464	2.695	5.560	7.830	9.808	9.344	280.11
Price (CRSP)	2.163	4.042	5.229	6.996	7.929	5.766	130.88
(High-Low)/High (CRSP, quotes)	4.873	6.457	6.017	5.255	3.755	-1.118	-17.43
Absolute Return (CRSP)	5.147	5.907	5.794	5.353	4.158	-0.988	-14.95
Quoted Spread Percent (TAQ)	9.736	7.796	5.589	2.648	0.599	-9.136	-258.84
Quoted Spread Cents (TAQ)	7.668	6.936	6.014	4.327	1.422	-6.246	-140.06
Effective Spread Percent (TAQ)	7.943	7.901	6.087	3.597	1.083	-6.860	-154.97
Effective Spread Cents (TAQ)	4.904	7.700	6.836	5.180	1.989	-2.914	-54.26
Bid Depth (TAQ)	5.307	5.233	5.334	5.277	5.217	-0.090	-1.93
Relative Order Imbalance in Percent (TAQ)	7.462	7.141	6.171	4.355	1.482	-5.981	-136.08
StDev Midquote Returns (TAQ)	9.185	7.861	5.539	2.856	0.936	-8.249	-198.62

TABLE 5. DARK POOL ACTIVITY IN THE CROSS SECTION

The table reports the results of regressions of RELDP on contemporaneous market characteristics based on monthly cross-sectional Fama-Macbeth regressions for the SIFMA sample in columns (1) through (5) and for the O'Hara and Ye (2009) sample in column (6). RELDP is 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Market quality measures are described in Table 5. We take the log of market capitalization, volume, depth and price. Cent spreads are divided by 100. NASDAQ, and AMEX are dummy variables which take the value of 1 for stocks whose primary listing exchange is NASDAQ, and AMEX respectively. We report the average monthly coefficients on top and t-statistics below .

Explanatory variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Intercept	5.250	-3.029	-2.967	1.944	6.554	5.961
	20.67	-12.91	-12.47	6.48	10.77	10.42
NASDAQ	0.250	0.424	0.424	0.044	0.204	0.341
	4.80	6.88	7.13	0.32	2.38	7.17
AMEX	-1.320	-0.982	-0.973	-2.097	-1.352	-1.758
	-10.10	-16.84	-15.75	-15.51	-22.33	-14.00
Market Capitalization (CRSP)	0.660					
	10.29					
Share Volume (CRSP)		0.512	0.501	0.257		
		27.21	28.58	8.12		
Price (CRSP)		0.549	0.572			
		5.78	5.79			
Quoted Spread Cents (TAQ)			-0.090			
			-4.34			
Quoted Spread Percent (TAQ)				-0.315	-0.143	-0.420
				-15.80	-5.06	-21.18
Bid Depth (TAQ)			0.008	0.027	0.028	-0.010
			1.16	3.18	2.94	-1.19
Relative Order Imbalance in Percent (TAQ)					-0.057	-0.034
					-6.78	-3.82
(High-Low)/High (CRSP, quotes)					-0.152	-0.021
					-2.85	-0.62

TABLE 6. DARK POOL ACTIVITY IN THE TIME-SERIES

The table reports the results of regressions of RELDP on contemporaneous and lagged market characteristics based on panel regressions for the SIFMA sample in columns (1) to (4) and for the O'Hara and Ye (2009) sample in column (5). RELDP is 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Market quality measures are described in Table 5. We take the log of volume, depth, and order imbalance. All variables are de-meaned to control for stock fixed effects. We report the estimated coefficients on top and t-statistics below . T-statistics are based on standard errors clustered by stock and day.

Deviations from stock means	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Share Volume (CRSP)	0.335	0.318	0.509	0.437	0.760
	13.64	11.09	15.38	12.81	11.70
Quoted Spread Percent (TAQ)	-0.029	-0.028	0.026	0.026	0.021
	-5.12	-4.56	2.72	2.82	0.81
Bid Depth (TAQ)	1.346	1.372	1.068	0.813	1.321
	10.44	10.35	8.70	7.04	8.59
Relative Order Imbalance in Percent (TAQ)		-0.006	-0.009	-0.009	-0.007
		-10.14	-14.13	-14.76	-6.33
(High-Low)/High (CRSP, quotes)			-0.083	-0.073	-0.171
			-17.44	-16.78	-14.52
Absolute Return (CRSP)			-0.014	-0.012	-0.015
			-3.99	-4.01	-2.04
Lag RELDP				0.225	0.221
				43.40	36.04
Lag Absolute Return (CRSP)				-0.002	0.004
				-1.09	0.21
Number of Observations	1,010,710	1,004,410	1,001,295	996,879	505,006
Adjusted R-square	0.008	0.009	0.013	0.063	0.073

TABLE 7. SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND MARKET QUALITY BY EXCHANGE

The table reports the results of analyzing the relationship between Dark Pool activity and Market Quality. We measure Dark Pool activity as RELDP, which is defined as 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Market quality measures are described in Table 3. Due to the potential simultaneity between market quality and dark pool activity, we estimate the following two-equation simultaneous equation model for RELDP and the following market quality measures (MQMs): Time-Weighted Quoted Spread in Percent and Cents; Share-weighted Effective Spread in Percent and Cents; (log of) Time-weighted Bid-depth; (log) of Share volume; Standard deviation of Mid-quote Returns; and High-Low/High:

$$MQM_{i,t} = a_1 RELDP_{i,t} + a_2 MQMnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 MQM_{i,t} + b_2 RELDPnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock *i*). Similarly, as an instrument for $MQM_{i,t}$ we use $MQMnot_{i,t}$, which is the average Market Quality Measure for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model by pooling observations across all stocks and days in the sample. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the estimated coefficients on top and p-values below . The results for the SIFMA sample are reported in Panel A, for Nasdaq stocks in Panel B, and for NYSE stocks in Panel C.

A. SIFMA Sample

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Time-weighted Quoted Spread in Basis Points	-0.119	0.887	-0.102	0.300
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	-0.099	0.839	-0.064	0.329
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.357	0.551	-0.227	0.279
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.249	0.274	-0.275	0.316
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.092	0.950	0.093	0.291
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.233	0.646	-0.092	0.340
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.326	0.740	-0.181	0.262
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.283	0.736	-0.156	0.288
	(<.001)	(<.001)	(<.001)	(<.001)

B. NYSE-listed Stocks

Deviations from stock means	a1	a2	b1	b2
Time-weighted Quoted Spread in Basis Points	-0.179	0.860	-0.163	0.322
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	-0.239	0.666	-0.176	0.369
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.287	0.610	-0.212	0.354
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.169	0.370	-0.175	0.409
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.078	0.949	0.115	0.355
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.306	0.693	-0.145	0.392
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.298	0.771	-0.197	0.313
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.221	0.810	-0.153	0.352
	(<.001)	(<.001)	(<.001)	(<.001)

C. NASDAQ-listed Stocks

Deviations from stock means	a1	a2	b1	b2
Time-weighted Quoted Spread in Basis Points	-0.050	0.931	-0.058	0.328
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	-0.034	0.933	-0.041	0.340
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.343	0.575	-0.216	0.282
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.311	0.236	-0.371	0.297
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.084	0.961	0.076	0.297
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.126	0.652	-0.036	0.351
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.267	0.768	-0.152	0.281
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.241	0.744	-0.139	0.300
	(<.001)	(<.001)	(<.001)	(<.001)

TABLE 8. SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND MARKET QUALITY BY SIZE

The table reports the results of analyzing the relationship between Dark Pool activity and Market Quality. We measure Dark Pool activity as RELDP, which is defined as 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Market quality measures are described in Table 3. Due to the potential simultaneity between market quality and dark pool activity, we estimate the following two-equation simultaneous equation model for RELDP and the following market quality measures (MQMs): Time-Weighted Quoted Spread in Percent and Cents; Share-weighted Effective Spread in Percent and Cents; (log of) Time-weighted Bid-depth; (log) of Share volume; Standard deviation of Mid-quote Returns; and High-Low/High:

$$MQM_{i,t} = a_1 RELDP_{i,t} + a_2 MQMnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 MQM_{i,t} + b_2 RELDPnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (SMALL, MEDIUM, LARGE), and with the same two-digit SIC code (excluding stock i). Similarly, as an instrument for $MQM_{i,t}$ we use $MQMnot_{i,t}$, which is the average Market Quality Measure for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model by pooling observations across all stocks and days in the sample. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the estimated coefficients on top and p-values below . The results for the SMALL caps are reported in Panel A, for MEDIUM caps in Panel B, and for LARGE caps in Panel C. SMALL includes stocks with market capitalization less than $50 million, MEDIUM includes stocks with market capitalization between $50 million and $1 billion, and LARGE includes stocks with market capitalization of $1 billion and more.

A. SMALL

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Time-weighted Quoted Spread in Basis Points	-0.793	0.659	-0.113	0.069
	(<.001)	(<.001)	(<.001	(<.001)
Time-weighted Quoted Spreads in Cents	-0.537	0.481	-0.150	0.074
	(<.001)	(<.001)	(<.001	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.780	0.310	-0.179	0.074
	(<.001)	(<.001)	(<.001	(<.001)
Share-weighted Effective Spreads in Cents	-0.449	0.086	-0.354	0.075
	(<.001)	(<.001)	(<.001	(<.001)
Time-weighted Bid-depth in Shares	0.380	0.904	0.072	0.071
	(<.001)	(<.001)	(<.001	(<.001)
Share volume	1.041	0.274	0.245	0.064
	(<.001)	(<.001)	(<.001	(<.001)
Standard Deviation of Mid-quote Returns	-0.843	0.561	-0.128	0.071
	(<.001)	(<.001)	(<.001	(<.001)
High-Low/High (CRSP)	-0.658	0.329	-0.149	0.085
	(<.001)	(<.001)	(<.001	(<.001)

B. MEDIUM

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Time-weighted Quoted Spread in Basis Points	-0.042	0.918	-0.048	0.294
	(0.001)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	-0.035	0.889	-0.176	0.369
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.228	0.616	-0.163	0.264
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.245	0.265	-0.253	0.246
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.079	0.963	0.072	0.262
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.207	0.615	-0.051	0.305
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.241	0.741	-0.139	0.258
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.198	0.770	-0.112	0.270
	(<.001)	(<.001)	(<.001)	(<.001)

C. LARGE

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Time-weighted Quoted Spread in Basis Points	-0.077	0.933	-0.131	0.445
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	-0.064	0.876	-0.089	0.504
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.333	0.607	-0.273	0.406
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.181	0.407	-0.190	0.506
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.072	0.952	0.136	0.428
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.226	0.776	-0.136	0.494
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.215	0.848	-0.208	0.373
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.188	0.851	-0.169	0.429
	(<.001)	(<.001)	(<.001)	(<.001)

TABLE 9. STOCK-BY-STOCK ESTIMATION OF SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND MARKET QUALITY

The table reports the results of analyzing the relationship between Dark Pool activity and Market Quality. We measure Dark Pool activity as RELDP, which is defined as 100 times SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Market quality measures are described in Table 3. Due to the potential simultaneity between market quality and dark pool activity, we estimate the following stock-by-stock two-equation simultaneous equation model for RELDP and the following market quality measures (MQMs): Time-Weighted Quoted Spread in Percent and Cents; Share-weighted Effective Spread in Percent and Cents; (log of) Time-weighted Bid-depth; (log) of Share volume; Standard deviation of Mid-quote Returns; and High-Low/High:

$$MQM_{i,t} = a_1 RELDP_{i,t} + a_2 MQMnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 MQM_{i,t} + b_2 RELPDnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock *i*). Similarly, as an instrument for $MQM_{i,t}$ we use $MQMnot_{i,t}$, which is the average Market Quality Measure for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model by pooling observations across all stocks and days in the sample. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the median estimated coefficients on top and its p-value below .

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Time-weighted Quoted Spread in Basis Points	-0.014	0.939	-0.065	0.359
	(0.027)	(<.001)	(<.001)	(<.001)
Time-weighted Quoted Spreads in Cents	0.000	0.868	-0.048	0.381
	(0.627)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Basis Points	-0.140	0.556	-0.145	0.367
	(<.001)	(<.001)	(<.001)	(<.001)
Share-weighted Effective Spreads in Cents	-0.106	0.310	-0.114	0.395
	(<.001)	(<.001)	(<.001)	(<.001)
Time-weighted Bid-depth in Shares	0.153	1.025	0.058	0.371
	(<.001)	(<.001)	(<.001)	(<.001)
Share volume	-0.085	0.705	-0.015	0.403
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of Mid-quote Returns	-0.105	0.785	-0.109	0.344
	(<.001)	(<.001)	(<.001)	(<.001)
High-Low/High (CRSP)	-0.092	0.790	-0.087	0.377
	(<.001)	(<.001)	(<.001)	(<.001)

TABLE 10. DESCRIPTIVE STATISTICS PRICE EFFICIENCY

The table reports descriptive statistics based on monthly stock level data for the SIFMA sample. Standard Deviations are based on midquote returns and the Variance Ratio is defined as (σ22(2)/σ12 - 1), where σ22(2) is the adjusted estimator of the variance of 30-minute returns and σ12 is the adjusted estimator of the variance of 15-minute returns. Standard Deviations of midquote returns are multipled by 100.

Monthly Measures	Observations	Mean	StDev
Standard Deviation of 15-minute Returns	49,882	0.952	0.766
Standard Deviation of 30-minute Returns	49,881	1.223	0.937
Variance Ratio	49,878	-0.185	0.111
Absolute Variance Ratio	49,878	0.188	0.107
Autocorrelation of 15-minute Returns	49,882	-0.020	0.145
Absolute Autocorrelation of 15-minute Returns	49,882	0.068	0.130

TABLE 11. DARK POOL ACTIVITY AND PRICE EFFICIENCY

The table reports average dark pool activity for quintile portfolios sorted monthly by price efficiency measure based on stock level data for the SIMFA sample. RELDP is 100 times SIFMA reported monthly average daily Dark Pool single-counted share volume divided by average daily consolidated volume as reported in CRSP. COUNTDP is the monthly average daily number of active Dark Pools per stock. Standard Deviations are based on midquote returns and the Variance Ratio is defined as (σ22(2)/σ12 - 1), where σ22(2) is the adjusted estimator of the variance of 30-minute returns and σ12 is the adjusted estimator of the variance of 15-minute returns. Standard Deviations of midquote returns are multipled by 100. The table reports time-series averages of monthly means based on monthly sorts of stocks into quintile price efficiency portfolios. The t-tests are based on the time series of the difference in the monthly average dark pool activity between the High and Low portfolios.

A. RELDP

Monthly Measures	Low	2	3	4	High	High-Low	t-statistic
Standard Deviation of 15-minute Returns	5.319	5.529	5.074	3.966	2.035	-3.283	-10.89
Standard Deviation of 30-minute Returns	5.268	5.458	5.022	4.040	2.136	-3.133	-10.64
Absolute Variance Ratio	5.099	4.963	4.682	4.157	3.025	-2.074	-9.04
Absolute Autocorrelation of 15-minute Returns	4.576	4.669	4.603	4.453	3.623	-0.953	-3.61

B. COUNTDP

Monthly Measures	Low	2	3	4	High	High-Low	t-statistic
Standard Deviation of 15-minute Returns	7.720	6.797	5.455	3.751	1.787	-5.934	-24.49
Standard Deviation of 30-minute Returns	7.553	6.638	5.410	3.885	2.025	-5.528	-22.55
Absolute Variance Ratio	6.673	6.538	5.802	4.333	2.167	-4.506	-21.68
Absolute Autocorrelation of 15-minute Returns	5.706	5.803	5.535	5.114	3.352	-2.354	-11.15

TABLE 12. SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND PRICE EFFICIENCY BY EXCHANGE

The table reports the results of analyzing the relationship between Dark Pool activity and Price Efficiency at the monthly frequency. We measure Dark Pool activity as RELDP, which is defined as 100 times the monthly average SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Standard Deviations and Autocorrelations are monthly averags based on midquote returns. The monthly Variance Ratio is defined as ($\sigma_{22}(2)/\sigma_{12}$ - 1), where $\sigma_{22}(2)$ is the adjusted estimator of the variance of 30-minute returns and σ_{12} is the adjusted estimator of the variance of 15-minute returns (Lo and MacKinley (1988)). Standard Deviations of midquote returns are multipled by 100. Due to the potential simultaneity between Price Efficiency and dark pool activity, we estimate the following two-equation simultaneous equation model for RELDP and the following price efficiency (PEs): Standard Deviation of 15-minute Returns, Standard Deviation of 20-minute Returns, Variance Ratio, and Autocorrelation:

$$PE_{i,t} = a_1 RELDP_{i,t} + a_2 PEnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 PE_{i,t} + b_2 RELDPnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock *i*). Similarly, as an instrument for $PE_{i,t}$ we use $PEnot_{i,t}$, which is the average Price Efficiency for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model by pooling observations across all stocks and months in the sample. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the estimated coefficients on top and p-values below . The results for the SIFMA sample are reported in Panel A, for Nasdaq stocks in Panel B, and for NYSE stocks in Panel C.

A. SIFMA Sample

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.206	0.843	-0.227	0.451
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.198	0.850	-0.228	0.447
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.156	0.333	0.248	0.618
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.100	0.116	-0.556	0.617
	(<.001)	(<.001)	(<.001)	(<.001)

B. NYSE-listed Stocks

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.186	0.871	-0.292	0.413
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.163	0.887	-0.286	0.415
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.110	0.372	0.165	0.705
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.070	0.154	-0.365	0.706
	(0.001)	(<.001)	(0.007)	(<.001)

C. NASDAQ-listed Stocks

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.114	0.893	-0.160	0.522
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.114	0.894	-0.163	0.517
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.198	0.355	0.300	0.600
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.126	0.082	-0.956	0.576
	(<.001)	(<.001)	(0.002)	(<.001)

TABLE 13. SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND PRICE EFFICIENCY BY SIZE

The table reports the results of analyzing the relationship between Dark Pool activity and Price Efficiency at the monthly frequency. We measure Dark Pool activity as RELDP, which is defined as 100 times the monthly average SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Standard Deviations and Autocorrelations are monthly averags based on midquote returns. The monthly Variance Ratio is defined as ($\sigma_{22}(2)/\sigma_{12}$ - 1), where $\sigma_{22}(2)$ is the adjusted estimator of the variance of 30-minute returns and σ_{12} is the adjusted estimator of the variance of 15-minute returns (Lo and MacKinley (1988)). Standard Deviations of midquote returns are multipled by 100. Due to the potential simultaneity between Price Efficiency and dark pool activity, we estimate the following two-equation simultaneous equation model for RELDP and the following price efficiency (PEs): Standard Deviation of 15-minute Returns, Standard Deviation of 20-minute Returns, Variance Ratio, and Autocorrelation:

$$PE_{i,t} = a_1 RELDP_{i,t} + a_2 PEnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 PE_{i,t} + b_2 RELDPnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock *i*). Similarly, as an instrument for $PE_{i,t}$ we use $PEnot_{i,t}$, which is the average Price Efficiency for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model by pooling observations across all stocks and months in the sample. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the estimated coefficients on top and p-values below . The results for the SMALL caps are reported in Panel A, for MEDIUM caps in Panel B, and for LARGE capsin Panel C.

A. SMALL

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.265	0.676	-0.186	0.365
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.275	0.674	-0.197	0.357
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.241	0.170	0.427	0.379
	(<.001)	(<.001)	(0.002)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.318	-0.010	-3.853	-0.102
	(<.001)	(0.750)	(0.215)	(0.817)

B. MEDIUM

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.161	0.889	-0.180	0.392
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.151	0.895	-0.179	0.390
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.206	0.315	0.306	0.503
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.158	0.081	-0.931	0.468
	(0.001)	(<.001)	(0.002)	(<.001)

C. LARGE

Deviations from stock means	a1	a2	b1	b2
Standard Deviation of 15-minute Returns	-0.136	0.900	-0.326	0.433
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.118	0.913	-0.324	0.465
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.101	0.467	0.142	0.811
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.026	0.227	-0.111	0.830
	(0.116)	(<.001)	(0.238)	(<.001)

TABLE 14. STOCK-BY-STOCK SIMULTANEOUS EQUATION MODEL: DARK POOL ACTIVITY AND PRICE EFFICIENCY

The table reports the results of analyzing the relationship between Dark Pool activity and Price Efficiency at the monthly frequency. We measure Dark Pool activity as RELDP, which is defined as 100 times the monthly average SIFMA reported daily Dark Pool single-counted share volume divided by daily consolidated volume as reported in CRSP. Standard Deviations and Autocorrelations are monthly averags based on midquote returns. The monthly Variance Ratio is defined as ($\sigma_{22}(2)/\sigma_{12}$ - 1), where $\sigma_{22}(2)$ is the adjusted estimator of the variance of 30-minute returns and σ_{12} is the adjusted estimator of the variance of 15-minute returns (Lo and MacKinley (1988)). Standard Deviations of midquote returns are multipled by 100. Due to the potential simultaneity between Price Efficiency and dark pool activity, we estimate the following two-equation simultaneous equation model for RELDP and the following price efficiency (PEs): Standard Deviation of 15-minute Returns, Standard Deviation of 20-minute Returns, Variance Ratio, and Autocorrelation:

$$PE_{i,t} = a_1 RELDP_{i,t} + a_2 PEnot_{i,t} + e_{1,t}$$
$$RELDP_{i,t} = b_1 PE_{i,t} + b_2 RELDPnot_{i,t} + e_{2,t}$$

As an instrument for $RELDP_{i,t}$ we *use* $RELDPnot_{i,t}$, which is the average Dark Pool activity of other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock *i*). Similarly, as an instrument for $PE_{i,t}$ we use $PEnot_{i,t}$, which is the average Price Efficiency for other stocks listed on the same exchange, in the same market capitalization grouping (LARGE, MEDIUM, SMALL), and with the same two-digit SIC code (excluding stock i). We estimate the simultaneous equation model stock-by-stock. To make the pooling meaningful, we de-mean all variables by deducting the stock-specific average and scale all variables by dividing by the stock-specific standard deviation to control for stock fixed effects. We report the median estimated coefficients on top and associated p-values below .

Deviations from stock means	**a1**	**a2**	**b1**	**b2**
Standard Deviation of 15-minute Returns	-0.050	0.914	-0.151	0.477
	(<.001)	(<.001)	(<.001)	(<.001)
Standard Deviation of 30-minute Returns	-0.048	0.921	-0.147	0.474
	(<.001)	(<.001)	(<.001)	(<.001)
Absolute Variance Ratio	0.103	0.380	0.163	0.667
	(<.001)	(<.001)	(0.030)	(<.001)
Absolute Autocorrelation of 15-minute Returns	-0.003	0.282	0.050	0.579
	(0.956)	(<.001)	(0.450)	(<.001)

TABLE APPENDIX TABLE 2. UNIVARIATE DARK POOL ACTIVITY BY SAMPLE

The table reports univariate statistics based on daily stock level data for: the SIMFA sample (Panel A); the Weaver (2010) sample (Panel B); and the O'Hara and Ye (2009) sample. (Panel C) DPVOL is SIFMA reported daily Dark Pool single-counted share volume per stock in thousands. RELDP is 100 times DPVOL divided by daily consolidated volume as reported in CRSP. COUNTDP is the number of active Dark Pools per day per stock. IHERF is the inverse of the Herfindahl Index for Dark Pools.

A. SIFMA SAMPLE	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	85.08	794.03	0.21	5.55	37.90
RELDP	4.51	5.74	0.65	3.05	6.22
COUNTDP	5.27	3.97	1	5	9
IHERF	2.43	1.74	1.00	2.40	3.70
B. WEAVER SAMPLE	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	85.14	794.32	0.22	5.58	37.96
RELDP	4.52	5.74	0.65	3.05	6.22
COUNTDP	5.27	3.97	1	5	9
IHERF	2.43	1.74	1	2.4	3.7
C. O'HARA & YE SAMPLE	**Average**	**StDev**	**Q1**	**Median**	**Q3**
DPVOL	102.37	363.93	2.20	16.30	73.96
RELDP	5.27	5.30	1.95	4.11	7.01
COUNTDP	6.88	3.66	4	8	10
IHERF	3.03	1.65	1.93	3.07	4.19

APPENDIX TABLE 3. DESCRIPTIVE STATISTICS

The table reports descriptive statistics based on daily stock level data for the two benchmark samples: the sample matching Weaver (2010); and the sample matching O'Hara and Ye (2009). Market capitalization is in billion dollars, share volume is in million shares, bid depth, absolute depth imbalance and order imbalance are in thousand shares. Relative order imbalance is in percent of share volume. Percent spreads are in basis points. (High-Low)/High measures and returns are multiplied by 100. Standard Deviation of midquote returns are multipled by 10,000.

A. Weaver (2011) Sample

Daily Measures	Observations	Mean	StDev
Market Capitalization (CRSP)	1,011,004	2.55	12.34
Share Volume (CRSP)	1,011,004	1.59	13.26
Price (CRSP)	1,011,004	15.38	23.05
(High-Low)/High (CRSP, quotes)	1,011,004	5.88	5.21
Absolute Return (CRSP)	1,010,887	3.44	5.50
Quoted Spread Basis Points (TAQ)	1,009,986	175.08	352.39
Quoted Spread Cents (TAQ)	1,009,986	9.63	27.24
Effective Spread Basis Points (TAQ)	1,000,710	41.28	128.00
Effective Spread Cents (TAQ)	1,000,710	1.88	6.66
Bid Depth (TAQ)	1,009,986	124.11	751.08
Relative Order Imbalance in Percent (TAQ)	1,000,710	20.43	23.92
Standard Deviation of Midquote Returns (TAQ)	1,009,634	26.60	60.04

B. O'Hara and Ye (2011) Sample

Daily Measures	Observations	Mean	StDev
Market Capitalization (CRSP)	510,433	1.83	6.42
Share Volume (CRSP)	510,433	1.83	6.42
Price (CRSP)	510,433	74.67	2090.26
(High-Low)/High (CRSP, quotes)	510,433	4.12	2.71
Absolute Return (CRSP)	510,356	2.34	2.59
Quoted Spread Basis Points (TAQ)	510,115	62.91	169.51
Quoted Spread Cents (TAQ)	510,115	18.08	346.71
Effective Spread Basis Points (TAQ)	507,272	11.64	45.02
Effective Spread Cents (TAQ)	507,272	3.48	151.80
Bid Depth (TAQ)	510,115	119.99	82.77
Relative Order Imbalance in Percent (TAQ)	507,272	13.48	18.20
Standard Deviation of Midquote Returns (TAQ)	510,053	8.98	22.76

APPENDIX TABLE 4. DARK POOL ACTIVITY AND MARKET QUALITY IN THE CROSS-SECTION

This table replicates Weaver (2010) based on the SIFMA sample in Panel A and the O'Hara and Ye (2009) sample in Panel B. Average market quality statistics and Dark Pool activity across days for each firm is first estimated. We then run cross-sectional regressions with market quality measures on the left hand side and stock characteristics and Dark Pool activity on the right hand side. Market quality statistics include quoted and effective spreads in percent and cents, the Amihud measure calculated as the average daily absolute CRSP return divided by CRSP volume (|ret|/Vol)., and the standard deviation of midquote returns. Volume is average CRSP volume in million shares and Price is CRSP closing price in dollars. Dark Pool activity is measured as the average daily Dark Pool share volume in percent of CRSP volume. We allow for non-linear effects by estimating a third order polynomial in Dark Pool activity (RELDP). The columns report estimated cross-sectional regression coefficients on top and t-statistics below.

A. SIFMA Sample	**Intercept**	**Volume**	**Price**	**StDev MQR**	**No. Trades**	**RELDP**	**RELDP^2**	**RELDP^3**	**Adjusted R2**
Quoted Spread in Percent (TAQ)	1.644	-0.003		0.052		-0.545	0.056	-0.002	0.862
	17.83	-1.29		100.00		-10.63	6.10	-3.66	
Quoted Spread in Cents (TAQ)	28.683	-0.073	0.140	0.012		-8.638	0.809	-0.021	0.985
	16.43	-1.99	531.55	1.23		-8.91	4.68	-2.32	
Effective Spread in Percent (TAQ)	-0.016	0.000		0.016		0.012	-0.002	0.000	0.841
	-0.60	0.65		104.84		0.78	-0.75	0.50	
Effective Spread in Cents (TAQ)	5.344	-0.012	0.027	0.001		-1.597	0.161	-0.005	0.985
	15.77	-1.65	531.72	0.75		-8.49	4.80	-2.73	
Amihud Measure (CRSP)	0.036					-0.018	0.003	0.000	0.022
	11.15					-7.60	5.30	-4.31	
StDev Mid-quote Returns (TAQ)	48.990				-6.028	-44.987	6.259	-0.255	0.578
	89.4				-20.27	-35.69	26.95	-20.52	

B. O'Hara and Ye (2011) Sample	**Intercept**	**Volume**	**Price**	**StDev MQR**	**No. Trades**	**RELDP**	**RELDP^2**	**RELDP^3**	**Adjusted R2**
Quoted Spread in Percent (TAQ)	0.006	-0.001		0.076		0.017	-0.008	0.000	0.911
	0.10	-0.67		112.62		0.59	-1.73	1.96	
Quoted Spread in Cents (TAQ)	19.580	-0.169	0.140	0.779		-6.445	0.544	-0.012	0.991
	5.60	-1.59	496.11	19.76		-3.84	2.02	-0.98	
Effective Spread in Percent (TAQ)	0.064	0.000		0.012		-0.015	0.001	0.000	0.891
	6.03	-0.88		98.58		-2.85	1.23	-0.59	
Effective Spread in Cents (TAQ)	5.208	-0.025	0.027	0.108		-1.801	0.190	-0.006	0.989
	6.68	-1.05	430.44	12.26		-4.80	3.16	-2.14	
Amihud Measure (CRSP)	0.024					-0.010	0.001	0.000	0.189
	23.18					-16.48	11.83	-8.94	
StDev Mid-quote Returns (TAQ)	62.364				-3.717	-10.897	1.289	-0.046	0.532
	49.42				-24.23	-13.52	10.00	-7.50	

APPENDIX TABLE 5. MARKET QUALITY FOR MATCHED SAMPLES WITH HIGH AND LOW DARK POOL VOLUME

In this table we follow report the test of differences in mean and median daily market quality statistics between stocks with High and Low amounts of Dark Pool activity (RELDP). The sample is constructed by first selecting 10 stocks randomly from each size decile on NASDAQ and NYSE separately based on market capitalization on January 2, 2009. We then select matching securities as the stocks that minimize the sum of the absolute percent difference in market captialization and price, holding the exchange of primary listing constant. Finally, for each pair of stocks we classify the High and Low RELDP stock based on their Dark Pool activity during January - June, 2009. Market quality statistics are computed based on July - December, 2009 data. P-values for paired t-tests of means and Wilcoxon signed rank test for differences in median are reported in the last column.

Market Quality Measure	Test	High RELDP	Low RELDP	High-Low RELDP	p-value
Quoted Spread Percent (TAQ)	Mean	41.49	45.00	-3.51	0.00
	Median	11.16	11.86	-0.42	0.00
Quoted Spread Cents (TAQ)	Mean	7.71	8.93	-1.22	0.00
	Median	2.76	2.74	0.00	0.00
Effective Spread Percent (TAQ)	Mean	7.88	8.23	-0.53	0.01
	Median	3.39	3.72	-0.19	0.00
Effective Spread Cents (TAQ)	Mean	1.52	1.65	-0.15	0.00
	Median	0.80	0.80	0.00	0.00
Bid Depth (TAQ)	Mean	146.51	129.85	16.65	0.00
	Median	148.29	132.10	9.42	0.00
Relative Order Imbalance in Percent (TAQ)	Mean	11.54	11.82	-0.35	0.00
	Median	7.50	7.38	0.06	0.75
(High-Low)/High (CRSP, quotes)	Mean	3.32	3.31	0.01	0.58
	Median	2.89	2.90	0.00	0.89
StDev Midquote Returns (TAQ)	Mean	5.79	6.42	-0.64	0.00
	Median	2.56	2.68	-0.11	0.00
Absolute Return (CRSP)	Mean	1.74	1.77	-0.02	0.09
	Median	1.26	1.29	-0.03	0.04

APPENDIX TABLE 6. MARKET QUALITY FOR MATCHED SAMPLES WITH HIGH AND LOW DARK POOL VOLUME BY MARKET CAPITALIZATION

In this table we follow report the test of differences in mean and median daily market quality statistics between stocks with High and Low amounts of Dark Pool activity (RELDP) for the top 3 declies (Large Caps)a nd the lowest 3 declies (Small Caps). The sample is constructed by first selecting 10 stocks randomly from each size decile on NASDAQ and NYSE separately based on market capitalization on January 2, 2009. We then select matching securities as the stocks that minimize the sum of the absolute percent difference in market captialization and price, holding the exchange of primary listing constant. Finally, for each pair of stocks we classify the High and Low RELDP stock based on their Dark Pool activity during January - June, 2009. Market quality statistics are computed based on July - December, 2009 data. P-values for paired t-tests of means and Wilcoxon signed rank test for differences in median are reported in the last column.

		Large Caps		**Small Caps**	
Market Quality Measure	Test	**High-Low RELDP**	**p-value**	**High-Low RELDP3**	**p-value4**
Quoted Spread Percent (TAQ)	Mean	-1.32	0.00	-7.60	0.00
	Median	-0.18	0.00	-4.23	0.00
Quoted Spread Cents (TAQ)	Mean	0.04	0.72	-3.64	0.00
	Median	0.00	0.03	-0.45	0.00
Effective Spread Percent (TAQ)	Mean	-0.17	0.00	-0.71	0.28
	Median	-0.05	0.00	-0.80	0.00
Effective Spread Cents (TAQ)	Mean	0.02	0.11	-0.29	0.00
	Median	0.01	0.00	-0.06	0.00
Bid Depth (TAQ)	Mean	27.10	0.00	21.77	0.00
	Median	23.83	0.00	19.95	0.00
Relative Order Imbalance in Percent (TAQ)	Mean	-0.04	0.69	-1.31	0.01
	Median	0.09	0.39	0.00	0.08
(High-Low)/High (CRSP, quotes)	Mean	0.09	0.00	0.13	0.00
	Median	0.08	0.00	0.04	0.02
StDev Midquote Returns (TAQ)	Mean	-0.15	0.00	-1.34	0.00
	Median	-0.02	0.00	-0.52	0.00
Absolute Return (CRSP)	Mean	0.05	0.02	-0.01	0.75
	Median	0.00	0.24	-0.01	0.40

APPENDIX TABLE 7. DESCRIPTIVE STATISTICS PRICE EFFICIENCY

The table reports descriptive statistics based on daily stock level data for the two benchmark samples: the sample matching Weaver (2010); and the sample matching O'Hara and Ye (2009). Standard Deviations are based on midquote returns and the Variance Ratio is defined as (σ22(2)/σ12 - 1), where σ22(2) is the adjusted estimator of the variance of 30-minute returns and σ12 is the adjusted estimator of the variance of 15-minute returns. Standard Deviations of midquote returns are multipled by 100.

A. Weaver (2011) Sample

Monthly Measures	Observations	Mean	StDev
Standard Deviation of 15-minute Returns	49,846	0.952	0.766
Standard Deviation of 30-minute Returns	49,845	1.224	0.937
Variance Ratio	49,842	-0.185	0.111
Autocorrelation of 15-minute Returns	49,846	-0.020	0.145

B. O'Hara and Ye (2011) Sample

Monthly Measures	Observations	Mean	StDev
Standard Deviation of 15-minute Returns	24,761	0.593	0.293
Standard Deviation of 30-minute Returns	24,761	0.773	0.376
Variance Ratio	24,761	-0.170	0.100
Autocorrelation of 15-minute Returns	24,761	-0.012	0.079

APPENDIX TABLE 8. DARK POOL ACTIVITY AND PRICE EFFICIENCY IN THE CROSS-SECTION

This table replicates Weaver (2010) based on the three sub-samples: SIFMA sample and the O'Hara and Ye (2009) sample. Price efficiency measures are calculated based on the entire year of data. We then run cross-sectional regressions with price efficiency measures on the left hand side and stock characteristics and Dark Pool activity on the right hand side. Price efficiency measures statistics are described in Table 12. No. Trades is the daily average log of the number of trades divided according to CRSP. Dark Pool activity is measured as the average daily Dark Pool share volume in percent of CRSP volume. We allow for non-linear effects by estimating a third order polynomial in Dark Pool activity (RELDP). The columns report estimated cross-sectional regression coefficients multiplied by 1,000 on top and t-statistics below.

A. SIFMA Sample	**Intercept**	**No. Trades**	**RELDP**	**RELDP^2**	**RELDP(95)**	**Adjusted R2**
Standard Deviation of 15-minute Returns	21.170	-0.279	-4.120	0.447	-0.015	0.366
	80.85	-5.97	-20.82	12.25	-7.58	
Standard Deviation of 30-minute Returns	26.320	-0.249	-5.200	0.566	-0.019	0.336
	77.84	-4.13	-20.31	12.00	-7.51	
Variance Ratio	-298.170	16.110	3.640	0.084	-0.039	0.327
	-97.46	29.51	1.57	0.20	-1.71	
Autocorrelation of 15-minute Returns	-92.560	7.350	7.890	-0.009	-0.035	0.258
	-38.09	16.94	4.29	-0.06	-1.91	
B. O'Hara and Ye (2011) Sample	**Intercept**	**No. Trades**	**RELDP**	**RELDP^2**	**RELDP^3**	**Adjusted R2**
Standard Deviation of 15-minute Returns	8.730	-0.382	0.288	-0.064	0.003	0.153
	40.05	-14.42	2.07	-2.85	3.25	
Standard Deviation of 30-minute Returns	10.510	-0.422	0.510	-0.096	0.005	0.110
	37.84	-12.51	2.88	-3.38	3.49	
Variance Ratio	-322.190	13.500	23.450	-2.820	0.078	0.304
	-53.3	18.37	6.08	-4.57	2.69	
Autocorrelation of 15-minute Returns	-87.350	7.14	7.100	-0.302	-0.013	0.192
	-19.25	12.95	2.45	-0.65	-0.58	

APPENDIX TABLE 9. PRICE EFFICIENCY FOR SAMPLES WITH HIGH AND LOW DARK POOL VOLUME

In this table we follow report the test of differences in mean and median monthly price efficiency statistics between stocks with High and Low amounts of Dark Pool activity (RELDP) in Panel A, and for the top 3 declies (Large Caps) and the lowest 3 declies (Small Caps) in Panel B. The sample is constructed by first selecting 10 stocks randomly from each size decile on NASDAQ and NYSE separately based on market capitalization on January 2, 2009. We then select matching securities as the stocks that minimize the sum of the absolute percent difference in market captialization and price, holding the exchange of primary listing constant. Finally, for each pair of stocks we classify the High and Low RELDP stock based on their Dark Pool activity during January - June, 2009. Price efficiency statistics are computed based on July - December, 2009 data and standard deviations are multipled by 1,000. P-values for paired t-tests of means and Wilcoxon signed rank test for differences in median are reported in the last column.

A. Matched Sample

Price Efficiency Measure	Test	High RELDP	Low RELDP	High-Low RELDP	p-value
Standard Deviation of 15-minute Returns	Mean	7.149	7.692	-0.544	0.00
	Median	5.985	6.541	-0.230	0.00
Standard Deviation of 30-minute Returns	Mean	9.255	9.813	-0.557	0.01
	Median	7.898	8.308	-0.210	0.00
Variance Ratio	Mean	-0.172	-0.189	0.016	0.01
	Median	-0.159	-0.180	0.013	0.00
Autocorrelation of 15-minute Returns	Mean	-0.014	-0.024	0.010	0.01
	Median	-0.010	-0.018	0.007	0.02

B. Matched Sample by Capitalization

		Large Caps		Small Caps	
Price Efficiency Measure	Test	**High-Low RELDP**	**p-value**	**High-Low RELDP**	**p-value**
Standard Deviation of 15-minute Returns	Mean	-0.420	0.01	-0.980	0.00
	Median	0.060	0.10	-0.260	0.05
Standard Deviation of 30-minute Returns	Mean	-0.470	0.03	-1.070	0.01
	Median	0.070	0.13	-0.120	0.12
Variance Ratio	Mean	0.012	0.35	0.022	0.01
	Median	-0.003	0.71	0.021	0.01
Autocorrelation of 15-minute Returns	Mean	0.010	0.34	0.016	0.01
	Median	0.008	0.39	0.013	0.03

FIGURE 1. Dark Pool Share Volume

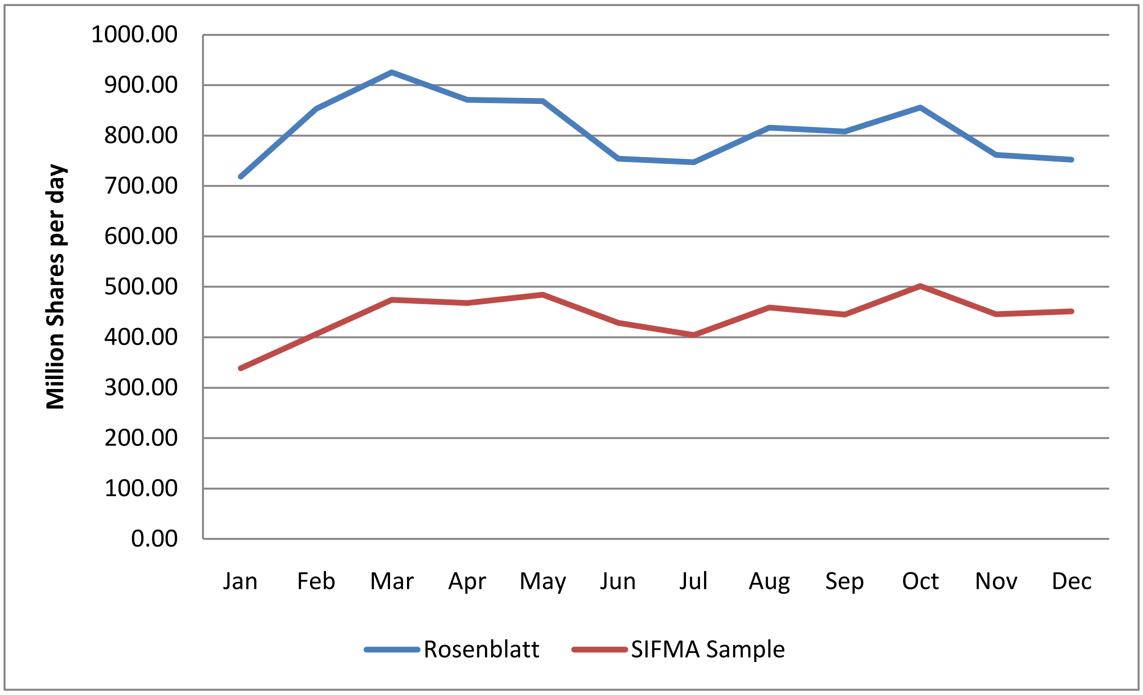

FIGURE 2. Dark Pool Relative to Consolidated Volume

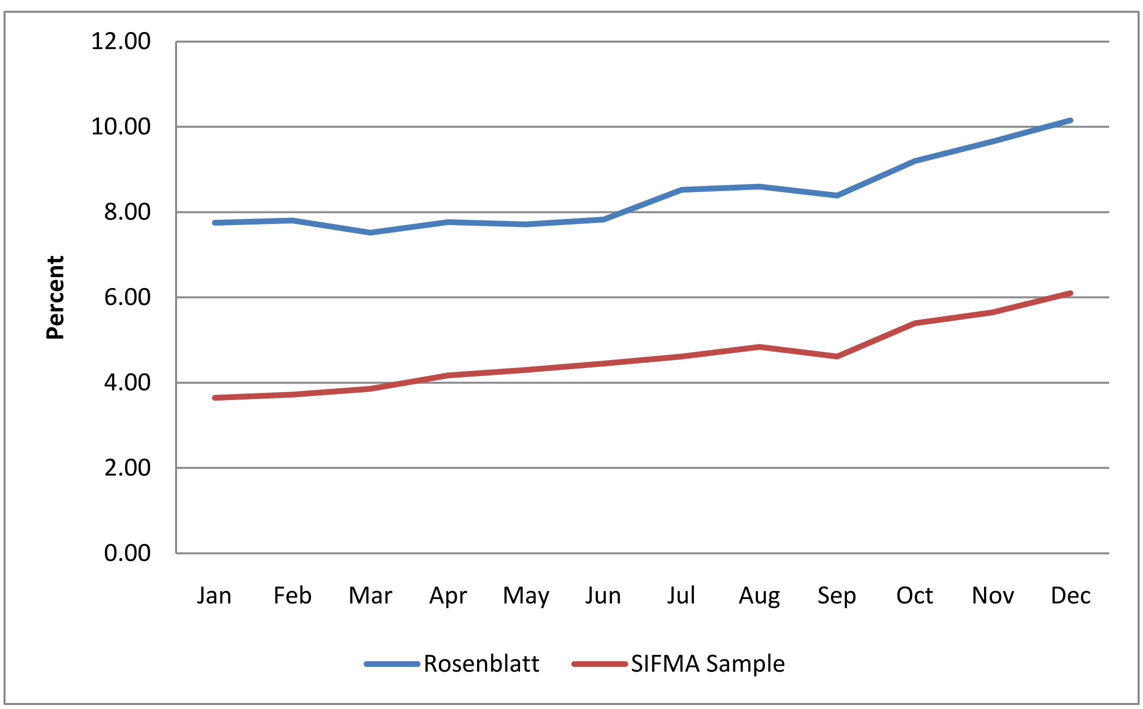

FIGURE 3. SIFMA Dark Pool Activity Relative to Rosenblatt

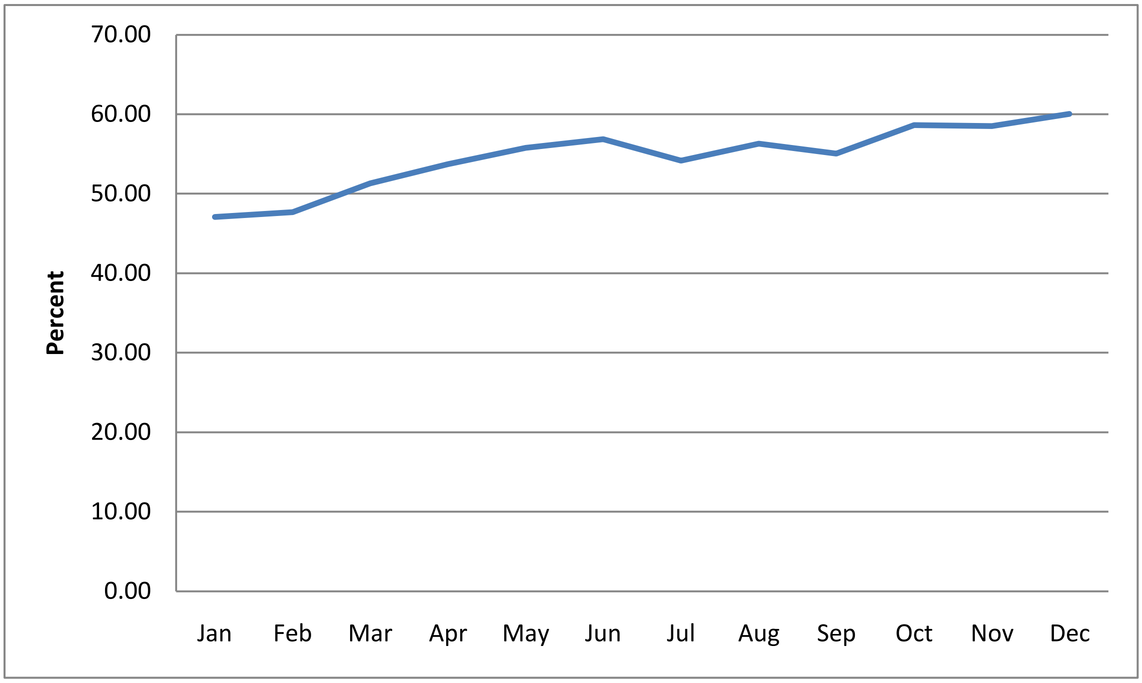

FIGURE 4. Percent Quoted Spread and RELDP – Functional Form

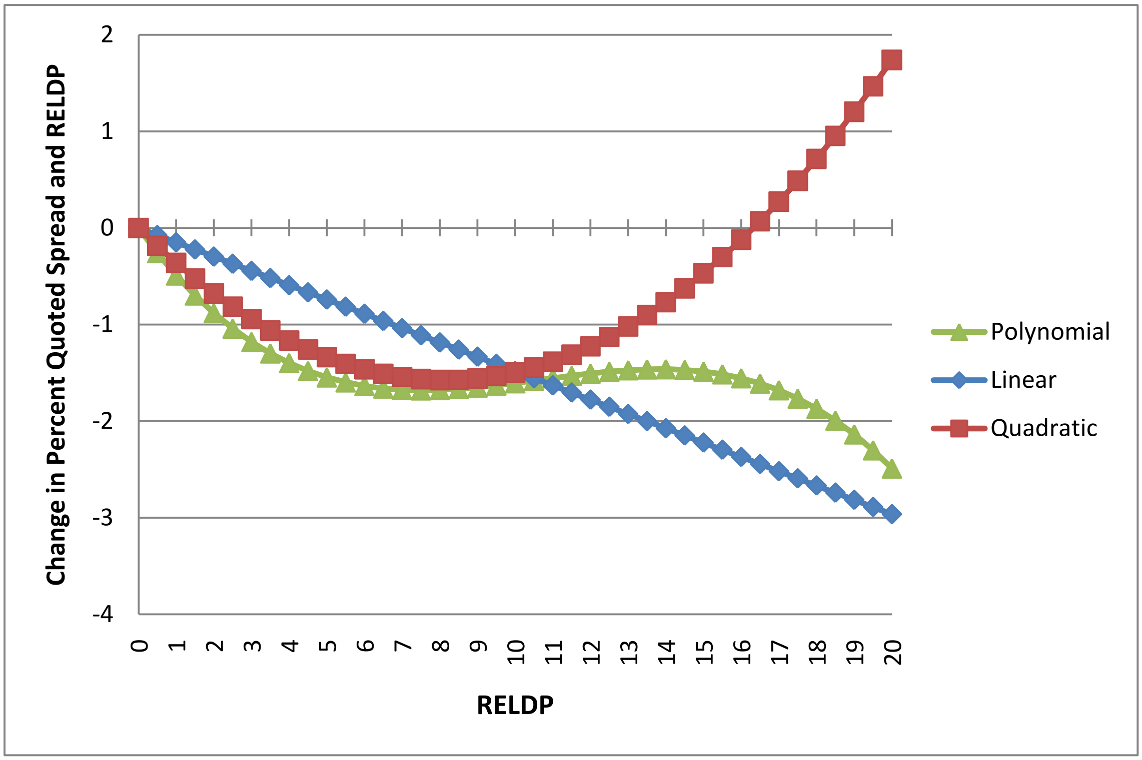

FIGURE 5. Histogram of RELDP