Making Dark Pools of Liquidity Visible

Sacrificing Price Discovery for the Sake of Competition in the Market Microstructure Environment

ERASMUS UNIVERSITEIT ROTTERDAM

Erasmus University of Rotterdam | **Master Thesis**
Erasmus School of Economics | **10-02-2011**
Capacity Group Business Economics | **Author: Yussel de Kraa**
Section Finance | **Student number 287572**
Supervisor: Dr. R.C.J. Zwinkels | **ydekraa@gmail.com**

Making Dark Pools of Liquidity Visible: *Sacrificing Price Discovery for the Sake of Competition in the Market Microstructure Environment**

Yussel de Kraa, 2011

Abstract

In this thesis we provide an overview of the field of research of Market Microstructure and describe the characteristics of dark pools of liquidity. Following Hasbrouck (1995) we define the US primary stock markets for a subset of stocks as two series: one containing all quotes on the NYSE, and the other as all other public stock venues. We define cointegration, covariance and correlation as indicators of the quality of the price discovery process in primary markets and accordingly test whether dark pools hampered the quality of this process in US stock markets over the period 2005 - 2010. We research cointegration by applying a Johansen cointegration test (1991), estimate a Vector Error Regression model to look at where price discovery in the markets happens and how it is influenced, and estimate a Bivariate GARCH (1,1) model to deduct covariance and correlation. We find the NYSE is leading in the price discovery process, covariance and cointegration between the NYSE and non NYSE series is not constant overtime, and an indicator for dark pool market share in consolidated US equity markets is significantly negatively related to correlation of the series. The latter implies dark pools indeed negatively influence the quality of the price discovery process.

* I thank Dr. R.C.J. Zwinkels of the Erasmus School of Economics for his comments and supervisory. Furthermore, I am grateful for the data and literature provided by J. Schack of Rosenblatt Securities Inc. Any errors are my own.

NON-PLAGIARISM STATEMENT

By submitting this thesis the author declares to have written this thesis completely by himself/herself, and not to have used sources or resources other than the ones mentioned. All sources used, quotes and citations that were literally taken from publications, or that were in close accordance with the meaning of those publications, are indicated as such.

Index

1. Introduction

Innovation and political lobbying for more liberalized financial markets have resulted in a competitive and quickly changing stock exchange landscape over the last 10 years. The media are increasingly reporting on so called ‘dark pools (of liquidity)’. In news flashes they mention these alternative stock trading venues ‘gain market share from incumbent stock exchanges’ but often remain vague what exactly these exchanges are and how they operate. Though originally designed for institutional investors seeking an opportunity to protect their block trades from adverse price movements, current clientele also consists of high frequency traders, hedge funds, and other professional investors. By 2008[1] estimations are about 40 (Mittal, 2008 and Degryse et al, 2008) or 50 (Tabb 2008) dark pools are in operation, by 2010 representing some 12% (Schack, 2010) of total consolidated trading volume in the United States and some 3% (Schack, 2010) in Europe. Dark pools can no longer be considered some exotic financial innovation and might well be an important part of stock trading in the future. As the trading community has been arguing, it is of vital importance to understand their impact on markets and security prices so regulators will be able to design proper regulation.

Many authors have raised questions about the impact dark pools might have on incumbent exchanges, price discovery, liquidity, and other aspects of market microstructure. They generally agree on the benefits dark pools provide to trading, such as lower transaction costs (Ende and Muntermann, 2010) and protection of market impact costs (Bikker et al, 2007 and Conrad et all, 2003). On the drawbacks literature is less consistent though. Assuming every quote in the market contains some information relevant for the price discovery process (Schwartz, 2010) liquidity disappearing to dark trading venues actually implies less information remains in the open market. Dark pools derive their prices from public markets, but information about prices within the dark pool remains opaque. Traders in the open market, often retail investors, do not have access to all demand and supply whereas an institutional investor in a dark pool has. As such an information asymmetry exists and the price discovery process in the primary markets might be influenced (Buti et al, 2010 and Giraud, 2009 and SEC 2010). Possible results might be large arbitrage possibilities and a loss of investor confidence. By steering markets towards more competition to lower transaction fees, regulators risk to have less liquid and less information efficient prices (Petrella, 2009 p.1). As long as limited macro-market research can be performed on dark pools it will remain difficult to adjust regulation. Consequently small investors become victims in the search for cheaper trading.

Exactly the question whether or not liquidity migrating to dark pools influences price discovery will be the main research question of this thesis. Whereas single dark pools occasionally provide data (often because of promotional reasons), macro data on this industry as a whole is basically not available (Sofianos, 2007 p.7). Therefore we research the other, visible markets. We use the distinction and relation between quotes on the NYSE on the one hand, and all other US public exchanges on the other to research price discovery in the US (following Hasbrouck, 1995) measured by correlation. We expect this correlation changed over the period January 2005 - June 2010 and find it indeed has not been constant over time. Though not analyzed as profound as Hasbrouck (1995) we also find indications of the NYSE having a dominant role in price adjustments over the non NYSEs.

[1] Presumably, the number of dark pools stayed about constant from 2008 till 2010 as on the one side new venues were started, but on the other M&A activity reduced the number of dark venues. Volume traded in dark pools increased a lot though (See Section 3.3).

Furthermore we test whether a small dataset of monthly estimations of dark pool trading volume in the US shows a pattern related to the change in correlation. Indeed, we find significant indications of this relation, implying the market share dark pools have is negatively related to price discovery.

The outline of this thesis is as follows. Following Madhavan (2000) in the second section we give an overview of the fairly new field of research called market microstructure to which the dark pool topic belongs. Then, though various authors explained the characteristics, definitions and (possible) effects of the rise of dark pools in both the US as well as in Europe, we experienced a lack of a complete description making clear distinctions between many confusing concepts and definitions. Therefore, in the third and fourth sections we will elaborate on both the working of dark pools as well as their impact and regulation within the stock exchange landscape in the United States and Europe separately. Next, in chapter four and five, by applying a Vector Error Correction, a Bivariate GARCH model and a combination of these we will test a number of hypothesis on 4 Dow Jones shares regarding price discovery, volatility and dark pool development in the markets. Finally, Section 7 contains a summary and the conclusions.

2. Market Microstructure: An Overview

2.1. Introduction

In order to transfer any asset men invented the concept of trading long ago. When voluntarily relinquishing something one wishes to get something back. A trade takes place. When the concept of a mutually accepted currency is added the two players are no longer equal, but become a seller and a buyer who both grant a value, represented by a price, to an asset[2]. When trading becomes more frequent or when assets become more difficult to physically trade a third player might be added to the process, facilitating the traders by creating a trading platform. This first Section will present an areal overview of the research field of market microstructure and the main concepts and definitions involved.

2.2. What is market microstructure?

The research field of market microstructure studies the process how asset prices are established as a result of investors' (latent) demands (Madhavan 2000 p.1). Other matters involved include the structure and transparency of markets and the role of information. Until a few years ago market microstructure mainly studied competition within markets whereas now competition between financial markets became important as well (Degryse, 2008 p.1). A central issue concerns the difference between the price of an asset and the value it has. Value represents the importance different investors grant to an asset, also called their latent demand. Price is the value that best reflects the entire market´s desire to hold an asset (Hooper and Schwartz, 2008 p.2). Trying to understand why asset prices change we could look at Brandt and Kavajecz (2004 p.2623) examining yield curve changes. They suggest two mechanisms should be distinguished. Firstly, public information flow, as for instance (periodically scheduled) macro-economic announcements, cause immediate changes in asset prices. Secondly, they mention heterogeneous private information, or individual interpretation of the public information mentioned before, and label this the concept of price discovery. As such one could distinguish a hedge fund employing a former board member of the Federal Reserve, which makes them differently informed than a pension fund rebalancing its portfolio containing over 5,000 stocks. Because of frictions such as information asymmetries of different investors, price is not a perfect resultant of value. Put differently, price might be the resultant of value, but the value for an investor is unsure as his information is never complete. As a third mechanism, the mutual behavior of investors could be distinguished. This implies an investor might react to other investor's behavior as a result of the latter's heterogeneous interpretation. Not (only) the real information underlying order flow influences prices, but more often it is the human interactions to trading events that influence order flow (Hooper and Schwartz, 2008 p.2-3). For instance Easly et al (1997) find trading behavior of uninformed traders is highly history dependent as i) they are more likely to trade when trades have recently occurred, ii) they are more likely to buy (sell) when the last trade was a buy (sell). The difference between price and value affects financing and capital structure decision making and explains an investor is willing to pay a certain price as he expects the value of the asset for him to be higher.

[2] As dark pools are currently mainly allowed to trade shares, this will be the default in this thesis.

2.3. Buy versus sell side traders

Buy side traders acquire securities from other buy side traders in the secondary market[3], whereas sell side firms provide and introduce many products and services in both the primary (for instance a public offering) and secondary market. Together these two markets are also indicated as the downstairs market. Next to retail investors the majority of buy side trading volume is done by large institutional investors like all kinds of mutual funds including pension funds.

2.4. Liquidity

Liquidity is often explained as the easiness at which you can buy a certain share. While this is true a specific addition should be made especially in the context of dark pools. Liquidity is moreover the degree to which an asset can be traded without affecting the price significantly. Note how this addition emphasizes how price depends on liquidity. Where the main indication of high liquidity in a certain asset used to be high trading volume, it recently became more the depth of the order book.

2.5. Market makers

A market offering investors continuous trading needs to offer liquidity at any moment the market is open. Market makers can be invited to ensure liquidity by continuously offering quotes for which they buy and sell assets (for instance NASDAQ). Their reward is the bid-ask spread (see Section 2.7). On other exchanges traders traditionally already had to provide liquidity themselves. As such individual investors can also place limit orders to supply liquidity. With the increasing use of computer power in trading, automated limit orders created a nearly continuous fluctuating liquidity. As such the need for market makers to solely guarantee liquidity does not exist anymore.

2.6. Informed trading

Bagehot (1971)[4] introduced the concept of informed traders versus liquidity traders. Investors trading based on heterogeneous private information are considered informed traders[5] and therefore add new information to the market and contribute to the price discovery. Liquidity (also called uninformed or noise) traders do not believe to possess information of which they could profit and their trading is driven by exogenous liquidity needs (Takayama, 2009 p.3) such as rebalancing of a portfolio of stocks. As such, a pension fund just rebalancing its portfolio can not be considered an informed trader whereas an individual retail investor that deducts a conclusion from studying some company press statement can. As the uncertainty of the information[6] decreases, informed traders generally expect to increasingly profit from trading with liquidity traders. The amount of informed trading can be measured by the probability of informed trading[7] (PIN) and is used as a proxy for the

---

[3] Distinguish primary market, secondary market, third market (see Section 3.3) and the fourth market (see Section 3.3). The third and fourth markets are also indicated as upstairs market.

[4] Pseudonym used by Jack Treynor.

[5] Informed trading only refers to the role of legal information underlying an investors´ decision and does not imply insider trading which is in fact illegal.

[6] Barber et al (2006 p.3) indicate an informed trader faces a variety of risks as i) information may be incorrect, ii) when obtaining to take an opposite position in an asset buying a substitute, the mutual correlation might change, iii) investor sentiment could change, iv) market liquidity might not be available when the informed trader wants to unwind his position in the future.

[7] This probability can be measured by the PIN method (Easly, Kiefer, O'Hara and Paperman, 1996) or the PROBINF method by Copeland and Galai (1983) and Popescu (2007). The latter shows more accurate results.

level of information asymmetry for stocks cross listed at more than one exchange (Easly et al 1996). The PIN as such measures the amount of trading adding information to the market which is important to see how much is added to price discovery.

2.7. Bid-Ask spread

The bid-ask spread is the difference between the current price asked by sellers of an asset and the current price a buyer offers in the market. The spread is traditionally the reward market makers require to provide liquidity to the market on a continuous basis. If an investor wishes to buy the asset he pays the spread to the seller, the liquidity provider. The liquidity demander and provider can both place their order as either market order[8], limit order[9] or any hybrid between the two. The total of outstanding limit orders is called the limit order book. A pair of a bid and an ask price offered on a specific moment for a specific security forms a quote. The bid-ask spread is the reward for both brokerage fees as well as a market maker's remuneration for creating the opportunity to make a transaction without delay. Thus investors pay a premium to market makers for offering the possibility of continuous over periodic trading (Madhavan 2000). Empirical evidence is contradictory though on what system investors ultimately prefer (Kalay et al 2002 p.524).

2.8. Order flow

The definition of net order flow is a widely discussed topic[10]. As suggested by Brandt and Kavajecz (2004, p.2628) we will here consider total order flow as the sum of the absolute value of both signed buy and sell trades. Consequently, net order flow is the sum of signed trade volume where buy (sell) orders have a positive (negative) sign. Assuming investors need to fill their order within a limited time-frame, positive (negative) net order flow will indicate an increase (decrease) in prices due to excess demand (supply). How much the price will be influenced depends on the liquidity of the asset. As the volume of the stock Royal Dutch Shell is very large, negative net order flow on some trading venue is not likely to influence the share price immediately. Thus, important notice is that the effect of net order flow on price depends on liquidity of the asset (see Section 0).

2.9. Trading of block orders

In nearly every market a large order can be executed in either one of two distinct ways of trading. Firstly an order could directly be sent to the 'downstairs' or secondary markets. The downstairs market consists of regular continuous intraday markets like Euronext and batch open (auction) markets such as openings (see Footnote 3). Prices are shown by quotes and trading can take place on a continuous basis. Secondly, a large order could be sent to the 'upstairs' markets (see footnote 3) where traditionally specialized brokers tried to find a counterparty. Next a price is negotiated and the Over The Counter (OTC) [11] trade is crossed according to the primary market regulation. Due to the liberalization of trading markets many hybrid venues between the down and upstairs market have

[8] A market order is an order placed on the exchange to be executed immediately against the current ask price (Petrella, 2009 p.6).

[9] Contrary to a market order, a limit order contains a value that will trigger either a sale or purchase of a share.

[10] For instance the SEC (Section 3.12.3).

[11] OTC trades are customized by two parties and not standardized and facilitated by an incumbent trading platform.

emerged. These alternative trading platforms, among others dark pools, will be outlined in Section 3.13.

2.10. Market fragmentation

As order processing costs for a trading platform decrease when trading volume increases[12], stock markets are recognized to have strong network externalities present[13]. As liquidity will attract liquidity, only one market will remain in the end. As a result, a logical expectation would be that stock markets in general tend to consolidate. The network externality puzzle refers to the fact that this not happens in practice (Madhavan 2000, p.23). Despite strong arguments for consolidation, markets tend to do the opposite and fragmented[14] especially over the last 5 years. Even if next to the primary market some satellite markets coexist, until today the first remained the major source of price discovery of an asset (Hasbrouck 1995, p.1197 and Menkveld et al 2006, p.20). This does not mean though, that because of migrating liquidity this price discovery process can not become disturbed. Schwartz (2009 p.4) states he is not concerned about dark pools, but only fears the ´s´ in the word. Dark pools are not eroding the quality of price discovery, but market fragmentation is.

2.11. Information

The majority of theoretical models, laboratory experiments and other economic publications agree on the rapid changes in investor behavior reacting to information (structure).The equal distribution of information among investors is in the US guaranteed by the consolidated data stream. All exchanges of any kind are obliged to inform the market regulator of any transaction completed on their platform[15]. This information is publicly available to the extent that the regulator requires this. In Europe regulation only requires price, time, and volume of trades is made public “on a reasonable commercial basis”, practically liberalizing the market for securities data (Petrella 2009, p.10). Data consolidation is a crucial component of the price discovery system. If there is a single and by government supervised reliable source this especially benefits retail investors to assess the quality and accurateness of prices.

2.12. Price discovery

Schwartz (2010 p.2) stresses the process of price discovery can be considered as much a public good as the beam of a lighthouse. He points out the discovered stock price shines light not only on trading itself, but also on valuation of underlying derivatives, comparable companies in corporate finance analysis, etc. Important other free riders deriving their prices from the public prices are dark pools. Securities in general are typically homogenous products as they all depend on the underlying stock price. Would investors be homogenous as well as assumed by many models (see for instance Fama and French´s famous CAPM model), they only care about the price they pay or receive for that stock. Because in the real world investors are actually heterogeneous (Carrie, 2008), next to price of an

[12] Higher trade volumes stimulate liquidity and therefore the holding period for market makers can be shorter (Madhavan 2000, p.23).
[13] Network externalities characterize those markets where the same assets (here securities) are traded simultaneously at the same price, but where the market with higher volume will grow at the cost of the competing one.
[14] Fragmentation of security trading implies scattering of trading volume over different markets.
[15] Both in the US as well as in the EU there exist some exceptions to these obligations (see Sections 4.2 and 4.4)

asset also other characteristics[16] can play important roles to determine the value (the price someone is willing to pay) for an individual investor. This suggests competition between exchanges for heterogeneous trading (demand), being the result of heterogeneous investors trading homogenous goods, actually exists. The ultimate goal of investigating market microstructure is to understand how and why prices change and what the role of information is in this process. Literature is generally consistent on price discovery being rather sticky and does not easily migrate from the home market to other markets. Hasbrouck (1995) finds the NYSE by then accounted for over 90% of price discovery (the Information Share) though trading had spread over many other public exchanges in the US. Furthermore Menkveld et al (2006 p. 5) find that price discovery of Dutch stocks is three times stronger during NYSE trading in Amsterdam, than during New York trading hours. Note that this all not implies though that the quality of the process can not change.

2.13. Conclusion

Though a fairly new field of research, researchers have developed a wide variety of literature on market microstructure. We understand better how many pieces of price discovery look like, but can still only solve small parts of the puzzle. High Frequency Traders (HFT) and other new players increased the speed and complexities of processes, but while changed, the old concepts such as market making remain. Liquidity, influenced by many factors and considered as the flow of information investors send to the market, is the main driver of price discovery. As such, a new, or more consciously present part of liquidity, dark liquidity, is an important concept to better understand as will be explained in the next Section.

[16] Speed of execution on the platform, available liquidity, etc.

3. Dark Pools of Liquidity

3.1. Introduction

The designation 'dark pool' is used for a range of different alternative trading platforms and is a relatively new and not by regulators defined concept. As such, unfortunately the terminology is dark itself and popular in usage by the media. In order to be correct and consistent it is therefore recommendable to look at characteristics instead of definitions. The different kinds that exist have some aspects in common: They provide liquidity, reduce market impact for large traders, compete with the traditional stock markets and they are opaque in releasing trade information to a varying degree. The number of dark pools operating is estimated to be over 40 in the US alone (Mittal 2008, p.2) and 60 worldwide (Degryse et al 2008, p.4) and the variety in characteristics of these pools seems to be just as large. In this chapter WE will explain what the working and motivation of dark pools is, which kinds exist and what benefits, risks and concerns arise when dark pools operate in the markets.

3.2. What is a Dark Pool?

Dark pools provide services to buy side traders (see Section 2.3) aiming to conceal trade intentions and reduce transaction costs. These alternative trading venues are part of the upstairs market (see footnote 3). In the USA dark pools are qualified as a type of Alternative Trading System (ATS) (see Section 4.2) and in the EU as either Systemic Internalizers (SIs) or Multilateral Trading Facilities (MTFs) (See Sections 4.3 and 4.5). Dark pools offer liquidity to institutional investors aiming to typically trade large blocks of for instance 100,000 shares. Rosenblatt Securities (2010) criticizes such absolute definitions by indicating a block trade can only be identified by expressing the order as a ratio of the firm's total market capitalization. As any trading venue, dark pools are obliged to display executed trades (volume, price and date) or quotes in a by regulation prescribed way to the public. In practice these requirements can be rather limited, due to exceptions applicable to protect their institutional clients (see Sections 4.2 and 4.4). In that case dark pools are recognized as proprietary markets and their trades are considered OTC. Transparency can also be influenced by denying certain more active players and brokers access to the platform and continuously police the platform[17]. The definition we will apply for dark pools during this thesis comprises: An alternative trading system originally designed for large block trades, which is in some way not completely transparent.

3.3. Dark Pools development

Since the beginning of the 1970s brokers have traded large blocks of stocks outside exchanges. In these third markets[18] they simply called potentially interested counterparties by phone. The brokers operating manually at the trading floor of the NYSE represented the biggest pool of invisible liquidity in the world (Haynes, 2008). As such OTC trading in the dark is nothing new. By becoming electronic in the 1980s Instinet, the Crossing Network and Posit defined the fourth market. The initial success remained rather limited though (Plexusgroup, 2004). The establishment of independent crossing

---

[17] Liquidnet is a more transparent dark pool but has a restricted entrance policy (Tabb, 2006).

[18] Distinguish primary market (see Section 2.3), secondary markets (see Section 2.3), third market and the fourth market (Block orders are commonly also traded directly between investors to avoid broker and transaction fees).

forums next to incumbent exchanges in the end of the 1990s was a new phenomenon in the fourth market. Originally created in 1987 POSIT electronic stock crossing systems came to Europe in 1998, providing a platform where every hour up to 7 moments a day a matching moment occurred. The platform provided services to pension funds and money management companies wanting to trade with limited market impact and pay low commission rates. In 2002 Liquidnet started trading in Europe as well and during the next period both venues experienced rapid growth as trading in the dark became more common and accepted. Whereas dark pool market share in equity trading was negligible in 2003, by December 2006 and 2010 Average Daily trading Volume (ADV)[19] equaled 8,92% and 12,14% of total consolidated trading in the United States respectively. A breakdown for all the major US markets can be seen in Figure 2. In Europe dark pools grew rapidly but by December 2008 dark pool ADV still equaled not even 1%, increasing to 2% by 2010. Following current trends experts indicate dark pool ADV in both the US and EU will further increase (Degryse et al 2008, p.4). During the financial crisis 2008 liquidity returned towards incumbent exchanges as traders prefer to trade in primary markets in distressed times[20]. Drivers of growth are likely to be in the first place changes in regulation, which by stimulating the rise of more alternative trading facilities gave extra impulses to all different kinds of liquidity pools. A second cause was the introduction of algorithmic trading, both because liquidity could be found easier (Degryse et al, 2008) as well as a presumably growing need to keep ones trading incentives of the market as the market becomes able to react to information increasingly faster. Saraiya and Mittal (2009) point out though protection of trade intentions is only a perception of investors as adverse selection risks in alternative trading venues can be significant (see also Section 3.10). Many other platforms were introduced, each having a slightly different target market and offering a faintly different strategy and product. Today dark pools can be registered and classified in many different categories (see Section 3.13).

3.4. Process

Traditionally a portfolio manager at some institutional investor deciding to buy a large amount of stocks would call his broker indicating his interest for a transaction. The broker then started calling counterparties trying to find one willing to sell for a reasonable price. In case the broker could not make the match, he would have to go to the open market to buy the shares. Dark pools changed this system. Nowadays brokers and markets (have to) use smart order routing algorithms[21] to deal with liquidity fragmentation in the markets and scan for the best available price or product. A dark pool aims to create a match before orders meet in a public exchange. Many players in the trading chain also integrate different functions. For instance an internalization pool, a combination of a broker and exchange, even first conglomerates all the order flow of not only the different departments of the investor which he has as a client, but also of all other clients. It then matches before only the remainder, the net order flow (see Section 2.8), goes to the public markets. The main hurdle for a dark pool is time, as trading is nowadays done in milliseconds. In order to handle this time constraint a dark pool needs to have 2 flows (Tabb 2008). Firstly there should be a large amount of orders in the waiting room, this would be called resident flow. Secondly transium flow heading for the market

[19] ADV denotes the average amount of traded stocks on a daily basis.

[20] Figure 4 on page 30 confirms this view, but market research institute Rosenblatt Securities stressed the crisis did not have much influence on the dark pool sector as a whole (Schack et al, 2009).

[21] In markets where stocks are dual listed, say on a public exchange and some ATSs, a smart order router can perform an instant search for the best price (as required by law RegNMS). These algorithms rapidly become more sophisticated and are used for internal optimization within trading systems as well. (Lin et al, undated).

should then be matched to the available liquidity in the resident flow. A dark pool can be both buy-to-buy side only as well as buy-to-buy and sell side. The first means only buy-side traders offer and take liquidity while the latter indicates also sell side traders are present to offer liquidity (see Section 2.3).

Some types of dark venues (for instance advertisement-based pools) became more aggressive and started to send out information about pending orders in the pool or receiving information from other pools to find a counterparty. So called automatic indicators of interest (also IOI, indications, alerts or conditional orders) are sent between pools in search of liquidity. The difference between these pools communicating IOIs and others or even the phone network brokers used to operate, is that these signals and the responding actions are no longer communication between individuals, but between trading engines. The risk of abuse or leaking of information is significant (see Section 3.9).

3.5. Transaction pricing

Firstly, common practice is the price in a dark pool transaction is derived automatically from the incumbent data information provider[22]. A second way of pricing concerns derived pricing from a selected public exchange (for instance the average price during the last 5 minutes of a stock at the LSE). A third way could be OTC price negotiation while the outcome of this is likely to be the market midpoint price mentioned above. In order to see why a possibly negotiated price in the dark pool is likely to be the same as the public market midpoint price[23], consider a situation with two buy side investors, party A looking to buy a large block of stock, and party B trying to sell an equal block of stock. If party A publishes his interest to the market the price would rise, if party B would indicate his interest the price would drop. As such both have an interest to keep the trade out of the market and minimize adverse price movements (see Section 3.6.3). To do so they could both turn to a dark pool. While in the dark pool a price can be discussed, both require the same volume and can basically not threat to fill the trade in the public market because of a possible adverse price movement. As such there is no alpha[24] creation possible from discussing the price in the OTC trade and the price can only be the value of the stock as for an investor in the public market. More specifically the midpoint price of the spread should be taken assuming a two-sided fee structure (see Section 3.6.1). The theoretical model above assumes though there is no information leakage and volumes of the two traders are equal as otherwise some potential for alpha value creation for one of both investors would exist. While a volume imbalance in public markets influences price, this imbalance can not be seen in a dark pool as not prices are discovered, but quantity (see Section 3.11). An unbalanced trade could simply be made and the residual buy or sell side order would remain waiting in the dark order book unknown to the outside world.

[22] National Best Bid and Offer (NBBO) in the US (Mittal, 2008 p13) or an incumbent exchange (often the home market) in Europe.
[23] The midpoint price is equal to the best bid price plus half the spread. It is exactly in between the best bid and best ask price.
[24] Alpha return indicates excess return meaning return above a reward for the cost of equity.

3.6. Motivation for investors

The fundamental motivations for traders to use alternative trading platforms, in this case in the dark, are twosome. In the first place, every trader both in small and large volume is always looking for alternative ways to trade when this enables saving on transaction costs. A second hurdle applicable to large block traders concerns the effect of signaling information to the market on the price of an asset. Madhavan (1995) describes block traders are afraid to be front-run as a result of information leakages and therefore prefer using the upstairs market. The upstairs market is shown to be a mechanism to diminish price impact of block-trading by risk sharing[25] (Madhavan and Keim 1996). Any dark pool can only try to minimize information leakage, but can not eliminate this (see Section 3.9). In their empirical research benchmarking Liquidnet trades with transactions in the home market of a stock, Ende and Muntermann (2010 p.1906) find i) Orders in a dark pool are less likely to be filled than in public markets, ii) If a counterparty can be found execution prices are significantly better than the best quotes in the primary market. Bikker et al (2007 p.976) research trades by pension fund ABP, one of the largest pension funds in the world. They discover the fund filling block trades in public markets cause average market impact costs of 20 basispoints (bps) for a buy and 30 bps for sell orders. Conrad et all (2003) show similar results indicating that dark pools indeed reduce both transaction and adverse price movement costs for institutional investors. Furthermore, Menkveld and Foucault (2006 p.24) as well as Degryse (2008 p.6) found public and dark liquidity markets together have a higher consolidated market depth than before the introduction of ATFs.

3.6.1. Transaction costs

Apart from the amount of transaction costs dark pools charge, fees are often charged according to a supplier-taker model (Mittal, 2008 p.5). This implies only the liquidity consumer, the buyer, is charged a fee for a trade whereas the liquidity supplier sometimes even receives a bonus (for instance BATS Trading). As every additional alternative trading facility fragments the market a bit more, dark pools are also contributing to decreasing transaction costs as a result of increased competition. Research shows dark pools do indeed provide better prices than on primary venues (Brandes et al, 2010 p.17)

3.6.2. Offering liquidity

Liquidity in markets used to be provided by market makers (see Section 2.5). Nowadays most regular financial markets offer continuous trading possibilities guaranteed by a limit order book. As a dark pool does not show any quotes, no market makers or visible limit order book can be present and liquidity needs to be provided in other ways. In dark pools liquidity is provided in three ways. Firstly, if the dark pool operator is a systemic internalizer (see Section 3.13) liquidity is fed by leveraging retail flow (Tabb, 2006). Secondly brokers cross public liquidity to the dark venues which means liquidity disappears from public markets which might in fact influence price discovery (Buti et al, 2010 p.4). Thirdly, the majority of liquidity is guaranteed, by connecting a dark pool with other dark pools (McEachern Gibs, 2007). A possible fourth but approximately resultant of the flow not matched to any of the 3 sources mentioned above, is the resident flow (see Section 3.4) ´already waiting' in

[25] The authors find that the price impact of block trades is temporary, and the effect is concavely related to the volume of an order. This also verifies the two tieredness of the market (see Section 3.5). This way the institutional tier is protected from informed traders in the public market tier.

the dark order book (Tabb 2008). As dark pools are subject to Reg NMS and MiFID regulation they need to guarantee best pricing respectively the best possible result for the client[26]. In order to skim all platforms where a stock is traded smart order routing algorithms have been developed[27]. Though currently under investigation by the SEC, one of the criticisms about dark pools in the US is the absence of a uniform reporting standard (Mehta 2009, Traders Magazine). This particularly concerns reporting of the quality of liquidity, which is not uncommonly also made public to attract more investors. Whereas it is generally agreed upon liquidity can be measured by the depth of the order book and the wideness of the spread, many dark pools apply a different definition of liquidity. They either double count, or include ‘touched’ volume[28] whereas traditional exchanges publish single-count matched trades. This obstructs investors to compare different dark pools.

3.6.3.Reducing market impact

The concept of market impact by volume on price goes back to the fundamental supply and demand framework presented by Adam Smith[29]. There, price is the resultant of equilibrium where supply equals demand. Market impact is therefore basically the effect of an increase in for instance demand on the equilibrium. The price will rise to the point when either the large volume buyer or other buyers are not willing to pay the high price anymore. The assumption no one is able to influence the market individually, as suggested in the famous paper by Fama and French (1970), in particular cases should be released. The price of a small stock might quite easily be influenced as long as the total volume wished to be traded is large enough compared to the daily traded volume. Empirical research has shown the market impact of a block trade is large in small cap stocks and is significantly influenced by trade volume and the size of the company (Loeb, 1983). Loeb for instance found stocks with less than $25 million market capitalization might experience a market impact of up to 15% as a result of a large block trade. For larger firms this rate drops to as low as 1%. Keim and Madhavan (1996) find market impact is a concave function of trade size, and a decreasing function of company size or liquidity. Empirical research done by Ready (2009) shows dark pool usage by institutional investors is lower as spreads on a stock are lower. This is consistent with the hypothesis that as more liquidity is available in public markets the possibility of adverse price movements decreases[30]. He further finds dark pool usage is higher for stocks trading at higher spreads indicating less liquidity available. As transaction costs are mostly decreasing with the volume of a trade, the motivation of institutional traders to trade a large volume at once is more likely to come from reducing market impact[31], than limiting transaction costs. For a large investor trading in a dark pool, the protection of information adds to the alpha a fund may believe to generate. A dark pool can only diminish the amount of information released after a trade, but not remove this (see Section 3.9).

---

[26] This obligation is mentioned in both Reg NMS as well as MiFID as the best execution practices obligation.
[27] The issue when scanning venues is discovering the amount of liquidity in every venue. The order flow get´s either completely filled by a platform, meaning more flow could have been sent, or only partly filled bouncing a number of orders. This problem is statistically known as censoring. Computeralgorithms can be designed to discover a near optimal allocation policy (Ganchev et al, 2010 p.1).
[28] Volume only routed elsewhere to be matched.
[29] Wealth of Nations, Adam Smith, 1776.
[30] Because a block trade will form a smaller part of available liquidity.
[31] The alternative of making a block trade is trading the same block in the public market, but sliced into a large amount of very small trades. This overcomes some modern risks for traders such as bad block trades or unattractive transaction costs becoming quickly evident. It is easier to blame an algorithmic trading program.

3.7. Motivation for operators

The operator of a dark pool is either an operator of a public exchange or a broker or consortium of brokers. As the operator has access to all information in the dark pool it should make sure firm internal policing is in place and well communicated to customers. The reason many dark pools are created is in the first place cost savings for the operator. Full automation and smart order routing in dark pools make sure phone-brokers are not necessary anymore. In the second place brokers try to consolidate all internal flow and match these either internally or in other ways to save transaction costs and cancel out opposing orders before hitting the public exchange. The redundant limit orders are kept in the dark order book and provide constant liquidity at hand.

3.8. Type of investors granted access

In order to protect retail investors, regulation allows only institutional and other large investors access to the less than public markets regulated dark pools. Except for wishing access to more information, a small investor would not have other motivation to trade in a dark pool. As some information transparency for retail investors is sacrificed enabling institutional investors to prevent price impact regulators should protect retail investors to this asymmetry. One thing is clear, high frequency and other short term gaining parties were not supposed to benefit from the situation. With current regulation in 2009 it was estimated 46% of trading volume in dark pools was accounted for by high frequency traders (Ortega, 2009).

3.9. Information leakage

As institutional investors seek protection of their trading intentions, an important issue for dark pools is the amount of information leakage. Most important information characteristics of an order are name, size, side, and time horizon in which it should be filled. Information leakage from dark pools can occur in three ways. Firstly, both in the United States as well as in Europe the regulator requires dark pools to display some information (see Section 4.6) which could be considered by law obliged 'leakage'. Secondly, information can be leaked to other potential investors by brokers for advertising purposes. Thirdly, leakage can occur depending on the type of counterparty a firm encounters in a trade. The latter two will be clarified.

3.9.1. Indicators of interest

A second way information may leak from a dark pool concerns leakage done intentionally by the dark pool. The party benefitting can either be the broker itself, the broker´s proprietary desk[32], or an external liquidity partner the pool has. Especially if IOIs or other ways of signaling are used information is supplied to market players, sometimes even without knowledge of the clients in the pool. Information leakage by the operator is inversely related to the amount of liquidity in the pool. Except for estimating the possibility of making a trade in a specific dark pool a client should therefore also check liquidity to assess (the urge for) information leakage.

[32] Proprietary trade flow is handled by a brokers´ proprietary desk. This means trading of a broker on its own account to generate alpha return.

3.9.2. Counterparty

As by definition any trade needs a counterparty, the latter will always be able do deduct some information. Mittal (2008, p.18) proposes a counterparty can in the first place be a non-informed investor seeking long-term alpha, in case he will not be interested in the signal. Secondly, it can be an informed trader who might actually use the information. Thirdly the transaction can be with a party trying to game (see Section 3.9.3) the dark pool. This is likely to try to use the investor´s trade intentions against him. Market impact is obviously the least when a trade is made with another buy-side, not information motivated trader (for instance a pension fund). Next to the counterparty in the trade, also the dark pool operator can see all trade intentions. This represents a risk for itself.

3.9.3. Gaming

Large outside market transactions are typically done between long term investors. Normally the price in a dark pool will equal the public exchange price and as such other players should not be interested in participating in the pool. Many short-term investors try to game dark pools though by trying to obtain information on trading volume present in the pool. If an institutional investor considers using a dark pool it is recommendable to try to find out who other constituents are and what their concentration is. Several techniques of gaming dark pools have been identified of which 'fishing' is the most known (Mittal 2008, p.15). In this technique a gamer determines the present volume in a dark pool by selling (buying) a small amount of stock. If such an order is filled it is quite likely more volume is present in the dark order book. Accordingly, the gamer buys (sells) large volume rapidly in the public market which makes the price move up (down). Finally the gamer offers (buys) a large amount of the stock in the dark pool against a higher (lower) midpoint price derived from the market. After this one minute process prices will revert to normal again. Other gaming techniques are midpoint gaming and market maker gaming inside the dark pool.

3.10. Adverse selection

Whereas primary, public markets are accessible for all investors, this does not hold for dark pools and other ATFs (see Section 3.8). Assuming trade-through possibilities are limited (especially the case in the EU as not required by MiFID) investors no longer have access to all information and inter-market differences exist. Liquidity providers are confronted with an adverse selection risk as institutional investors are likely to be more informed than the average trader (Hoffmann, 2010 p.2 and Degryse, 2008 p.6). As only larger institutional (informed) traders get access to dark pools liquidity providers will encounter better deals in the primary, open market where traders are less informed. Hoffmann (2010 p. 2) finds a typical dark pool indeed contains significantly more private information than primary markets. He further finds the competitive position of dark pools is negatively affected because liquidity providers recognize the adverse selection risk. When trading in a dark pool Degryse et al (2008 p.10) describe research has shown the advantage of lower trading costs than in primary markets can be completely offset by opportunity costs of restraining to trade due to adverse selection risk. As dark pools are increasingly operated by (consortiums of) banks providing their internal order flow as liquidity, the adverse selection risk for liquidity providers might be less present.

3.11. Hampered price discovery

To see why price discovery might be influenced by dark liquidity, note that as markets separate, by chance some platforms receive more buy orders whereas others receive more sell orders. If the amount of flow migrates to opaque alternative trading venues, price discovery is disrupted. That point will be reached if the balance between buy (demand) and sell (supply) orders displayed in the transparent markets does not represent the trend of all markets. Basically this means that prices do not represent the equilibrium of order flow anymore. As no one would have an overview of all markets to see the ´right´ price, massive arbitrage possibilities would exist as a result of information asymmetries. Traders possessing unique information could arbitrage to the real price. Estimations are that if traded volume by dark pools in the US reaches 15 to 20% the SEC will announce regulation to limit further growth as they fear price discovery might be disturbed (Tabb 2008). Schwartz (2009) argues though, sending large orders to a dark pool might as well facilitate price discovery as it is one of the least hampering ways to deal with block orders. In normal markets volume is given, and dealers discuss and ´discover´ a price (see Section 2.12). Now, by taking the price from the regular market as given, quantity will be the outcome and will be ´discovered´ in the dark pool. The alternative, sending a large order in a sequence of smaller tranches to the regular market, would still influence price discovery and might even give rise to momentum trading, further disrupting price discovery (Schwartz, 2010 p.2).

3.12. SEC worries

While dark pools are neither new nor very exotic, in 2009 the SEC expressed its concerns fed by the technological development and rapid growth of usage in dark pools. They faired transparency, efficiency and general fairness might be at stake. The commission started an investigation because if traded volume in dark pools would become to large, public prices might not be accurate anymore. Furthermore, if a significant amount of liquidity flows to dark pools, the SEC fears market participants not being a member of the pools might not be able to find enough contra-side interest anymore (Shapiro, 2010). Reacting as requested on the SEC's propositions the industry relativized the relation between dark liquidity and price discovery or stressed that dark pools do not represent such a large proportion of the market yet and that very little statistical proof can be given (Keegan, 2010). Others indicate additional data made public matters little to the crowd but would indeed benefit high frequency traders translating the information flow into algorithms (Spicer, 2009). They further indicate the funds of institutional investors, mainly pension funds, are owned by retail investors. By limiting the dark pool advantage for those traders this could actually have negative consequences for the public. Dark pool Liquidnet reacted by understating the false sense of security institutional investors might have after sending an order to the broker. They summon the SEC to force institutions to disclose all important handling practices to their clients (Merrin et al, 2010). Rosenblatt Securities (Schack et al, 2010) doubts whether this disclosure should be mandated by the government, or by the already slowly started process of market forces. The 2009 SEC proposal to change the existing rules consisted of three aspects which will be presented in the next Sections (Owens et all, 2010 p.5-10). An exception to all three rules would be applicable to very large transactions involving trades over $200,000 to disguise market impact for really large traders[33]. Also private IOIs remain allowed as long as they are only sent to credible potential counterparties.

3.12.1. Restate the meaning of the term ´quotes´

Many dark pools use IOIs to inform potential counterparties of outstanding orders in their dark trading book. As this way of providing information only informs selected investors, information asymmetry, especially with retail investors, exists. The SEC therefore suggested recognizing IOIs as offers and bids, which are already required to be made publicly available. This way the two-tiered information distribution, the public market and the private information in the dark pool is prevented. Some opponents of the SEC plans actually relativize and defend these two tiers, claiming the institutional investor market should have nothing to do with the public retail markets (Spicer, 2009). Obviously, dark pools not sending out IOIs will not be affected.

3.12.2. Decrease the non-display threshold

Furthermore, the SEC proposed to expand the range of ATSs obliged to publicly display best-priced quotes, this way making sure the market price takes into account all available information (The Daily Bell, 2009). Currently, if a dark pool displays orders to more than one party, and its average daily trading volume in that stock exceeds 5% of market capitalization, best-priced orders must be made public. The SEC obtains to lower this to 0.25%. In practice nearly all dark pools do not exceed this 5% threshold implying no dark pool ever publishes quotes.

3.12.3. Introduce reporting standards

Lastly, to answer complains of many large investors requiring transparency and comparability of different dark pools the SEC proposed to change the existing joint-industry reporting rules to after-trade publish the name of the dark pool where a transaction took place. This requirement would be applicable to all ATSs not only dark pools.

3.13. Different types of dark pools

Whereas dark pools often have many common characteristics, Mittal (2008) distinguishes 5 different types of dark pools.

One of the first types and in the US still popular are (public) crossing networks (CNs), mostly operated by brokers to generate commissions. Mittal (2008 p.3) states that clients are buy-side traders and one of the most distinguishing properties is the absence of proprietary flow from the operator. They are quite often innovative, uniquely designed platforms but after a few years of rapid innovation the market for CNs seems rather saturated. Some large players dominate the landscape and not many new venues are created. While CNs were originally ‘true’[34] dark pools, a variant can now also be observed as the advertisement based pools sending out IOIs to potential counterparties. The SEC defines CNs as “systems that allow participants to enter unpriced orders to buy and sell securities, these orders are crossed at a specific time at a price derived from another market” (SEC, 1998). As

[33] It would by advisable to define this boundary only as a ratio of a firm’s total market capitalization as explained in Section 3.2 and applied in other requirements as mentioned in Section 3.12.2.
[34] Distinguish dark pools not sending information whatsoever (true dark pools) from dark pools sending IOIs to possibly interested contra-parties.

the merchandise for a broker owned CN might be influenced by the performance of the parent firm this can have devastating results (Schack et al, 2009). In December 2008 Citimatch, owned by Citigroup, lost 37% of its trading volume, presumably because of the troubling situation of the parent company.

Another by regulation driven innovation to dark pools became popular over the last few years. Internalization pools are constructed mainly to internalize the trade flow of the operator itself, often large brokers (investment banks). Whereas these pools were initially intended to internally cross trade flow and save costs, soon also buy-side investors were attracted to the liquidity of the venues. The flow in these pools is likely to consist of consolidated retail flow, proprietary flow and buy-side customer flow. Next to cost savings, alpha generation and commission income, also the exposure of broker and investment banking services by the operator to buy-side investors is an advantage. As operator of the dark pool the broker can deny other sell-side parties access to the platform monopolizing the customer base inside. Internalization pools often attract liquidity partners for extra volume.

A third kind of dark pool are ping destinations. They have only proprietary flow which is tried to be matched to immediate or cancel (IOC) order of customers. These types of dark pools only contain the operator's own flow as volume in the dark order book. Operators are normally hedge funds and electronic market makers. As the matching chance for a buy-side trader is not very large in these pools, most clients are large sell-side traders. Prices for trading are normally very low and clients use the pools to check (ping) their volume.

Officially registered as ATS, dark pools operated by exchanges are called ´exchange based pools´. A very alike type of dark pool are 'hidden pools' (run within and by ECNs, the latter basically being exchanges as well), which do not directly match traders but only facilitate the ECN. The pricing of exchange-registered dark pools is typically on a per share basis whereas hidden pools apply supplier-taker models[35]. Quite typically both types tend to interact with orders displayed in the public market. Main purpose of running either one of both types is to increase available liquidity on the exchange or ECN.

A hybrid of CNs and internationalization pools are consortium based pools. A group of brokers creates a separate, independent organization having as a result a more transparent dark pool for clients[36]. For instance sell-side traders are not specifically denied access because of personal interest of a broker. Very often these broker partners also have their own internalization pool and use the consortium pool as a next level if an order can not be filled in the own dark pool. Consortium based pools are very price driven as the time to fill an order is often fairly short.

3.14. Conclusion

This section touched upon the many concepts and definitions involved with dark pools. Whereas the concept of dark liquidity has existed over 30 years, they recently became widely popular with beside their original public, institutional investors, also HFTs. By now they represent a (still growing) significant part of some 12% of US equity trading. In Europe, dark pools currently have some 3%

[35] In a supplier taker model the liquidity supplier is rewarded a premium and the liquidity consumer pays a (typically higher) fee.
[36] The different stakeholders are for instance likely to supervise each other.

market share, but are rapidly catching up with the US partly stimulated by a little more liberalized regulation. Recently, US and EU regulators showed their concerns about market liquidity, fair competition and the influence dark pools might have on the price discovery process. For now, tighter regulation has been suggested, but not yet introduced. In general dark liquidity is likely to be a permanent addition to the industry and complex regulation is needed to not let markets spin out of control. The next section will give an elaborate overview of RegNMS regulation in the US and MiFID in EU.

Figure 1: Development over time of ownership of dark pools. Values between January and June 2010 are estimations as data is not available (Source: Rosenblatt Securities)

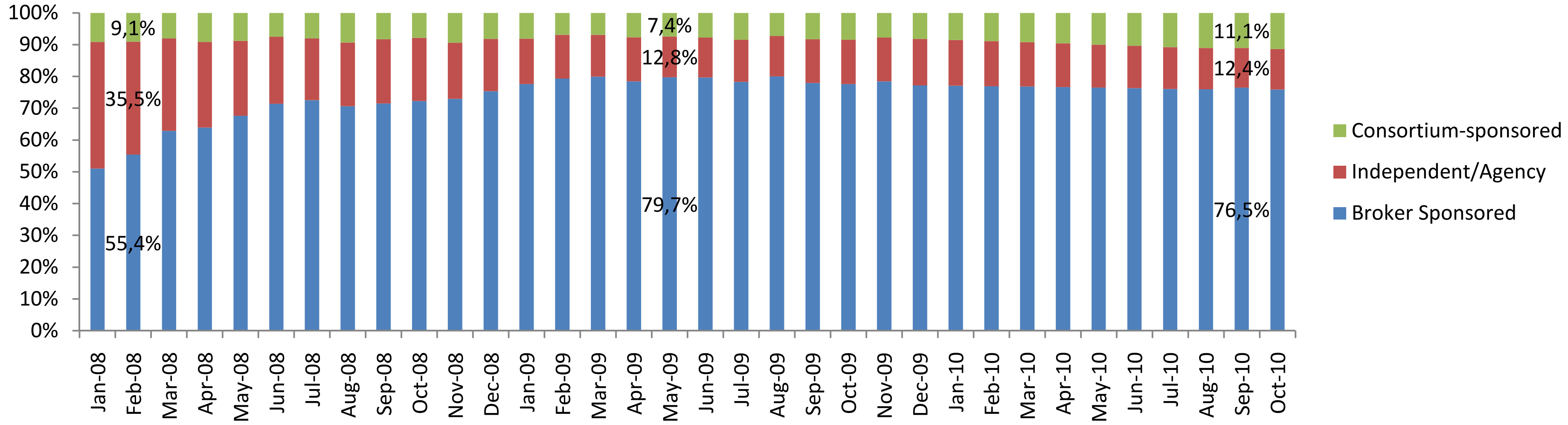
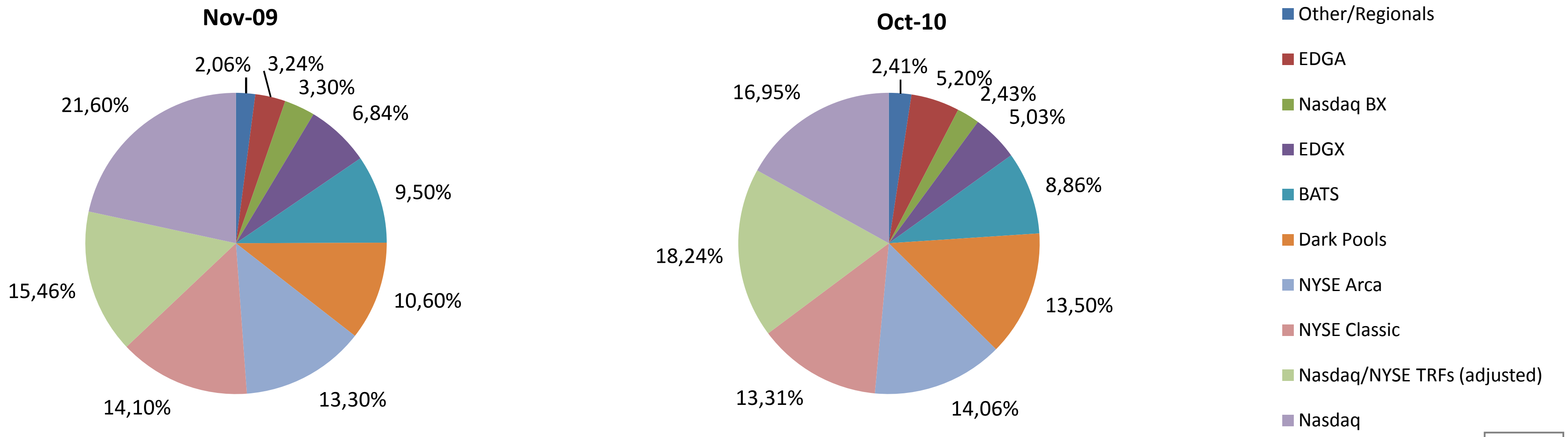


Figure 2: Changing of market shares over one year (Source: Rosenblatt Securities)



4. MiFID and Reg NMS

4.1. Introduction

In 2004 both the European Commission as well as the SEC introduced new legislation regarding trading facilities. While obtained to accomplish approximately the same results (mainly being stimulation of competition), the way legislation was set up was quite different. Contrary to traditional liberal practices, US legislation in this case is more defined and less liberalized than the European counterpart. In this chapter we will give an overview of the current legislator landscapes for trading venues in both the US as well as the EU.

4.2. Regulation in the United States (Reg NMS) 2005

Regulation National Market System (Reg NMS) originally introduced in 1975 but adjusted in 2005, recently intended to stimulate competition by making markets directly responsible for a best trade execution practice[37]. By doing so US Congress reconfirmed their rejection of a single unitary market model but safeguarded availability of information[38] by prescribing the linking of all trading platforms through Securities Information Processors (SIPs) [39]. Brokers and markets use smart order routing algorithms (see Footnote 21) to scan all available prices and find the 'best deal' for the investor. Assuming orders can always be executed immediately, the best execution practice only involves a best price criterion. This makes markets only compete on price. Another issue appointed by Reg NMS concerns the obligation for markets to provide best quotes and trades to one of the SIPs while they remain free to distribute their data to others. Exchanges, ECNs and dealers are furthermore obliged to report on a monthly basis several statistics regarding the quality of the executions[40]. Dark trading venues however, are not requested to report to the SIPs in real time, and can furthermore report anonymously (Schack et al, 2010). In order to provide the necessary access to quotations, Reg NMS enabled the use of private linkages between broker-dealers and trading venues[41]. To improve comparability of quotes as required by the Law, fees to have access to quotations were limited to $ 0.003 per share. This implies that if a share is displayed as $ 8.000, the actual costs of accessing the offer and completing the transaction may not exceed $ 8.003. As such, displayed prices can approximately be considered actual prices.

[37] A best trade execution practice obliges subjects to execute each order at the best price available in the market. If this is not possible the order should be redirected to a competitor offering better quotes (Petrella 2009 p.2).
[38] Guaranteeing information is available concerning which market is most attractive to execute an order is called pre-trade transparency.
[39] The most important SIPs are the Consolidated Tape Association (CTA), Consolidated Quotation Plan (CQ) and NASDAQ Unlisted Trading Privileges Plan (NASDAQ UTP).
[40] This post-trade transparency enables society to check which orders are finally executed (and check best execution prevision.
[41] This pre-trade transparency gives investors access to quotations of stocks on different trading venues.

4.3. Trading systems in the United States

The total of regulated markets include traditional stock exchanges (NYSE and NASDAQ), regional securities exchanges (National Stock Exchange and Philadelphia Stock Exchange), Alternative Trading Systems (ATSs), market makers and automated matching systems. ATFs bring together buyers and sellers of securities, and may make the traditional broker role redundant. The ATSs can be divided into i) Electronic Communication Networks[42], ii) Crossing networks, iii) Call markets and iv) Matching systems.

As from 1998[43] the SEC admitted Electronic Communication Networks (ECNs) to operate as ATFs. ECNs are digital trading venues offering clients access to buy and sell orders. Trades are typically anonymous where trade reports show the ECN as the party doing a transaction. Clients are not only large institutional funds, but also retail investors. ECNs are electronic markets often considered as the future of public stock markets. Trading through an ECN enables traders to directly trade without interference of middlemen[44]. Access to information can be limited to only the top of the trading book or the entire potential order flow giving valuable information about the depth of interest. Another benefit of ECNs is the possibility to trade after regular trading hours. ECNs can choose to be either i) a more costly[45] regulated exchange having own access to the ITS, or ii) a regulated (sub-enterprise of) a broker. ECNs are supervised by the SEC and obliged to display quotes on the consolidated tape. In case shares traded are listed on the NASDAQ, supervision is performed by the National Association of Securities Dealers (NASD)[46] and quotes are required to be reported to the NASDAQ UTP (see footnote 39). Fee structures may be two-sided or supplier taker models. Recently ECNs enjoyed a resurge after the introduction of Reg NMS, which required protection of orders in the market by obliging “trade-through” [47], regardless of where those orders are placed. ECNs contribute to price-discovery as they are basically digital variants of public markets.

Crossing Networks (CNs) are trading venues matching buy and sell orders electronically without first routing to an exchange or ECN which display a public quote. By doing so price impact is reduced. Different types of CNs exist and in the US the term is approximately equivalent to what can be considered dark pools (see Section 3.13).

Call markets are markets where orders are assembled and transacted in batches at pre-specified times. Pricing depends on the number of shares offered and is set by the exchange often at the market midpoint on some moment during the trading day. Normally exchanges are auction markets

---

[42] Broker owned ECNs are not equal to broker internal (electronic) crossing networks being internalization pools (see Section 3.13). An ECN is a regulated type of broker or exchange providing continuous matching of orders in direct contact with the public market (NASDAQ). A CN is a dark pool that electronically matches orders at from public markets derived prices.

[43] ECNs were allowed to trade stocks as a result of the OTC Pricing Scandal in 1997. At that time the SEC discovered that information about order flow was used to artificially increase spreads. As spreads were too high, this created large profits for dealers. The SEC decided these activities were in conflict with anti-trust rules.

[44] Middlemen can be described as high frequency, often arbitrage traders. They for instance buy a large amount of shares from a sell-side trader and resell these to numerous smaller investors. (Jovanovic and Menkveld, 2010 p.1)

[45] In 2002 the NYSE and NASDAQ spent over $142m and $500m respectively to control and supervise themselves (Wall Street Journal 2003). By now those costs are likely to be much higher.

[46] The parent organization of NASDAQ

[47] An order is not ´traded-through´ if it has been exercised at a less than optimal price in the market. The Order Protection Rule (Rule 611) aims to ensure that all investors get the best price in their trade.

implying orders are filled as soon as a buyer and seller together set a price for a given number of shares.

Matching systems are typically used by retail investors offering a venue where buyer and seller meet without interfering of brokers. As such these markets are very fair. Prices are derived from primary markets.

4.4. Regulation in the EU (MiFID 2007)

In order to create more competition, maintain transparency and protection of retail investors the EU introduced the Markets in Financial Instruments Directive (MiFID) in 2004, to become of force in 2007. MiFID tended to solve problems as internationally operating trading platforms obliged to report to different supervisors and abolished the concentration and related default rules[48] that used to be still present in a few countries[49].

MiFID introduced the "country of origin" concept indicating financial services providers in possession of a European passport do only have to report to the regulator in their own country. This way, a pan-European operating firm does not have to comply to (often conflicting) regulation of different countries anymore. The passport can be obtained from the own regulator. As the EC was like its American counterpart concerned about market fragmentation, to protect investors a best execution and handling practice was ordered to the investment firms. As such investment firms are obliged to at all times do everything possible to obtain the "best possible result" for a client. This best result should be an optimization of price, costs, speed, likelihood of execution, size, or any other for an investor relevant characteristic. By introducing a more extensive best execution practice, MiFID created a range of possibilities for players to compete by recognizing the heterogeneity of traders (Degryse, 2008 p.5). As only investment firms are responsible to obtain the best result for their clients, they use order routing systems to scan for the best deals. It is actually the preferences and characteristics of the investor that define the best deal a broker should look for. On the upside, this means trading venues can compete on many different aspects, whereas on the downside it might be difficult to route an order to the one best execution possibility (Petrella, 2009 p.2). On the contrary, trading venues in the US need to be linked to always optimize only prices. MiFID determines trading data can be reported to the competent authority by either the investment firm, a third party hired to do so, or by some other kind of authorized reporting system. Consequently the fragmentation of data reporting platforms does not contribute to transparency and the accurateness of price discovery. The market for assembling and publishing trading data is more liberalized than in the US. European trading venues are not obliged to (monthly) report any statistics on quality of trading. Degree of client protection in MiFID is based on 3 scales being (from most to least protected) non-professional (retail) investors, professional investors, and qualifying counterparties. The multidisciplinarity of the best execution practice and the absence of mandatory statistics on execution quality seem to make the European security trading less transparent than the US counterpart Reg NMS.

[48] Obliging broker-dealers to always execute an order on the incumbent exchange unless an investor had opted-out.

[49] For instance France, Italy, Spain and The Netherlands (until 2001) (Petrella, 2009 p.1)

4.5. Trading systems in the EU

In order to provide potential trading facilities opportunities to better adjust to investor preferences MiFID recognizes regulated markets, Multilateral Trading Facilities, Systematic Internalizers and a residual category containing other (foreign) systems.

A Regulated Market (RM) is an exchange for all kinds of financial products and is recognized and licensed by official authorities to facilitate trading in these products on a regular basis. In recent years the number of these traditional exchanges has shown a downward trend due to consolidation (for instance the merger of NYSE and Euronext). RMs experience fierce competition from alternative trading venues.

A Multilateral Trading Facility (MTF) is defined as a "multilateral system, operated by an investment firm[50] or a market operator[51], which brings together multiple third-party buying and selling interests in financial instruments – in the system and in accordance with non-discretionary rules – in a way that results in a contract in accordance with the provisions of Title II of MiFID". The main difference with a RM is that (the operation of) a MTF is considered as an investment activity meaning firms can not register their stock at a MTF (for instance public offerings) whereas they can on a RM (Haynes, 2008). This means a MTF can only trade stocks registered on some RM. Furthermore requirements to a prospectus publication of figures and price-sensitive information are less strict for MTFs (Degryse, 2008 p.3). As RMs, MTFs are obliged to display quotes and depth of the market.

MiFID created a third possible entity for security trading in "investment firms[52] which, on an organized, frequent and systematic basis, deal on own account by executing client orders outside a regulated market or MTF"[53]. Systematic Internalizers (SIs) are allowed to facilitate trading in all kinds of securities but article 27 prescribes publication of trade information in case of share trading. SIs are obliged to publish quotes of shares traded on regulated markets but only if three conditions are satisfied (Avgouleas, 2005 p.196) i) for which they act as SIs, ii) for which there is a liquid market and iii) for orders smaller than the standard market size[54]. An SI is allowed to have an own restricted policy on the entrance of clients to their services. Contrary to a RM or MTF a SI itself has the counterparty position in a transaction and does not only facilitate a trade of buyer and seller. SIs are only allowed to trade stocks and no other securities. This might recognize the fact that most clients of SIs will be large institutional investors. For instance a pension fund, is less likely to invest in derivatives than in the stock itself. Haynes (2008) states an SI is essentially nothing else than a system for banks "to make markets internally", off an exchange or MTF. Accordingly, MiFID trades SIs as mini exchanges, centralizing equal treating of customers in the pool. Internalization pools can not directly interact with public markets. As transparency requirements are less strict for large volume trades and

[50] Any legal person whose regular occupation or business is the provision of investment services to third parties and/or the performance of professional investment activities within the scope of MiFID. This includes portfolio management, investment advise, transmission and execution or orders relating to financial instruments and the underwriting of financial instruments (Hamilton, 2007 Wolters Kluwer MiFID review).

[51] Only a licensee of a RM is allowed to operate a MTF. An exchange can not directly operate a MTF (Kasbank MiFID review).

[52] Organization that provides financial services (not only services regarding financial products) to third parties, either professional or retail investors.

[53] Article 4(7) of the Directive.

[54] Not defined in MiFID.

trading venues with low trade frequency SIs provide scope for dark pools while transparency is required for retail investors trading small volume.

Financial service providers established and supervised outside the EU that wish to provide trading services in the MiFID-region are regulated as ‘other systems’. They can only provide services if their trading venue is comparable to either a RM or MTF. As foreign firms can not obtain a European passport, for each member state where such a system wishes to operate a license has to be obtained. Naturally, a separately incorporated affiliate of a non European firm would be considered an EU company and so, it would be a MiFID firm like any other European company.

4.6. Competition in trading

A central question in the debate for renewal of trading regulation, especially in Europe, concerns the tradeoff between on the one hand encouragement of competition in order to stimulate innovation and accomplish lower transaction fees, and on the other concentration of trading to optimize liquidity. A scattered trading landscape is likely to eventually result in less liquid markets and less accurate price discovery (see Section 3.11). Schwartz (2010 p.4) explains competition as a goal is ambiguous and a trade-of should be made. Firstly, if people speak about increasing competition related to the concept ‘markets’ they often mean competition between different markets. But secondly, one can identify competition for order flow (liquidity) and the process of competition for the best price. The fragmentation of order flow does not necessarily have to lead to less quality of the market though (Petrella, 2009). Especially now advanced technology (for instance smart order routing) is available to scan different trading venues in theory all scattered markets could be interconnected. Thus, all depends on the elasticity of the order flow[55]. Others such as both Schwartz (2010 p.6) as well as Sarkar et al (2009 p.4) confirm the theory of elasticity in order flow, but are skeptical about the interconnection of over 50 or more different markets as they expect this will always be imperfect. Therefore the collective information content, the access to liquidity and the quality of price discovery according to them will be harmed. In 2004 both the SEC and the EC decided to let the balance lean to some more fragmentation in the expectation competition between markets would increase. As a result in theory the competition for order flow decreases. And so does the quality of price discovery.

4.7. Conclusion

Both in the US as well as in the EU regulators more or less simultaneously introduced new regulation, RegNMS and MiFID respectively, aiming for more competition for incumbent stock exchanges. Whereas defined differently, characterization of the types of stock exchanges is quite alike. Important difference between RegNMS and MiFID though is the design of best execution policies: In the US defined as a trading platform obliged to guarantee the “best price” for a client, whereas in the EU a client should get the “best result available”. If a venue in the US is not able to offer the best price, the trade-through rule in RegNMS commissions to re-route the order to the cheapest platform. MiFID applies a broader definition and a trade might execute at an inferior price if this provides the best deal. The absence of a trade-through obligation gives rise to adverse selection risks (see Section

[55] The elasticity of order flow is the availability of information flow from different trading platforms, and the possibility to accordingly redirect order flow to the most efficient trading venue (Petrella, 2009 p.1). In practice though, the elasticity is never as perfect as when having one market.

3.10). Regarding the general effects of the introduction of more competition, the argumentation of Schwartz (2010) and Sarkar (2009) mentioned in section 4.6 could also be reversed. Fragmentation of the market will in theory at least not improve the access to liquidity and the resulting price discovery. The rise of dark pools amongst other ATFs has fragmented the exchange landscape and liquidity has moved away from incumbent exchanges (see Figure 2 on page 24). Assuming price discovery can be measured by quotes, in the next sections we will research whether the relation between quotes published on the NYSE on the one hand and non-NYSE (but still public) exchanges on the other, has changed over time. Comparing this to the increasing share of trading volume represented by dark pools and volatility, the effects of moving liquidity and related price discovery become visible. This might produce some statistical proof for the discussion described in Section 3.11.

5. Data Description and Methodology

5.1. Introduction

The main problem we encounter when researching the effects of dark pools on the process of price discovery, is the fact that nearly no data for dark pools is available (Keegan, 2010). For this reason the only way to make their effects visible is to consider the changes in how the other, visible markets interact. We will estimate the relationship between quotes on the NYSE on the one hand and all other exchanges on the other. This distinction was introduced by Hasbrouck (1995) in his research to identify the contribution of the NYSE to price discovery in the US equity markets (see Section 2.12). We will further research this relation by comparing it to volatility and the estimated share dark pools have of consolidated US equity volume over time. Especially the latter could provide some statistical evidence regarding the discussion whether or not dark pools indeed influence the price discovery process as described in Section 3.11. Section 5.2 will describe the US trading data from visible markets and the manipulation of that data to obtain usable series for statistical analysis. Furthermore the little data directly obtained on dark pools in the US is clarified after which the VIX volatility index will be introduced. Section 5.3 gives a brief introduction to OLS estimation after which Section 5.4 will illustrate how difficulties with OLS arise when variables such as stock quotes of different markets are cointegrated. In Section 5.5 we explain how to deal with cointegration by estimating a VEC model and Section 5.6 adds a GARCH model to review the covariance of quote series over time. Section 5.7 will finally introduce a total of 8 hypothesis to be tested by the models mentioned above.

5.2. Data description

In this research we measure the quality of price discovery by creating series for quotes on the NYSE and a combined average of all other public US stock exchanges following Hasbrouck (1995, p.1187). He states researching quotes instead of actual trades is expected to give better results as quotes are nearly continuously updated and trades happen less frequently. Furthermore, there is always a quote 'valid' to give the actual price of an asset, whereas the price resulting from a trade is by definition historical information. For this research a database was created containing all available quote data for 4 large, randomly selected S&P500 stocks trading on the NYSE[56]. The time frame is restricted to all quotes between 15.59.45 and 16.05.00[57] hours from 01/2005 – 06/2010. Data is obtained from the NYSE Trade And Quote (TAQ) database made available by Wharton Research Database Services (WRDS). The TAQ database contains all quotes of the 9 large public stock exchanges[58]. Firstly all quotes not belonging to the NYSE are cumulated to form two series containing NYSE quotes and non

[56] The stocks are IBM (IBM), Coca Cola (KO), Bank of America (BAC) and Walmart (WMT), tickers between brackets.

[57] Some quotes are posted and recorded after closing of the NYSE.

[58] The 8 public equities exchanges reported in NYSE TAQ are: i) NYSE, ii) NASDAQ, iii) NYSE AMEX Equities (formerly AMEX), iv) NASDAQ OMX BX (formerly Boston Stock Exchanges), v) National Stock Exchange, vi) Chicago Stock Exchange, vii) NYSE Arca (formerly Archipelago), viii) NASDAQ OMX PHLX (formerly Philadelphia Stock Exchange), ix) CBOE. BATS and DirectEdge's quotes are not reported in the NYSE TAQ database as they are not members of the Consolidated Tape Association (CTA).

NYSE quotes. Next daily closing quotes are created by first deleting so called zero bid[59] quotes of either $0,00 or $0,01. As this dataset includes all registered quotes for the time frame, the frequency is extremely high. In order to find real trends and ignore possible noise, accordingly the frequency is decreased by averaging all quotes (in some cases up to 400) to one per second. To create a closing quote per day the bid or ask closest to closing time of the NYSE, 16.00.00 hours, is selected. The result are four series: two containing daily closing bid and asks for the NYSE, and the other non NYSE closing bid and asks, during the sample of 5,5 years. Most tests will be performed on growth in the quotes, defined as the natural logarithm (ln) and referred to as either ln B for bid, ln O for offer, or together ln(quote). A second possibility is testing on the return of the quote, the difference of the natural logarithm of the quote, presented as Dln B, Dln O, Dln(quote). Finally, many outliers are present in the price, especially offer, series. Therefore we change any quotes showing a spread larger than 2,0 to NA. As the Biv GARCH model we describe in Section 5.6 requires continuous samples, only there we replace all NAs by interpolating.

Quite some empirical research on individual dark pools providing data to academics has been performed (See for an overview Degryse et al, 2008 and a recent paper by Hoffmann, 2010). At this moment though, only two (commercial) research institutes provide quantitative macro-economic data on the total number of dark pools and their daily trading volume processed. In the first place The TABB group, a financial markets' research and strategic advisory firm and a pioneer in dark pool research has published many research articles including quantitative data. Secondly Rosenblatt Securities, a top-20 brokerage and furthermore advisory firm for institutional investors, has been publishing one of the industry's leading sources of information with their monthly reports on dark pool development[60]. From especially the Rosenblatt monthly reports we constructed both a general overview of the major US dark pools, as well as a monthly estimation for the cumulated share dark pools had in US total consolidated equity volume (US equity volume) from 2007 till 2010 as can be seen in Figure 3 and Figure 4. We estimate a linear approximation for the monthly dark pool share (see Figure 4, coefficients are highly significant and R^2 is 0.845 on 35 observations). As the linear approximation seems to be very strong we generally assume dark pool market share indeed grows linearly and therefore interpolate the monthly observations to daily estimates.

The financial markets' leading volatility indicator is the Chicago Board Options Exchange's (CBOE) Volatility Index (VIX) introduced in 1993. The index is supposed to be 30 day forward looking and is calculated using the implied volatilities of both call and put options on S&P500 stocks. We obtained daily closing prices from the CBOE website. A graph of the VIX from 2005 till 2010 can be observed in Figure 5. From this we constructed a dummy variable taking the value 0 when VIX is lower than it's third quartile, and 1 if VIX is higher.

[59] Zero bid quotes can either be withdrawn quotes and broken trades (TAQ 3 User's Guide, p. 26) or so called stub quotes (SEC, 2010 p.63). A stub quote is submitted by a market maker with no liquidity available, but by exchange rules obligated to submit a quote continuously. The offer is made so far away from the market price that it is not intended to and will not be executed.
[60] The reports named "Let There Be Light" are available against payment.

Figure 3: Graphs showing the average daily trading volume in absolute value for all equities traded in the US and for all equities traded in dark pools in the US.

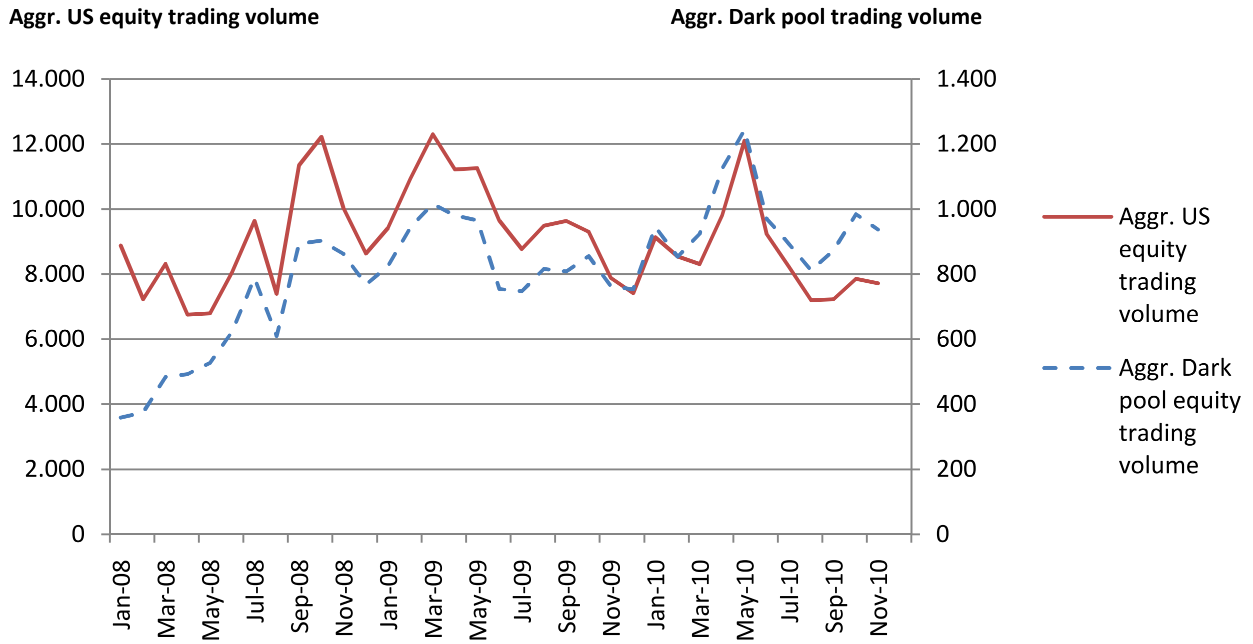


Figure 4: This figure shows the dark pool share in US equity trading as well as a linear approximation.

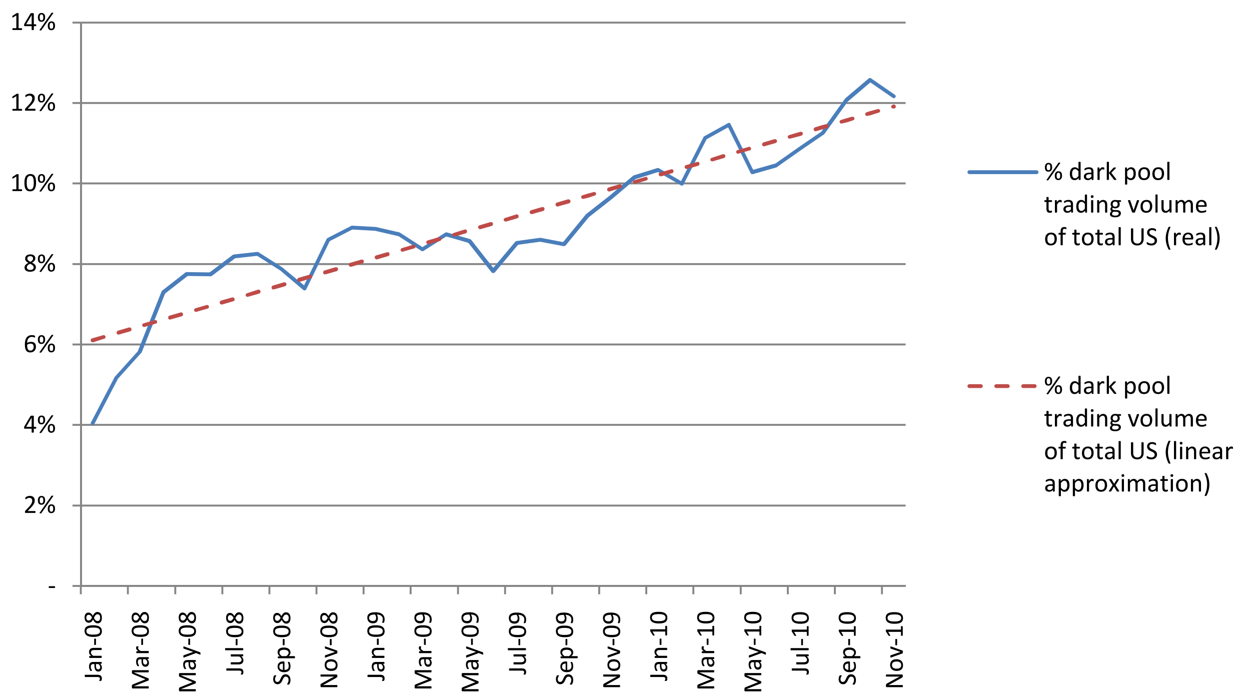
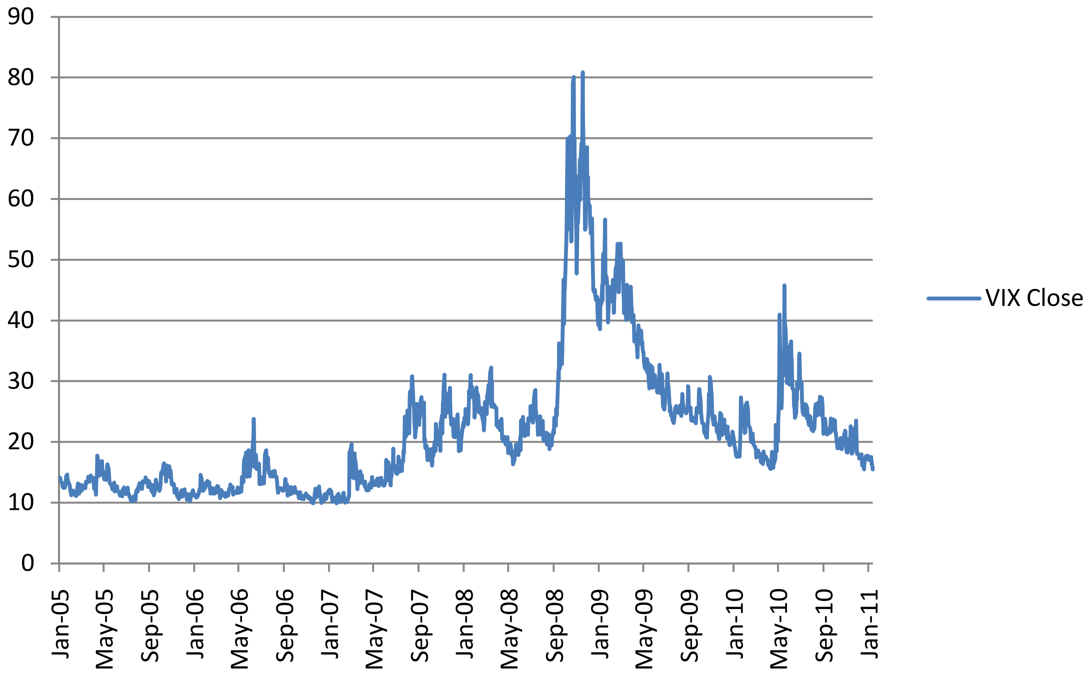


Figure 5: This figure presents the VIX index from 2005 till June 2010.



5.3. Ordinary Least Squares

As explained in Section 2.2 different markets can contain small permanent price differences as a result of the quality of services provided. The random trend in a quote as a result of the underlying stock value fluctuating will be present in all markets though. The most commonly used method to describe a relationship between two variables (here stock quotes for two defined markets) is the Ordinary Least Squares (OLS) method. OLS assumes a time series is a linear function of some variables over time which can be represented by:

$$Y_t = \beta_1 + \beta_n * X_t + u \tag{1}$$

Where Y_t is the explained variable, β_1 is a (constant or intercept) coefficient and β_n is a (slope) coefficient both to be estimated by OLS, X_t is the explanatory variable and u is the error or residual term.

OLS assumes residuals u of the equation are only temporary, random disturbances having a zero mean and being homoskedastic[61]. As such the variables will converge to their long term values as the sample size increases. In this case the variables are said to be stationary and the regression to be consistent. If these assumptions do not hold, at least one of the variables contains a trend and is said to be integrated of order one, often represented by I(1), or said to contain one unit root and to be non-stationary (Brooks 2007, p.387-389). In case of I(1) the impact of an external factor does not fade out and values will not return to the long term mean. As such the trend will be visible in the residuals. The problem with these series is they can not be modeled using OLS as this method can not model a trend in the residuals. The concept is illustrated by Figure 7. In case both variables contain a (different) trend both variables and their residuals are said to be I(1). Stationarity of variables can be tested by performing a unit root test such as the Augmented Dicky Fuller (ADF) and Phillips-Perron (PP) test (Brooks, 2002 p.379-381). The ADF test has as null hypothesis the presence of a unit root in the series, so non-stationarity whereas the PP test has the opposite.

5.4. Cointegration

Most financial time series are I(1). After having confirmed this by a unit root test as described in Section 5.4, the relation between these series can be investigated. In case two series are both I(1) and contain the same trend, the combination of these series in one equation will normally be I(1) as well. Brooks (2007, p.388) states though that in case a linear combination of I(1) variables sharing a same, common trend is itself I(0) the variables are said to be cointegrated. One could imagine cointegrated variables, for instance disposable income and consumption, as moving together over time influenced by market forces. Cointegrated variables are also said to be super consistent if the sample is large enough, referring to Stock's (1987) finding that they actually convert faster to their long-term values than combined stationary series. Combining the variables containing that trend isolates the trend, reducing the residuals to I(0). Whereas most financial data includes one unit root, higher orders of integration exist as well. Berenguer-Rico et al (2007) explain in case the cumulated, by definition I(1) residuals[62] are cointegrated with the I(1) original variables a second order of cointegration occurs. A higher order of cointegration might for instance imply correlated

---

[61] Homoskedasticity implies a constant variance. Variance is the square of the errors (residuals) of an OLS regression.

[62] As the original variables are I(1).

accelerations, indicating not only the variables contain the same trend, but also the residuals are adjusted to that trend with the same speed. Berenguer-Rico et al (2007) mention cointegration of more than two variables or by variables of different orders is also possible. This is called multicointegration. Note though, that the finding of several cointegrating relations in a set of variables would still be cointegration of order 1 as this a different concept from multicointegration.

Deciding on applying OLS, one should first make sure the series are not cointegrated. In case of cointegration a more sophisticated model is necessary. Tests for cointegration include the Engle-Granger 2-step method (1987)[63], the Gregory-Hansen method (1996)[64] and the Johansen VAR technique[65] (1991). As suggested by Brooks (2007, p.395) we here apply the latter as this allows for cointegration testing directly on the variables instead of the residuals. The null hypothesis is the number of cointegrating relationships is r against an alternative of $r + 1$ (Johansen, 1991). The test will iterate until reaching the maximum amount of cointegrating vectors present. An extended explanation of applying a Johansen test can be found in Appendix 9.1. The Johansen test will be performed on the whole sample, as well as two smaller samples of each 2 years. The first are the years 2005 and 2006, before the rapid growth of dark pools, the second 2008 and 2009[66], after dark pools started do gain significant market share. These samples will accordingly be indicated by 'before dark pools' and 'after dark pools'. As this is a fairly rough method to test for a change in cointegration, only in case dark pools have great impact on the relation between the NYSE and non NYSE series the second sample might in fact show indications of clearly less or even a lack of cointegration.

Brooks (2007, p.389) explains the earliest assumed solution to deal with cointegrated variables was to simply take first differences. But though statistically valid this does give the problem these first differences do not have a long term mean to which they revert. As such OLS would neither give a valid solution nor equilibrium relationship in the long run. An alternative, modeling the trend as autocorrelation, gives the same problem. The solution to give a good representation of two cointegrated series is to use a VEC model as will be explained in Section 5.5.

The log-likelihood provided in the tests indicates the fit of the model. This indication is not as general as R^2 though as models can only really be compared if one is an extension of the other. Generally speaking, a higher log-likelihood indicates a better fit. Whereas the number of coefficients in different models is unequal though, the values can not directly be compared. The difference between two models multiplied by 2 is chi^2 distributed with the difference in number of coefficients as degrees of freedom.

---

[63] Which first estimates a cointegrating regression by OLS and then performs a unit-root test on the residuals of the first regression (Brooks 2007 p.393).

[64] The Gregory and Hansen (1996) test additionally takes into account the cointegrating vector (the underlying trend) is actually likely to not be constant over time. The method applies an unknown structural brake in the cointegrating relationship. The method is especially advisable in case the sample period is long.

[65] A Vector AutoRegressive (VAR) model models the relationship of both the explained variable itself as well as multiple explanatory variables on their previous lags. The Johansen VAR test also allows for multicointegration.

[66] A correction for the financial crisis is not necessary as we consider normalized price values and furthermore both series will be equally influenced.

Figure 7: The red line contains a unit root and does not return to the long term mean whereas the blue line does. The effects of the cointegration can be dealt with by correcting the residuals of the red line using a VEC M.

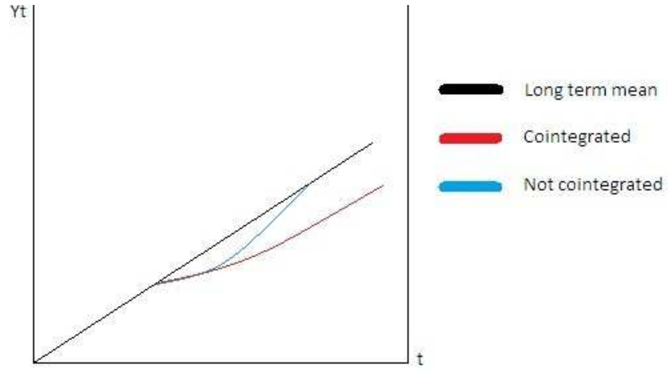


Figure 6: The VEC M corrects every value of Y for all t.

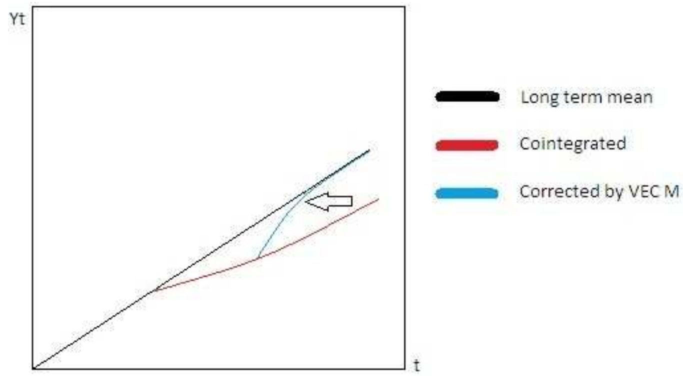


Table 1: This table contains the possibilities regarding trends in variables. Assume regressions in all 5 cases are found to be significant and furthermore different trends such as A, B, etc could be present. Only in case 5, if both variables contain the same trend cointegration, occurs.

	X variable	Y variable	Regression	Possibilities
Case 1	Stationary,	Stationary	Consistent	Use OLS
Case 2	Trend A, I(1)	Stationary	Spurious	Detrend or take first differences
Case 3	Stationary	Trend A, I(1)	Spurious	Detrend or take first differences
Case 4	Trend A, I(1)	Trend B, I(1)	Spurious	Detrend or take first differences
Case 5	Trend A, I(1)	Trend A, I(1)	Consistent	Cointegration, use VEC M (and OLS) or other model

5.5. The VEC model

Having found the series are cointegrated, a Vector Error Correction Model (VEC M) estimates the trend in the residuals by correcting the deviation from the long-run equilibrium by using an error correction term. The latter contains short run corrections to get back to the long term mean as illustrated by Figure 6. A VEC model is a type of Vector Auto Regression Model (VAR M). Several VEC Models, often also called Equilibrium Correction Models, will be estimated, represented as by Brooks (2007, p.390).

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) + u_t\end{aligned} \tag{2}$$

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) + \beta_6 * \Delta \ln DPs + u_t\end{aligned} \tag{3}$$

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) + \beta_6 * \Delta VIX + u_t\end{aligned} \tag{4}$$

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) + \beta_6 * \Delta X_{t-1} * \Delta VIX + u_t\end{aligned} \tag{5}$$

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) + \beta_6 * \Delta \ln DPs + \beta_7 * \Delta VIX + u_t\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta Y_t = \beta_0 + \beta_1 * \Delta X_{t-1} + \beta_2 * \Delta X_{t-2} + \beta_3 * \Delta Y_{t-1} + \beta_4 * \Delta Y_{t-2} \\ + \beta_5 * (Y_{t-1} - \gamma X_{t-1}) * \Delta \ln DPs + u_t\end{aligned} \tag{7}$$

Assuming Y_t and X_t are cointegrated with coefficient γ, the equation implies changes in the dependent variable Y_t (NYSE) are determined by changes in X_t (non NYSE), changes in Y_t and by the error correction term $(Y_{t-1} - \gamma X_{t-1})$. Obviously, X_t is estimated as dependent variable as well though here note depicted. $\Delta \ln DPs$ represents the change in the normalized value of the dark pools' share in US consolidated equity volume and ΔVIX the change in the dummy for high volatility. As explained in Section 5.4 the error term containing a linear combination of the original variables will now be I(0) even though the variables are I(1). As such it would again be possible to use OLS and other statistical estimation methods.

Naturally in both the equation as well as the error term an intercept (constant) could be included depending on whether economic arguments suggest this. Specifying the VEC M we choose to include an intercept (β_0 in the equations) but no trend (Eviews 6 User Guide II, p.369) as the movements in the quotes on either the NYSE or the non-NYSE series are assumed to be random (see Section 2.2). The number of cointegrating vectors is 1 as the cointegration test indicated 1 cointegrating relationship. Furthermore we choose to specify the lags as 1 2, implying testing the influence of $\Delta X_{(t-1)}$, $\Delta X_{(t-2)}$, $\Delta Y_{(t-1)}$ and $\Delta Y_{(t-2)}$ on $\Delta Y_{(t)}$ and visa versa. Furthermore, there is no reason to restrict any of the coefficients (Eviews 6 User Guide II, p.369). Dependent variable Y_t will be either the ln(bid) or ln(offer) value of the NYSE quotes whereas the independent variable $\mathrm{X_t}$ will be the equivalent of the non NYSE quotes (equation 2). The VEC M will furthermore be extended with indicators for high market volatility and share of dark pool trading in total US equity trading volume. This implies 2 extra beta's are added and equations 3 and 4 are formed above. In Equation 5 the dummy for high market volatility is combined with the lagged value of the NYSE series. Equation 6 is a combination of Equations 3 and 4. The ECT is combined with the growth in the dark pool market share in Equation 7. The β coefficients can be estimated by the VEC model in Eviews.

5.6. The Bivariate GARCH model

An AutoRegressive Conditional Heteroskedasticity (ARCH) model was originally presented by Engle (1982) and models heteroskedastistic residuals, u_t. The movements in the conditional variance[67] of u_t are essentially the same as the movements in the conditional variance of the variable Y_t. Hence modeling the variance of u_t models the variance of Y_t as well (Brooks 2007, p.482). Later the ARCH model was generalized to the GARCH model by Bollerslev (1986) and by Taylor (1986) and by now many different variants exist. When researching the relationship between several assets one could estimate a GARCH model for more than one variable and apply a Multivariate GARCH (M GARCH) model for two variables. Estimating a GARCH model for both series together shows how the errors of the variables move over time. Also indicated as Bivariate GARCH model (Biv GARCH) we will apply this both on i) the ln(quote) and ii) return series, as well as iii) on the residuals of the VEC M. The Biv GARCH model used here is the BEKK GARCH (1,1) proposed by Engle and Kroner (1995). Note that a GARCH model describes the correlation between two variables based on heteroskedastistic modeling of their residuals. In case iii described above we will use the GARCH equation to model the residuals of the residuals. That is: the correlation between two variables (in itself residuals of the VEC M) by modeling the residuals of the variables. A technical explanation of an ARCH model and the BEKK GARCH (1,1) model as described by Van der Wall (2008, p.9) can be found in appendix 9.2 and 0.

5.7. Hypothesis to be tested

In this section the rationales for a total of 8 hypothesis will be introduced, whereas an overview can be found in Table 2 on page 45. The results of the tests can be found in Section 6.

In order to check whether cointegration is present we will apply a Johansen test as described in Section 5.4. We expect cointegration of order 1 as both quote series for the NYSE as well as non NYSE share a common trend, the underlying value of the stock, and any difference should be related to market imperfections (see Section 2.12). The test will both be directly applied on the quotes as well as first differences(returns). Quote series are nearly identical and are likely to be cointegrated, whereas the return series can differ more. Hypothesis I has a null hypothesis of r cointegrating relations, against an alternative of r+1. The Johansen test iterates till the present number of cointegrating relations, the first iterations being r=0. For the first hypothesis we expect 1 cointegrating relationship. In Section 5.4 we announced to also perform the Johansen test on a 'pre dark pools' sample and an 'after dark pools' sample. We expect hypothesis II to show less clear, no, or at least changed, cointegration in the second sample.

[67] Conditional variance of a series X is the expected variance given another or more series (Brooks, 2002 p.446).

Incumbent exchanges are likely to still contribute most to price discovery as described in Section 2.12. The informational content is highest in the primary market. As such we expect changes in the non NYSE series are stronger influenced by changes in the NYSE series than visa versa. This results in null hypothesis III of a change in the NYSE series having a stronger impact on the non NYSE series in the VEC M, than visa versa.

As much deliberated upon in Sections 2.10 and 3.11, dark pools retract liquidity from displayed markets. As less information enters the price discovery process the quality of the process declines and price differences between markets grow. Consequently the covariance between NYSE and non NYSE markets should be inversely related to the market share dark pools have in consolidated equity volume. This leads to null hypothesis IV of a coefficient for dark pools not being significant in explaining changes in either the NYSE or non NYSE series in the VEC M.

As return often increases slowly in times of low volatility, but decreases strongly when markets endure high volatility, we created a dummy variable equal to 1 in case VIX is higher than 3 quartiles of its distribution. The corresponding null hypothesis V proclaims a coefficient for the high volatility dummy added to the VEC M will not be significant in explaining changes in either the NYSE or non NYSE series.

In theory, as was argued in Section 2.12, we would expect the correlation between the NYSE and non NYSE series to be constant as this would imply arbitrage opportunities are constant (to zero/ absent). Absent technological innovations or other external influencing factors, correlation should be constant, ideally perfect to 1 as this would imply one, the same price in all markets. This results in the null hypothesis VI stating the Biv GARCH M covariance calculated between the NYSE and non NYSE is constant.

The main argument of this thesis regards the possible influence dark pools might have on the price discovery process, represented by the covariance between NYSE and non NYSE quotes (See Section 5.1). The test is relevant as this relation might confirm the disturbance of price discovery (See Section 3.11). As such the null hypothesis vii states dark pool's share in US consolidated equity volume is not negatively related to covariance between the NYSE and non NYSE series.

Rosenblatt Securities (Schack et al, 2009) suspects a negative relationship between volatility on the one hand, and the % DPs claim of US cons.eq. volume. In their monthly reports Rosenblatt often suggests a close inverse relationship between dark pools share in US equity volume and market volatility[68]. They never mentioned having tested this empirically. They state that if volatility increases, the share of dark pools drops for two reasons. Firstly, if volatility and market instability increase, institutional investors trading in dark pools "stick to their gut's and wait on the sideline" (Schack et al, 2010). Secondly, if volatility is high, institutional investors prefer to trade in a primary market above waiting for liquidity in a dark pool and experiencing more risk. Therefore null hypothesis VIII consists of the dummy variable indicating volatility was high not being positively related to the share dark pools have in consolidated US equity volume.

[68] Volatility or variance is the squared difference of the return on a financial instrument and is generally interpreted as the main measurable indicator for risk during a specific time frame.

6. Results

6.1. Introduction

This Section includes a presentation and explanation of the results from the tests proclaimed in Section 5. The output is available in overview in the appendices, Section 9.4. An overview of the outcome of the hypothesis announced in Section 5.7 are presented Table 2.

6.2. Johansen covariance test

The results of the Johansen cointegration test for the sample 01/2005 – 06/2010 are reported in Table 3, Table 6, Table 9 and Table 12. The null hypothesis of zero cointegrating relations is rejected in all cases. The test on the ln(quote) series indicates nearly consistently 1 cointegrating relation whereas the Dln(quote) series show 2 cointegrating relations in the return. As expected, the price series show a common trend, the value of the stock, and the difference is probably due to small intermarket differences (see Section 2.12). The return series we expect to not share a trend and to be randomly distributed around a zero mean. Clearly the ln(quote) series are both I(1) and together cointegrated of order 1 whereas the return series are I(0) and together give a test result of 2 cointegrating relationships which is inherent to being stationary. In general we can conclude hypothesis I, being a lack of a cointegrating relationship in the price series, can be rejected in case of the ln(quote) series for all shares, for all reported data trends. The ln(quote) series of the NYSE and non NYSE are indeed cointegrated.

Log-likelihoods are shown for all tests, for 0, 1 and 2 lags. As these values give an indication of the fit of the model for the data, they can be used to compare different data trends (see Section 5.4). Contrary to R^2 though, this method for comparison can only be trustfully used in case one model is a restricted version of another, which here is indeed the case. Clearly, the differences between the datatrends (multiplied by 2) in none of the cointegration estimates will ever exceed the chi^2 distribution, 5% critical value of 3,84 (1 degree of freedom) or 5,99 (2 degrees of freedom. As a conclusion we can not decide on which data trend best describes the cointegrating relationship between the NYSE and non NYSE series.

The results of a Johansen test performed on two samples, 'before dark pools' and 'after dark pools', can be observed in Table 4 and Table 5 for IBM, Table 7 and Table 8 for KO, Table 10 and Table 11 for WMT, Table 13 and Table 14 for BAC. We observe cointegration is less clearly revealed for both samples as the Johansen test now indicates 1 or 2 cointegrating relationships in the ln(quote) series. When estimating for a shorter sample of 1 year results become even less clear and cointegrating relations are of order 0, 1 or 2. In general the differences are not consistent with hypothesis II and are clearly due to shortening of the sample in which we look for cointegration. As cointegration is typically present in time series, testing for this indeed requires samples over multiple years in fairly highly frequented data (for instance on a daily basis). From these results we can not conclude cointegration changes between the 2 samples and hypothesis II of cointegration between the NYSE and non NYSE series does not change from one sample to the other should not be rejected. A finding of 0 cointegrating relations in the 'after dark pools' sample would in fact indicate really large changes as the relation of bids and offers with the underlying value of the share would be hampered in such a way that they do not move together anymore. Unless exchanges can really distinguish from others and differences are caused by offering a different product (See Section 2.11), this result would be fairly impossible to obtain. Considering the changing cointegration over 2 samples is only a rough method.

6.3. VEC Model

The estimated VEC models for the ln(quote) series can be observed in Tables 15-22. VEC models are estimated on only the natural logarithm of the series as presented in Equations 2-7 (Section 5.5), with addition of an external variable for dark pools' trading volume (Equation 3), with addition of an external variable for volatility (Equation 4), with addition of an external variable for volatility, combined with a NYSE lag (Equation 5), one model for dark pools and volatility present together (Equation 6) and one model where dark pool share forms part of an ECT (Equation 7)[69]. As the Dln(quote) series are shown to not be cointegrated (See Section 6.2), a VEC M is not estimated for these. The VEC M output tables show independent variables in rows, and dependent variables in columns. We assume a critical t-value of 1,96 corresponding to a 5% significance level.

Firstly, and fairly consistent for all shares in both bid and offers except for offers IBM, the ECT for the non NYSE series is highly significant and positive. This implies the ECT corrects these series towards the NYSE series. The influence of ECT for the NYSE is not significantly present and mainly negative. Therefore we may conclude the VEC M indicates the non NYSE series move towards the NYSE series as corresponding to the price discovery theory (see Section 2.12).

Secondly, an even more clear indication of the non NYSE following the NYSE series might be found by looking at the effect of changes in the lagged NYSE series. We expect these to have significant impact on the non NYSE series whereas both the NYSE and non NYSE series as dependent variables should have an equal sign. In general, indeed both dependent variables have equal signs for the lagged variables. This implies both series react in the same direction on changes in the NYSE series, which is consistent with the observed cointegration in Section 6.2. The t-values for the lagged NYSE series on the non NYSE series are not consistently significant though. Conclusion regarding hypothesis III of the NYSE having stronger influence on the non NYSE series than visa versa (stronger price discovery) is that we find indeed the non NYSE moves towards the NYSE series, but the lagged values of the NYSE series do not have significant influence.

Thirdly, a coefficient for dark pool trading volume as external variable in the VEC M (See Equation 3) shows not significant, though a consistently negative effect on the dependent variables. This would imply an increase in dark pools' market share results in lower return for the NYSE and non NYSE. An explanation might be found in liquidity draining away from primary markets. As such a liquidity premium could squeeze returns in primary markets. A more sophisticated model to test for the impact of dark pools is adding an extra error correction term where the original correction term of the lagged NYSE is strengthened or weakened by multiplying with the dark pool market share indicator (see Equation 7). The results of this test are separately provided in Table 23 and Table 24. Only for the stock lnB^{WMT} , lnO^{WMT}, lnB^{IBM} and lnB^{KO} this parameter is significant and negative as expected. The factor implies an increase in the share of dark pool trading volume results in a stronger negative correction effect for both the NYSE and non NYSE series, meaning they move away

[69] The results of testing Equation 7 are provided separately in Table 23 and Table 24.

from each other. For the other stocks though, the coefficient is either positive or not significant. In general we conclude hypothesis IV of no influence for dark pool share can not be rejected.

Fourthly, a coefficient for high volatility as external variable in the model (Equation 4) is for the majority of bid and offers of all shares negative, but not significant. Combining the volatility dummy with a lagged NYSE series (Equation 5), implying high volatility strengthens the effect of the lagged NYSE series, shows similar results. The influence of this combined lagged NYSE and VIX is consistently stronger for the non NYSE series. The negative relation can be explained as increasing volatility, leading to lower return in the dependent variables as return often slowly increases in periods of low volatility, to then decrease rapidly in times of exceptionally high volatility. In economic context this for instance means markets rise slowly over a longer period, to then fall rapidly in times of a crisis. As a conclusion we can not reject hypothesis V of no influence for volatility though the signs of the coefficients are as expected.

6.4. Biv GARCH

The results of the Biv GARCH estimation including the values for the parameters $\omega_{11}, \omega_{12}, \omega_{22}, \alpha_{11}, \alpha_{22}, \beta_{11}, \beta_{22}$ and and μ_1 and μ_2 can be seen in Table 25 for IBM, Table 26 for KO, Table 27 for WMT and Table 28 for BAC. We neglect the unconditional means μ_1 and μ_2 in the Biv GARCH model as these long term mean estimates are approximately zero for the return series, and equal to an average of the ln(stock price) when considering the quote series, when using daily observations (Van der Wal, 2008). From the p-values for all other parameters in all models we may conclude that all parameters are generally statistically significant. Covariance between several combinations of the NYSE and non NYSE series is clearly not constant and the hypothesis VI is rejected. Furthermore it follows that both values of matrix A, being α_{11}, α_{22}, (see Section 9.3) do not equal zero, confirming general expectations that variance is deducted from return (see Footnote 68). Next, as both values in matrix A are positive, return in equal direction (positive or negative) has positive influence on the covariance, meaning the series indeed move together. The coefficients for the lagged variance variables β_{11}, β_{22} in matrix B (see Section 9.3), are positive as well and make the lagged variance contribute most to the observed variance and covariance. Some differences in coefficients A and B can be observed. In general the models on the residuals of the VEC M show a weaker influence of matrix A, return, in determining the covariance. Consistent with an often applied 'rule of thumb' by academics, the values for matrix B are higher in those models, implying the lagged variance has a stronger influence in determining the covariance. In the model related directly to the quotes the starting values in matrix W (see Section 9.3) and the return coefficients in matrix A are more important. Further studying of the Biv GARCH models shows the maximum likelihood parameters resulting from estimation based on the VEC M residuals are highest, but as the models are not simply extensions of each other we can not conclude anything from these estimations.

Finally the covariance $h^2_{XY_t}$ found in the results of the Biv GARCH model can be regressed on the series containing observations on the growth of dark pool share in consolidated US equity volume. As correlation as an indicator for the relation between 2 variables is more obvious we repeat the test for correlation as dependent variable. The results of the simple regressions can be found in Table 29, Table 30, Table 31 and Table 32. Except for tests on the covariance for ticker BAC, in general we find a significant to highly significant, negative relation between covariance and correlation on one side, and the share dark pools have in consolidated US equity volume on the other. The economic significance of this relationship becomes clear when considering an example. In Table 29 the coefficient for dark pool market share's impact on the correlation of NYSE and non NYSE return of $Dln\ B^{IBM}$ is estimated to be -0,017336. Assuming current dark pool share is 15% and increases with 5% to 15,75%, the impact on the estimated correlation would be -0,00234. As mentioned in Section 0, regulators are expected to allow a 15 to 20% market share by dark pools before strict regulation is introduced. According to the example this implies an effect of -0,01565 on correlation for IBM bid series. R^2 is between 0 an 8% for all tests, but in time series (simple) regressions it is rather common to not find a good fit. Regarding the hypothesis VI we may conclude indeed there seem to be indications of a relation between dark pools and covariance or correlation between public markets as an indication of price discovery.

The results of a simple regression regarding hypothesis VII are shown in Table 33. As the dark pool trading volume data is originally only available per month, we estimate the relation on that frequency. The coefficients are positive and not significant, R_2 is near zero and so the hypothesis that volatility is negatively related to dark pool market share should be rejected. The coefficient of the test directly on the values of the series confirms the expected negative relationship, but coefficients are not significant.

Table 2: This table contains an overview of the hypothesis to be tested as introduced in Section 5.7. Extensive analysis of the results can be found in Section 6.

#	Hypothesis	Expectation	Way of testing	Result	Statistics available
I)	H_0=The quotes of the NYSE and non NYSE series are not cointegrated (number of cointegrating relations=r) H_1= H_0 does not hold (number of cointegrating relations=r+1)	Reject H_0	Johansen test	Reject H_0	Table 3, 6, 9, 12
II)	H_0=Cointegration of sample 'pre dark pools' is stronger than of sample 'after dark pools' H_1= H_0 does not hold	Do not reject H_0	Johansen test	Reject H_0	Table 4, 5, 7, 8, 10, 11, 13, 14
III)	H_0=Applying the VEC M the NYSE series will influence the non NYSE more than visa versa H_1= H_0 does not hold	Do not reject H_0	VEC M	Do not reject H_0 as nonNYSE moves to NYSE, Reject H_0 as lags NYSE have no effect on nonNYSE	Tables 15-24
IV)	H_0=The % DPs claim of US cons.eq. volume is not significant in explaining changes of either the NYSE or non NYSE series in the VEC M H_1= H_0 does not hold	Reject H_0	VEC M	Do not reject H_0	Tables 15-24
V)	H_0=A dummy for the VIX being higher than 3 quartiles is not significant in explaining changes of either the NYSE or non NYSE series in the VEC M H_1= H_0 does not hold	Reject H_0	VEC M	Do not reject H_0	Tables 15-24
VI)	H_0= Covariance between the NYSE and non NYSE is constant over time H_1= H_0 does not hold	Reject H_0	BV GARCH	Reject H_0	Tables 25-28
VII)	H_0= Covariance between the NYSE and non NYSE is negatively related to the % DPs claim of US cons.eq.volume H_1= H_0 does not hold	Do not reject H_0	Simple OLS regression	Do not reject H_0	Tables 29-32
VIII)	H_0=The VIX is not related to the % DPs claim of US cons.eq.volume H_1= H_0 does not hold	Reject H_0	Simple OLS regression	Do not reject H_0	Table 33

7. Summary and Conclusions

Dark liquidity has been a part of the stock trading market for a long time but did not become industrialized and successful until a few years ago. In the early 2000's regulators in the US and EU both advocated more competition to lower transaction prices. As a result, alternative trading venues, among others dark liquidity pools, flourished and gained market share from incumbent exchanges. Financial innovation gave an extra impulse and dark pool market share reached some 13% in the US and 3% in the EU, of total consolidated equity trading by 2010. The traditional objective of a dark pool, protecting block orders from price impact, allows dark pools to release limited or delayed trading information. Therefore, dark liquidity gives rise to information asymmetry as investors in primary, public markets have less access to information than investors in a dark pool. Information represents investors' latent demand and supply (liquidity) and is therefore the main driver of the price discovery process. Many studies have shown the price discovery process mainly takes place in primary markets (See Section 2.12). As liquidity drains away to dark pools in which trading information is only limitedly released, the quality of the price discovery process is likely to be influenced when dark pools have gained a substantial share of the market. The main research question in this thesis is whether US dark pool development is indeed related to a possible changing price discovery in the US markets.

Assuming cointegration, covariance and correlation are indicators for the quality of the price discovery process in the market, we test whether indicators for volatility and dark pool trading volume are related to that cointegration in the US market by looking at daily closing quotes for 4 large volume Dow Jones stocks. Following Hasbrouck (1995), we gather data for all 9 primary US stock exchanges and create 1 series containing NYSE quotes and 1 series containing quotes from all other exchanges (non NYSE) over the sample 01/2005 – 06/2010.

Presuming a unit root in the quote series we test for cointegration using a Johansen Test (See Section 6.2). We expect the NYSE and non NYSE series are very closely related as no large arbitrage possibilities between markets can exist and indeed find cointegration, confirming hypothesis I. Following Hasbrouck (1995) a changing relation between the 2 price series might indicate changing price discovery. Therefore we consider whether cointegration in 2 samples of each 2 years ('before dark pools' and 'after dark pools') differs. Strong influence of dark liquidity might result in hampered cointegration as a rough indication for less price discovery. We find less clear cointegration present in both samples. Presumably the samples become too short and therefore cointegration is not well observable anymore. We reject hypothesis II of changing cointegration between the two samples.

We then apply a VEC model to both get an indication of whether price discovery happens primarily on the NYSE, or in the non NYSE markets (See Section 6.3). We also test for a possible influence dark pool development and volatility might have on the NYSE and non NYSE series. We find the Error Correction Term (ECT) strongly influences the non NYSE series towards the NYSE series whereas the opposite effect does not seem to be present. The lagged values of the NYSE series do not have clear significant impact on the non NYSE series though. From the first perspective our hypothesis III of stronger price discovery in the NYSE can not be rejected, whereas from the second perspective it should. Furthermore, coefficients for external variables for dark pools' market share and a dummy for high volatility are consistently negatively related to return of the NYSE and non NYSE, but results are not significant. Applying the dark pool market share in an ECT shows significantly negative results only in case of some tickers. Combining the volatility dummy with the lagged values of the NYSE series gives again negative, but not significant results. Negative and significant parameters would confirm our expectation of both dark pool share in US consolidated equity volume and volatility having a negative effect on the NYSE and non NYSE return series. We can not reject hypothesis IV and V of no explanatory power for these parameters though.

Estimating a Bivariate GARCH(1,1) model we directly devise the covariance and correlation (See Section 6.4). We expect the correlation (and covariance) between the NYSE and non NYSE series to decrease between 2005 and 2010. This might indicate decreased price discovery as markets would 'react' less instantly to each other. We find the Biv GARCH parameters have values consistent with literature and are significant. The covariance and correlation are subsequently tested for being linearly related to the development of dark pools' market share over time. The results show a consistently negative relation between the 2 series as expected, that in many cases is also significant. Regarding our hypothesis IIV we may therefore conclude there does seem to be a relation between covariance and correlation on the one hand, and dark pool share in trading volume on the other. Ultimately, we do not find any relation between volatility and dark pool market share and so can not reject hypothesis VIII.

Our results have a number of important implications for both academics as well as practitioners. Our study provides a clear overview of dark pools in a Market Microstructure environment and attempts to include all common definitions and aspects. Though we found many studies to individual dark pools have been performed, only a few report on the industry as a whole, and none perform empirical research of the dark industry on markets in general. Using basic data as indicators we do find some important evidence. Firstly, consistent with literature, NYSE and non NYSE quotes seem to be strongly cointegrated over time as long as the sample is large enough. Secondly, the NYSE still seems to lead the price discovery process before the non NYSE trading venues. Third and most importantly, we seem to find basic evidence of a negative relationship between dark pools' presence in the markets and the correlation of other primary markets. We therefore emphasize the need for obtaining more insights on the effect dark liquidity has on the market microstructure environment.

As for recommendations for further research, we firstly recognize phenomenon's such as cointegration are only revealed over long samples. Extending the dataset is likely to easier reveal changing cointegration and correlation between the NYSE and non NYSE series over time. Secondly, volatility is generally most informative on a daily basis. Therefore our regressions of dark pool trading volume and volatility could better be performed on a daily instead of monthly frequency. Thirdly, more observations on dark pool market share, especially before 2008, would increase the reliability of the regression tests.

8. References

References are provided firstly for academic literature and secondly for other literature.

Academic Literature:

Avgouleas, E. (2005). “A Critical Evaluation of the New EC Financial-Market Regulation: Peaks, Troughs and the Road Ahead”. University of Piraeus. Working paper.

Bagehot, W. (1971). “The Only Game in Town”. *Financial Analysts Journal*. Vol 22. Pages 12-14.

Berenguer-Rico, V. and J.L. Carrion-i-Silvestre. (2007). “Multicointegration, Polynomial Cointegration and I(2) Cointegration with Structural Brakes. An Application to the Sustainability to the US External Deficit”. University of Barcelona, Research Institute of Applied Economics. Working paper.

Bikker, J.A. and L. Spierdijk, P.J. Van der Sluis. (2007). “Market Impact Costs of Institutional Equity Trades”. *Journal of International Money and Finance*, Vol 26, No 6. Pages 974-1000

Brandes, Y. and I. Domowitz (2010). “Alternative Trading Systems in Europe – Trading Performance by European Venues Post-MiFID”. Investment Technology Group.

Brandt, M.W. and K.A. Kavajecz. (2004). “Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve”. The Journal of Finance, Vol 59, No. 6.

Brooks, C. (2002). “Introductory Econometrics for Finance”. *Cambridge University Press*, sixth reprint, 2005.

Buti, S. and B. Rindi, I.M. Werner. (2010). “Dynamic Dark Pool Trading Strategies in Limit Order Markets”. Working paper.

Carrie, C. (2008). “Illuminating the dark influence on trading and U.S. market structure”, The Journal of Trading, Vol. 3, Nr. 1. Pages 40-55

Conrad, J. and K.M. Johnson, S. Wahal. (2003). “Insitutional Trading and Alternative trading Systems”. *Journal of Financial Economics*, Vol 70, No 1. Pages 99-134.

Degryse, H. (10-2008). “MiFID: Competitie op Financiële Markten en Financieel Toezicht”. Katholieke Universiteit Leuven.

Degryse, H. and M. Van Achter, G. Wuyts. (2008). “Shedding Light on Dark Liquidity Pools”. TILEC Dicussion Paper.

Easly, D. and N.M. Kiefer, M. O’Hara. (1997). “The Information Content of the Trading Process”. *Journal of Empirical Finance*, 4, p. 159 – 186.

Ende, B. and J. Muntermann (2010). “Opacity and Exclusivity in Electronic Securities Trading: The Case of Dark Pools – A Comparison of Liquidnet with Electronic Exchanges” *Informationssysteme in der Finanzwirtschaft*.

Engle, R.F. (1982). “Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation”. *Econometrica*, Vol 50. Pages 987-1008.

Foucault , T. and A.J. Menkveld (2006). "Competition for Order Flow and Smart Order Routing Systems". *Journal of Finance*, 63, p.119-158.

Ganchev, K and M. Kearns, Y. Nevmyvaka, J. Wortman, (2010). "Censored Exploration and the Dark Pool Problem". Computer and Information Science, University of Pennsylvania.

Hasbrouck, J. (1995). "One Security, Many Markets: Determining the Contributions to Price Discovery"; J. Hasbrouck; *The Journal of Finance*, Vol 50, No 4.

Hoffman, P. (2010). "Adverse Selection, Transaction Fees and Multi-Market Trading". Job Market Paper.

Hooper, M. and R.A. Schwartz (2008). "The Limits of Liquidity". *Journal of Trading*, Summer 2008, Volume 3, Number 3.

Johansen, S. (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models". *Econometrica*, Vol 59. Pages 1551-1580.

Kalay, A. and L. Wei, A. Wohl. (2002). "Continuous Trading or Call Auctions". *The Journal of Finance*, Vol 57, No 1.

Keim, D. and A. Madhavan. (1996). "The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects". *The Review of Financial Studies*, Vol 9, No 1.

Loeb, T.F. (1983). "Trading cost: the critical link between investment information and results". *Financial Analysts Journal*, Vol 39. Pages 39-43.

Madhavan, A (2000). "Market Microstructure: A Survey". University of Southern California. Working paper.

Madhavan, A. (1995). "Consolidation, fragmentation, and the disclosure of trading Information". *Review of Financial Studies*, Vol 8. Pages 579-603.

Menkveld , A.J. and S.J. Koopman, A. Lucas (2007). "Modeling Round-the-Clock Price Discovery for Cross-Listed Stocks using State Space Methods". *The Journal of Business and Economic Statistics*.

Mittal, H. (2008). "Are You Playing in a Toxic Dark Pool? – A Guide to Preventing Information Leakage". Investment Technology Group (ITG) INC.

Owens, A.E. and S.J. Yim, A. Stekler, C.L. March. (2010). "Proposed national market system changes aim to lighten dark pools". *Journal of Investment Compliance*, Vol 11, No 1.

Petrella, G. (2009). "MiFID, Reg NMS, and Competition Across Trading Venues in Europe and United States". Working paper. 14th of September, 2009.

Popescu , M. (2007). "Two Essays on the Probability of Informed Trading". PhD publication.

Ready, M.J. (2009). "Determinants of Volume in Dark Pools". University of Wisconsin-Madison. Working paper.

Saraiya, N. and H. Mittal. (2009). "Understanding and Avoiding Adverse Selection in Dark Pools". Investment Technology Group. Working paper.

Schwartz , R.A. (2010). “Dark Pools, Fragmented Markets and the Quality of Price Discovery”. *The Journal of Trading*, Spring 2010.

Schwartz, R.A. (13-09-2009). “Dark Pools and Fragmented Markets”. Speech Annual Meeting of the World Federation of Exchanges.

Sofianos, G. (2007). “Dark Pools and Algorithmic Trading”. Algorithmic Trading, A Buy Side Handbook, 2nd edition. Goldman Sachs

Takayama, S. (2009). “The Dynamic Strategy of an Informed Trader with Market Manipulation”. *Annals of Finance*, Vol 6. Pages 287-294.

Van der Wall, M. (2008). “Modeling the Conditional Correlation Between United States Stock and Bond Returns”. Rijksuniversiteit Groningen. Faculty of Mathematics.

Other Literature:

Anonymous (2009). “The Consolidated Tape: Yes, the Dark Pools are in there”. Wall Street Journal, 19-11-2009

Anonymous. (2009). “SEC Issues Proposals to Shed Greater Light on Dark Pools”. U.S. Securities and Exchange Commission.

Anonymous. (undated) “MiFID Glossary”. Kasbank.

Appenzell, (22-10-2009). “Derivatives don’t need regulation?...Dark pool trade limits to be reduced 95%”. Industry journal.

Economist, the (2009). “Finance and Economics: Attack of the clones; The rise of dark pools”.

Haynes, A. (2008). “Europe Enters New Dark Age”. *Ftmandate.com*

Keegan, D. (2010). “Re: Concept Release on Equity Market Structure”. Citigroup. 5th of May, 2010.

Lin, G. and B. Valentine. (undated). “Alternative Liquidity Sources”. *The Trade Magazine, The Trade Asia*, No.3

McEachern Gibbs, C. (2007). “STA Panel: Bright Future for Dark Pools”. *Advanced trading.com*

McIntyre, H. (DATE). “ECN and ATS…The Electronic Future”. The Summit Group

Merrin, S and H. Meyerson, V. Khandros. (2010). “Re: Market Structure Round Table”. Liquidnet. 23rd of June, 2010.

Ortega, E. and J. Kearns, E. Martin. (2009). “High Frequency Traders say Speed Works for Everyone (Update 1)”. *Bloomberg.com*. 28th of July, 2009.

Quantitative Microstructure Software. (2007). “EViews 6 User Guides I and II”.

Plexus Group. (2004). “The Changing Role of the Buy-side Trader”.

Schack, J. and J. Gawronski. (2007). "Trading Talk – The Decline and Fall of Displayed Markets". Rosenberg Securities Inc. 02-08-2007

Schack, J. and J. Gawronski. (2007). "Trading Talk – The Decline and Fall of Displayed Markets". Rosenberg Securities Inc. 11-10-2007

Schack, J. and J. Gawronski. (2008-2010). "Trading Talk – Let There Be Light". Rosenberg Securities Inc. Volumes 3, 6, 10, 22, 30 and 31.

Schapiro, M.L. (2010). "Speech by SEC Chairman, Stenghtening Our Equity Market Structure". U.S. Securities and Exchange Commission. Speech.

SEC. (Sept 2010). "Findings Regarding the Market Events of May 6th 2010". *SEC and CFTC*.

Spicer, J. (2009). "Industry warns of consequences of dark pool reform". Reuters.

Tabb, L. (2006). "Dark is Hot". Wallstreet and Tech.com.

Tabb, L. (2008). "Search, Dark Pools and Disappearing Traders". O'Reilly Money: Tech Conference. Presentation.

TAQ 3 User's Guide. (Oct 2008). NYSE, Version 1.1.9.

The Daily Bell. (2009). "Derivatives don't need Regulation?". 22nd of October, 2009.

9. Appendices

9.1. Applying a Johansen cointegration test

The Johansen cointegration test basically proposes a VAR model (VAR M) should be turned into a VEC M by writing the VAR M in first differences. Brooks (2007, p.403) writes the VEC M of a VAR containing g variables as:

$$\text{VEC M:} \quad \Delta \mathrm{Y_t} = \Pi * \mathrm{Y_{t-k}} + \Gamma_1 * \Delta \mathrm{Y_{t-1}} + \Gamma_2 * \Delta \mathrm{Y_{t-2}} + \cdots + \Gamma_{\mathrm{k-1}} * \Delta \mathrm{Y_{t-(k-1)}} + u_t \qquad (8)$$

$$with\, \Pi = \left(\sum_{i-1}^{k} \beta_i \right) - I_g$$

Where in the long run changes in the dependent variable $\mathrm{Y_t}$ are determined by the long-run coefficient matrix Π. To see why, note all the other in differences expressed dependent variables $\Delta \mathrm{Y_{t-i}}$ and attached coefficient matrices Γ_i will equal zero in the long-run as $\mathrm{Y_t}$ reverts to its mean. Furthermore, the long-run expectation of the residual u_t is also zero. Π itself is a matrix of the summarized lung-run coefficients β_i of the VAR M.

The Johansen test than considers whether the rank of the matrix Π is different from zero, which would indicate cointegration. The rank of a matrix indicates the number of linearly independent rows or columns (Brooks, 2002 p. 663). Every independent row in a matrix adds 1 to the rank. For no unit root to be present all rows should be linearly dependent. In case of an independent trend in a row, the rank would be different from zero and the Johansen test would indicate cointegration. The Johansen test checks whether cointegration is present and estimates the number of cointegrating relations, the VEC M estimates a cointegrating system on the first differences of the variables.

9.2. The ARCH model

The ARCH models model the relation variance has with previous errors. It does no longer depend on a mean and residual sum of squares as with OLS. An ARCH(1) model can be represented by the equations (Brooks 2002, p.447):

$$Y_t = \beta_1 + \beta_n * X_{nt} + u_t \qquad (9)$$

Where the error term u_t is normally distributed with mean 0 and variance σ_t^2.

$$\sigma_t^2 = \alpha_0 + \alpha_1 * u_{t-1}^2 \qquad (10)$$

Where σ_t^2 denotes the variance at time t, α_0 and $\alpha_1$denote coefficients and u_{t-1}^2 represents the error one period back. Very often σ_t^2 is renamed h_t in the literature. As variance h_t is constructed from squaring the positive and negative errors it should by definition be positive. The model should always restrict h_t to be positive as any other results would not make sense.

9.3. The BEKK GARCH(1,1) model

Engle and Kroner slightly modified the GARCH model developed by Baba, Engle, Kraft and Kroner (1990) and introduced the BEKK GARCH(1,1) model in 1995. The GARCH model allowed the conditional variance h_t to rely on its previous own lags as well. The formula for h_t was changed to:

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 * u_{t-1}^2 + \beta * h_{t-1} \tag{11}$$

Now the conditional variance relies on a weighted formula of firstly α_0 being a constant that can be interpreted as a long term variance. Secondly, $\alpha_1 * u_{t-1}^2$ being volatility one period back and thirdly on $\beta * h_{t-1}$ as the observed conditional variance one lag back.

In order to jointly estimate $h_t$for both the NYSE and non NYSE series and their conditional covariance the BEKK GARCH(1,1) model can be represented as Van der Wal (2008) suggested:

$$H_t = {\omega^*}' \omega^* + A_1^{*\prime} h_{t-1} A_1^* + B_1^{*\prime} r_{t-1} {r'}_{t-1} B_1^* \tag{12}$$

$$H_t = \begin{pmatrix} var(X_t) & covar(Y_t X_t) \\ covar(X_t Y_t) & var(Y_t) \end{pmatrix} = \begin{pmatrix} h_{XX_t}^2 \\ h_{XY_t}^2 \\ h_{YY_t}^2 \end{pmatrix}$$

Where h_t denotes the conditional variance, ω^* denotes the upper triangular and A_1^* and B_1^* represent coefficient matrices containing α and β. The matrices r_{t-1} are the return matrices for X_t (here NYSE), Y_t (here non NYSE) and their combination. Note the covariance of X_t, Y_t is equal to taking the covariance of the reverse. Therefore the final vector H_t contains only three parameters. ω is a vector containing 4 parameters and $A_1$and $B_1$are 2 X 2 parameter matrices. As the matrix ω contains 3 unique values the total number of parameters to be estimated equals 3+4+4=11. The three variance equations from formula 12 can be directly estimated by the BEKK GARCH(1,1) model and are as follows:

$$h_{XX_t}^2 = \omega_{11}^2 + \alpha_{11}^2 * r_{X,t-1}^2 + \beta_{11}^2 * h_{X,t-1}^2 \tag{13}$$

$$h_{XY_t}^2 = \omega_{12} * \omega_{11} + \alpha_{11} * \alpha_{22} * r_{X,t-1} * r_{Y,t-1} + \beta_{11} * \beta_{22} * h_{XY_t,t-1}^2$$

$$h_{YY_t}^2 = \omega_{22}^2 + \omega_{12}^2 + \alpha_{22}^2 * r_{Y,t-1}^2 + \beta_{22}^2 * h_{Y,t-1}^2$$

Tables 25-28 show the estimated value for the parameters $\omega_{11}, \omega_{12}, \omega_{22}, \alpha_{11}, \alpha_{22}, \beta_{11}, \beta_{22}$ and μ_1 and μ_2, the two unconditional means of $r_X$andr_Y. These latter parameters can be seen as the long run return for the series X and Y, and can be considered more or less equal to zero as all data here is on a daily basis.

The GARCH(1,1) model can be generalized to a GARCH(p,q) model. Considering the iterative proves of passed squared errors influencing the current conditional variance as modeled by the GARCH(1,1) a higher order of lags is rarely necessary in finance (Brooks 2007, p.455). A GARCH model is estimated by maximum likelihood as estimating the relationship most likely given the data. OLS can not be employed estimating a GARCH model as OLS minimizes the sum of the squared errors (RSS), finding a linear relationship around a constant mean. The RSS does not separately model the variance and only minimizes the errors around a mean.

9.4. Eviews output

Table 3: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 06-2010			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{IBM}_{NYSE}$ & $\ln B^{IBM}_{nonNYSE}$	1138	Number of cointegrating relations by model:					
		Trace	1	1	1	1	2
		Max-Eig	1	1	1	1	2
		Log likelihood by number of relations and model:					
		0	7991.320	7991.320	7991.469	7991.469	7991.873
		1	8049.253	8049.268	8049.415	8051.345	8051.611
		2	8049.379	8049.948	8049.948	8054.511	8054.511
$\ln O^{IBM}_{NYSE}$ & $\ln O^{IBM}_{nonNYSE}$	1029	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	7267.520	7267.520	7267.918	7267.918	7268.524
		1	7315.750	7315.858	7316.215	7316.991	7317.570
		2	7316.092	7316.310	7316.310	7319.401	7319.401
$\text{Dln } B^{IBM}_{NYSE}$ & $\text{Dln } B^{IBM}_{nonNYSE}$	1106	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	7434.289	7434.289	7434.337	7434.337	7434.411
		1	7664.380	7664.390	7664.431	7664.589	7664.590
		2	7775.375	7775.465	7775.465	7775.986	7775.986
$\text{Dln } O^{IBM}_{NYSE}$ & $\text{Dln } O^{IBM}_{nonNYSE}$	991	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	6681.643	6681.643	6681.748	6681.748	6681.831
		1	6898.676	6898.707	6898.812	6898.827	6898.835
		2	7011.597	7011.797	7011.797	7012.660	7012.660

Table 4: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 12-2006			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{IBM}_{NYSE}$ & $\ln B^{IBM}_{nonNYSE}$	398	Number of cointegrating relations by model:					
		Trace	1	1	1	0	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	2898.236	2898.236	2898.355	2898.355	2900.309
		1	2918.205	2918.267	2918.283	2918.487	2920.267
		2	2918.224	2919.540	2919.540	2920.925	2920.925
$\ln O^{IBM}_{NYSE}$ & $\ln O^{IBM}_{nonNYSE}$	372	Number of cointegrating relations by model:					
		Trace	1	0	0	0	0
		Max-Eig	1	0	1	0	0
		Log likelihood by number of relations and model:					
		0	2695.824	2695.824	2695.841	2695.841	2697.179
		1	2712.518	2712.764	2712.779	2713.044	2714.305
		2	2712.535	2712.944	2712.944	2714.361	2714.361
$\text{Dln } B^{IBM}_{NYSE}$ & $\text{Dln } B^{IBM}_{nonNYSE}$	384	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2691.257	2691.257	2691.303	2691.303	2691.312
		1	2771.242	2771.329	2771.357	2771.423	2771.425
		2	2805.090	2805.277	2805.277	2807.425	2807.425
$\text{Dln } O^{IBM}_{NYSE}$ & $\text{Dln } O^{IBM}_{nonNYSE}$	354	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2467.356	2467.356	2467.397	2467.397	2467.434
		1	2540.412	2540.412	2540.445	2540.447	2540.447
		2	2572.092	2572.219	2572.219	2573.726	2573.726

Table 5: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2008 – 12-2009			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{IBM}_{NYSE}$ & $\ln B^{IBM}_{nonNYSE}$	442	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	3014.771	3014.771	3015.124	3015.124	3015.158
		1	3037.438	3041.648	3041.724	3041.901	3041.925
		2	3037.522	3042.698	3042.698	3042.908	3042.908
$\ln O^{IBM}_{NYSE}$ & $\ln O^{IBM}_{nonNYSE}$	439	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	3048.920	3048.920	3049.891	3049.891	3049.916
		1	3079.464	3081.339	3082.285	3083.343	3083.343
		2	3080.374	3082.907	3082.907	3083.966	3083.966
$\mathrm{Dln}\, B^{IBM}_{NYSE}$ & $\mathrm{Dln}\, B^{IBM}_{nonNYSE}$	433	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2815.804	2815.804	2815.827	2815.827	2815.833
		1	2907.140	2907.474	2907.497	2907.546	2907.552
		2	2952.638	2953.120	2953.120	2953.195	2953.195
$\mathrm{Dln}\, O^{IBM}_{NYSE}$ & $\mathrm{Dln}\, O^{IBM}_{nonNYSE}$	431	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2840.115	2840.115	2840.182	2840.182	2840.184
		1	2952.009	2952.031	2952.072	2952.127	2952.127
		2	3006.535	3007.441	3007.441	3007.504	3007.504

Table 6: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 06-2010			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{KO}_{NYSE}$ & $\ln B^{KO}_{nonNYSE}$	988	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	6867.266	6867.266	6867.369	6867.369	6867.552
		1	6917.523	6920.084	6920.132	6921.439	6921.481
		2	6917.553	6921.525	6921.525	6922.955	6922.955
$\ln O^{KO}_{NYSE}$ & $\ln O^{KO}_{nonNYSE}$	920	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	6591.517	6591.517	6591.781	6591.781	6591.943
		1	6626.359	6626.699	6626.881	6627.939	6628.018
		2	6626.518	6627.479	6627.479	6628.540	6628.540
$\mathrm{Dln}\, B^{KO}_{NYSE}$ & $\mathrm{Dln}\, B^{KO}_{nonNYSE}$	941	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	6238.484	6238.484	6238.567	6238.567	6238.791
		1	6444.788	6444.993	6445.011	6445.520	6445.527
		2	6545.820	6546.038	6546.038	6546.550	6546.550
$\mathrm{Dln}\, O^{KO}_{NYSE}$ & $\mathrm{Dln}\, O^{KO}_{nonNYSE}$	864	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	5917.735	5917.735	5917.745	5917.745	5917.943
		1	6148.158	6148.260	6148.266	6148.320	6148.343
		2	6236.486	6236.688	6236.688	6236.774	6236.774

Table 7: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 12-2006

Ticker and Series	N		Data Trend: None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{KO}_{NYSE}$ & $\ln B^{KO}_{nonNYSE}$	242	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	1777.622	1777.622	1777.732	1777.732	1777.938
		1	1792.444	1794.017	1794.018	1794.754	1794.959
		2	1792.446	1794.237	1794.237	1795.038	1795.038
$\ln O^{KO}_{NYSE}$ & $\ln O^{KO}_{nonNYSE}$	202	Number of cointegrating relations by model:					
		Trace	1	0	1	0	0
		Max-Eig	1	1	1	0	1
		Log likelihood by number of relations and model:					
		0	1629.753	1629.753	1629.937	1629.937	1630.800
		1	1637.111	1637.974	1638.061	1639.411	1639.755
		2	1637.200	1638.099	1638.099	1639.791	1639.791
$D\ln B^{KO}_{NYSE}$ & $D\ln B^{KO}_{nonNYSE}$	214	Number of cointegrating relations by model:					
		Trace	2	2	2	1	2
		Max-Eig	2	2	2	1	2
		Log likelihood by number of relations and model:					
		0	1499.277	1499.277	1499.585	1499.585	1499.622
		1	1542.875	1543.315	1543.449	1543.563	1543.566
		2	1567.306	1567.814	1567.814	1567.962	1567.962
$D\ln O^{KO}_{NYSE}$ & $D\ln O^{KO}_{nonNYSE}$	171	Number of cointegrating relations by model:					
		Trace	2	1	2	1	2
		Max-Eig	2	1	2	1	2
		Log likelihood by number of relations and model:					
		0	1321.450	1321.450	1321.513	1321.513	1321.897
		1	1356.041	1356.113	1356.149	1356.503	1356.624
		2	1375.261	1375.409	1375.409	1375.913	1375.913

Table 8: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2008 – 12-2009			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{KO}_{NYSE}$ & $\ln B^{KO}_{nonNYSE}$	474	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	3395.070	3395.070	3395.098	3395.098	3396.634
		1	3426.483	3427.762	3427.784	3428.234	3429.765
		2	3426.494	3429.094	3429.094	3430.214	3430.214
$\ln O^{KO}_{NYSE}$ & $\ln O^{KO}_{nonNYSE}$	456	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	3223.232	3223.232	3223.306	3223.306	3224.363
		1	3239.680	3239.804	3239.812	3240.556	3241.523
		2	3239.683	3240.791	3240.791	3241.909	3241.909
$\mathrm{Dln}\, B^{KO}_{NYSE}$ & $\mathrm{Dln}\, B^{KO}_{nonNYSE}$	469	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	3205.651	3205.651	3205.730	3205.730	3205.737
		1	3317.982	3317.982	3318.052	3318.052	3318.059
		2	3366.391	3366.392	3366.392	3367.949	3367.949
$\mathrm{Dln}\, O^{KO}_{NYSE}$ & $\mathrm{Dln}\, O^{KO}_{nonNYSE}$	448	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	3036.410	3036.410	3036.455	3036.455	3036.585
		1	3160.776	3160.776	3160.784	3160.796	3160.856
		2	3207.254	3207.267	3207.267	3208.177	3208.177

Table 9: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 06-2010			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B_{NYSE}^{WMT}$ & $\ln B_{nonNYSE}^{WMT}$	1123	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	7930.830	7930.830	7930.981	7930.981	7930.987
		1	7990.475	7992.910	7992.944	7993.085	7993.090
		2	7990.535	7997.043	7997.043	7998.279	7998.279
$\ln O_{NYSE}^{WMT}$ & $\ln O_{nonNYSE}^{WMT}$	1084	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	7645.266	7645.266	7645.374	7645.374	7645.433
		1	7711.220	7711.463	7711.464	7711.545	7711.597
		2	7711.228	7715.993	7715.993	7716.890	7716.890
$D\ln B_{NYSE}^{WMT}$ & $D\ln B_{nonNYSE}^{WMT}$	1102	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	7388.992	7388.992	7389.029	7389.029	7389.154
		1	7686.406	7686.526	7686.538	7686.550	7686.587
		2	7805.229	7805.428	7805.428	7805.445	7805.445
$D\ln O_{NYSE}^{WMT}$ & $D\ln O_{nonNYSE}^{WMT}$	1058	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	7163.880	7163.880	7163.983	7163.983	7164.176
		1	7367.803	7368.000	7368.010	7368.234	7368.252
		2	7484.455	7484.658	7484.658	7484.921	7484.921

Table 10: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 12-2006			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{WMT}_{NYSE}$ & $\ln B^{WMT}_{nonNYSE}$	410	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	3035.014	3035.014	3035.067	3035.067	3035.268
		1	3062.636	3062.920	3062.948	3074.985	3075.174
		2	3062.671	3067.131	3067.131	3079.389	3079.389
$\ln O^{WMT}_{NYSE}$ & $\ln O^{WMT}_{nonNYSE}$	372	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	2829.490	2829.490	2829.545	2829.545	2829.647
		1	2860.614	2864.563	2864.564	2873.578	2873.663
		2	2860.615	2868.347	2868.347	2878.148	2878.148
$\text{Dln } B^{WMT}_{NYSE}$ & $\text{Dln } B^{WMT}_{nonNYSE}$	401	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2843.438	2843.438	2843.604	2843.604	2843.924
		1	2953.946	2954.020	2954.118	2954.149	2954.203
		2	2982.584	2982.670	2982.670	2982.747	2982.747
$\text{Dln } O^{WMT}_{NYSE}$ & $\text{Dln } O^{WMT}_{nonNYSE}$	358	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2642.718	2642.718	2643.005	2643.005	2643.409
		1	2752.119	2752.621	2752.647	2753.068	2753.072
		2	2777.616	2778.121	2778.121	2778.597	2778.597

Table 11: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2008 – 12-2009			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{WMT}_{NYSE}$ & $\ln B^{WMT}_{nonNYSE}$	469	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	3312.373	3312.373	3312.505	3312.505	3312.639
		1	3352.086	3352.985	3353.020	3353.341	3353.428
		2	3352.107	3355.974	3355.974	3357.164	3357.164
$\ln O^{WMT}_{NYSE}$ & $\ln O^{WMT}_{nonNYSE}$	475	Number of cointegrating relations by model:					
		Trace	1	1	2	1	2
		Max-Eig	1	1	2	1	2
		Log likelihood by number of relations and model:					
		0	3397.074	3397.074	3397.177	3397.177	3397.396
		1	3440.258	3440.521	3440.580	3440.799	3440.975
		2	3440.301	3443.652	3443.652	3445.039	3445.039
$\mathrm{Dln}\ B^{WMT}_{NYSE}$ & $\mathrm{Dln}\ B^{WMT}_{nonNYSE}$	463	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	3106.590	3106.590	3106.603	3106.603	3106.629
		1	3217.285	3217.429	3217.443	3217.492	3217.514
		2	3275.560	3275.728	3275.728	3275.948	3275.948
$\mathrm{Dln}\ O^{WMT}_{NYSE}$ & $\mathrm{Dln}\ O^{WMT}_{nonNYSE}$	470	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	3197.859	3197.859	3197.883	3197.883	3197.895
		1	3305.145	3305.200	3305.212	3305.270	3305.282
		2	3365.851	3365.958	3365.958	3366.249	3366.249

Table 12: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 06-2010			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{BAC}_{NYSE}$ & $\ln B^{BAC}_{nonNYSE}$	1148	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	5512.934	5512.934	5513.202	5513.202	5513.331
		1	5583.971	5599.213	5599.476	5604.020	5604.141
		2	5584.362	5599.870	5599.870	5605.933	5605.933
$\ln O^{BAC}_{NYSE}$ & $\ln O^{BAC}_{nonNYSE}$	1057	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	5608.766	5608.766	5608.995	5608.995	5609.106
		1	5664.872	5678.108	5678.267	5685.572	5685.676
		2	5665.141	5678.651	5678.651	5687.340	5687.340
$\text{Dln } B^{BAC}_{NYSE}$ & $\text{Dln } B^{BAC}_{nonNYSE}$	1118	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	5002.241	5002.241	5002.250	5002.250	5002.260
		1	5241.666	5241.666	5241.673	5241.682	5241.691
		2	5358.139	5358.396	5358.396	5358.472	5358.472
$\text{Dln } O^{BAC}_{NYSE}$ & $\text{Dln } O^{BAC}_{nonNYSE}$	1018	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	5049.616	5049.616	5049.636	5049.636	5049.650
		1	5292.945	5292.953	5292.968	5293.000	5293.000
		2	5400.388	5400.604	5400.604	5400.661	5400.661

Table 13: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2005 – 12-2006

			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{BAC}_{NYSE}$ & $\ln B^{BAC}_{nonNYSE}$	346	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	2562.505	2562.505	2564.475	2564.475	2565.503
		1	2591.371	2592.140	2593.977	2594.337	2595.349
		2	2593.214	2593.986	2593.986	2596.423	2596.423
$\ln O^{BAC}_{NYSE}$ & $\ln O^{BAC}_{nonNYSE}$	325	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	2420.316	2420.316	2422.152	2422.152	2422.788
		1	2451.737	2456.546	2458.350	2458.637	2459.273
		2	2453.533	2458.353	2458.353	2460.374	2460.374
$\mathrm{Dln}\, B^{BAC}_{NYSE}$ & $\mathrm{Dln}\, B^{BAC}_{nonNYSE}$	326	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2323.086	2323.086	2323.127	2323.127	2323.658
		1	2399.139	2399.211	2399.248	2399.249	2399.773
		2	2420.269	2421.256	2421.256	2423.089	2423.089
$\mathrm{Dln}\, O^{BAC}_{NYSE}$ & $\mathrm{Dln}\, O^{BAC}_{nonNYSE}$	303	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	2167.818	2167.818	2167.857	2167.857	2168.073
		1	2252.709	2252.826	2252.826	2252.833	2253.045
		2	2272.748	2273.599	2273.599	2274.678	2274.678

Table 14: Tables 3-14 contain the results of a Johansen cointegration test for 4 randomly selected Dow Jones stocks traded on the NYSE. Samples are 01/2005 – 06/2010, 01/2005 – 12/2006 and 01/2008 – 12/2009. Series are firstly the ln of daily closing bids and offers in absolute value and secondly first differences of those same log(quotes). The tables show whether cointegrating relations are present on a significance level of 0,05.

Sample: 01-2008 – 12-2009			Data Trend				
Ticker and Series	N		None	None	Linear	Linear	Quadratic
		Test Type:	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
$\ln B^{BAC}_{NYSE}$ & $\ln B^{BAC}_{nonNYSE}$	453	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	1848.365	1848.365	1848.612	1848.612	1848.862
		1	1875.301	1887.698	1887.945	1889.675	1889.923
		2	1875.798	1889.006	1889.006	1890.775	1890.775
$\ln O^{BAC}_{NYSE}$ & $\ln O^{BAC}_{nonNYSE}$	453	Number of cointegrating relations by model:					
		Trace	1	1	1	1	1
		Max-Eig	1	1	1	1	1
		Log likelihood by number of relations and model:					
		0	2095.389	2095.389	2095.668	2095.668	2096.117
		1	2119.293	2131.134	2131.339	2131.585	2131.745
		2	2119.781	2132.422	2132.422	2132.773	2132.773
$\mathrm{Dln}\ B^{BAC}_{NYSE}$ & $\mathrm{Dln}\ B^{BAC}_{nonNYSE}$	448	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	1682.015	1682.015	1682.018	1682.018	1682.018
		1	1783.684	1783.693	1783.694	1783.695	1783.695
		2	1830.063	1830.251	1830.251	1830.500	1830.500
$\mathrm{Dln}\ O^{BAC}_{NYSE}$ & $\mathrm{Dln}\ O^{BAC}_{nonNYSE}$	448	Number of cointegrating relations by model:					
		Trace	2	2	2	2	2
		Max-Eig	2	2	2	2	2
		Log likelihood by number of relations and model:					
		0	1924.903	1924.903	1924.907	1924.907	1924.909
		1	2033.603	2033.603	2033.605	2033.648	2033.648
		2	2080.669	2080.892	2080.892	2081.290	2081.290

Table 15: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

IBM, BIDS	**Equation 2**		**Equation 3**		**Equation 4**		**Equation 5**		**Equation 6**	
Cointegrating Equation:										
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000		1.000000		1.000000		1.000000		1.000000	
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-1.000333		-0.997128		-1.000042		-1.000002		-0.999474	
	[-1121.46]		[-801.396]		[-1161.53]		[-1159.86]		[-622.846]	
C	-0.001470		-0.016965		-0.002808		-0.002993		-0.005888	
N	1210		556		1210		1210		556	
Error Correction:										
	Dependent Variable									
Independent Variable	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nonNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	0.193225	0.801985	0.232939	0.971881	0.241314	0.871648	0.244201	0.874511	0.444540	1.193938
	[1.06419]	[4.34426]	[0.57802]	[2.39175]	[1.31042]	[4.66021]	[1.32630]	[4.67645]	[1.09905]	[2.93317]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	-0.304768	-0.088546	-0.573552	-0.444360	-0.332525	-0.128859	-0.334129	-0.130428	-0.704301	-0.581093
	[-1.84498]	[-0.52721]	[-1.65004]	[-1.26783]	[-2.00218]	[-0.76389]	[-2.01219]	[-0.77336]	[-2.02388]	[-1.65929]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{NYSE}}$	-0.176448	-0.044431	-0.361235	-0.283603	-0.190845	-0.065346	-0.191619	-0.066078	-0.427421	-0.352469
	[-1.34469]	[-0.33303]	[-1.36556]	[-1.06326]	[-1.45168]	[-0.48938]	[-1.45784]	[-0.49498]	[-1.61660]	[-1.32470]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	0.259629	0.052929	0.500537	0.381101	0.285326	0.090257	0.286827	0.091724	0.627501	0.513712
	[1.59845]	[0.32050]	[1.45106]	[1.09571]	[1.74844]	[0.54454]	[1.75796]	[0.55351]	[1.81710]	[1.47820]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	0.157796	0.026474	0.314513	0.228467	0.170292	0.044633	0.170972	0.045272	0.377975	0.294267
	[1.22797]	[0.20263]	[1.19717]	[0.86247]	[1.32350]	[0.34152]	[1.32903]	[0.34650]	[1.43945]	[1.11359]
C	0.000294	0.000299	0.000959	0.002601	0.000731	0.000914	0.000761	0.000943	0.003696	0.005619
	[0.66675]	[0.66649]	[0.19091]	[0.51366]	[1.43201]	[1.76075]	[1.48973]	[1.81750]	[0.69951]	[1.05674]
% DPs	-	-	-0.007205	-0.026642	-	-	-	-	-0.025088	-0.043822
	-	-	[-0.12451]	[-0.45661]	-	-	-	-	[-0.43029]	[-0.74688]
VIX	-	-	-	-	-0.001757	-0.002469	-	-	-0.002387	-0.003063
	-	-	-	-	[-1.69555]	[-2.34600]	-	-	[-1.43973]	[-1.83595]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$*	-	-	-	-	-	-	-0.000405	-0.000558	-	-
VIX	-	-	-	-	-	-	[-1.81043]	[-2.45921]	-	-
R-squared	0.004458	0.029783	0.012100	0.018669	0.006788	0.034027	0.007113	0.034437	0.016527	0.026861

Table 16: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

IBM, OFFERS	**Equation 2**		**Equation 3**		**Equation 4**		**Equation 5**		**Equation 6**	
Cointegrating Equation:										
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000		1.000000		1.000000		1.000000		1.000000	
$\text{Dln O}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-1.000066		-1.001243		-1.000239		-1.000275		-0.997373	
	[-1104.29]		[-904.669]		[-1134.44]		[-1133.95]		[-706.172]	
C	0.003211		0.009009		0.004011		0.004176		-0.009269	
N	1109		545		1109		1109		545	
Error Correction:										
	Dependent Variable									
Independent Variable	$\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{nonNYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	-0.397424	0.199378	-1.131283	-0.376793	-0.414476	0.198403	-0.416909	0.196272	-1.497931	-0.709456
	[-2.04452]	[1.02156]	[-2.57395]	[-0.86274]	[-2.10912]	[1.00538]	[-2.12149]	[0.99452]	[-3.41098]	[-1.62088]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}$	0.306679	0.541807	0.976867	1.105996	0.317602	0.542381	0.319177	0.543742	1.212064	1.322556
	[1.72564]	[3.03643]	[2.57670]	[2.93583]	[1.77715]	[3.02220]	[1.78597]	[3.02965]	[3.21804]	[3.52304]
$\text{Dln O}(-2)^{\text{WMT}}_{\text{NYSE}}$	0.130681	0.251388	0.229131	0.277121	0.136323	0.251682	0.137109	0.252385	0.348276	0.387336
	[0.91840]	[1.75961]	[0.80183]	[0.97593]	[0.95568]	[1.75700]	[0.96125]	[1.76193]	[1.22612]	[1.36816]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-0.385377	-0.624696	-1.085062	-1.214360	-0.396666	-0.625288	-0.398270	-0.626678	-1.322892	-1.433343
	[-2.17342]	[-3.50895]	[-2.84353]	[-3.20258]	[-2.22373]	[-3.49072]	[-2.23273]	[-3.49833]	[-3.48995]	[-3.79387]
$\text{Dln O}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	-0.153865	-0.275016	-0.276082	-0.319421	-0.159893	-0.275343	-0.160721	-0.276082	-0.396762	-0.430600
	[-1.09572]	[-1.95059]	[-0.96770]	[-1.12672]	[-1.13538]	[-1.94700]	[-1.14133]	[-1.95224]	[-1.39848]	[-1.52278]
C	0.000520	0.000509	0.004666	0.003520	0.000696	0.000534	0.000721	0.000557	0.007226	0.005530
	[1.15219]	[1.12463]	[0.97196]	[0.73783]	[1.32651]	[1.01422]	[1.37537]	[1.05807]	[1.43366]	[1.10090]
% DPs	-	-	-0.044703	-0.031035	-	-	-	-	-0.054731	-0.041731
	-	-	[-0.80472]	[-0.56222]	-	-	-	-	[-0.98049]	[-0.75008]
VIX	-	-	-	-	-0.000689	-9.64E-05	-	-	-0.003474	-0.002238
	-	-	-	-	[-0.65809]	[-0.09173]	-	-	[-2.05639]	[-1.32919]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}*$	-	-	-	-	-	-	-0.000171	-4.04E-05	-	-
VIX	-	-	-	-	-	-	[-0.75399]	[-0.17763]	-	-
R-squared	0.010939	0.038940	0.030607	0.042055	0.011297	0.038962	0.011412	0.038985	0.041002	0.046584

Table 17: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

KO, BIDS	Equation 2	Equation 3	Equation 4	Equation 5	Equation 6
Cointegrating Equation:					
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	1.000000	1.000000	1.000000	1.000000	1.000000
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.994787 [-628.857]	-0.998426 [-669.845]	-0.994750 [-628.044]	-0.994767 [-628.402]	-1.002260 [-541.046]
C	-0.022780	-0.008111	-0.022922	-0.022857	0.006969
N	1091	584	1091	1091	584

Error Correction:

Dependent Variable

Independent Variable

	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$
ECT 1	-0.019386 [-0.13811]	0.629466 [4.23663]	0.068489 [0.19880]	0.806825 [2.28296]	-0.024187 [-0.17222]	0.625099 [4.20429]	-0.023650 [-0.16844]	0.625431 [4.20772]	0.210402 [0.60171]	0.983775 [2.74734]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	-0.214925 [-1.67756]	0.052459 [0.38683]	-0.296728 [-0.96813]	-0.167372 [-0.53233]	-0.212438 [-1.65785]	0.054733 [0.40346]	-0.212805 [-1.66097]	0.054470 [0.40159]	-0.391873 [-1.26479]	-0.287928 [-0.90747]
$\mathrm{Dln\ B}(-2)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	-0.112978 [-1.16324]	0.038741 [0.37684]	0.004650 [0.01997]	0.053145 [0.22247]	-0.111164 [-1.14437]	0.040394 [0.39280]	-0.111237 [-1.14526]	0.040391 [0.39282]	-0.042949 [-0.18381]	-0.005340 [-0.02232]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	0.078831 [0.63305]	-0.194202 [-1.47333]	0.153360 [0.50798]	0.019174 [0.06191]	0.076236 [0.61210]	-0.196582 [-1.49090]	0.076657 [0.61557]	-0.196269 [-1.48879]	0.248320 [0.81275]	0.140207 [0.44812]
$\mathrm{Dln\ B}(-2)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	0.047480 [0.50867]	-0.089815 [-0.90902]	-0.087271 [-0.38141]	-0.123935 [-0.52801]	0.045023 [0.48218]	-0.092057 [-0.93128]	0.045061 [0.48266]	-0.092109 [-0.93196]	-0.039733 [-0.17295]	-0.065002 [-0.27629]
C	6.13E-05 [0.15684]	7.66E-05 [0.18516]	0.001710 [0.43523]	0.002920 [0.72443]	0.000312 [0.66868]	0.000304 [0.61549]	0.000338 [0.72341]	0.000339 [0.68472]	0.002231 [0.52918]	0.004697 [1.08799]
% DPs	- -	- -	-0.022547 [-0.49898]	-0.036716 [-0.79210]	- -	- -	-	-	-0.027152 [-0.58694]	-0.050019 [-1.05585]
VIX	- -	- -	- -	- -	-0.000838 [-0.98185]	-0.000760 [-0.84142]	- -	- -	-0.000244 [-0.17476]	-0.001231 [-0.86260]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}*$ VIX	- -	- -	- -	- -	- -	- -	-0.000238 [-1.08224]	-0.000225 [-0.96823]	- -	- -
R-squared	0.022698	0.079179	0.027815	0.045056	0.023567	0.079759	0.023753	0.079963	0.028359	0.048900

Table 18: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

KO, OFFERS	**Equation 2**	**Equation 3**	**Equation 4**	**Equation 5**	**Equation 6**
Cointegrating Equation:					
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	1.000000	1.000000	1.000000	1.000000	1.000000
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.998489 [-720.501]	-0.998426 [-669.845]	-0.994750 [-586.173]	-0.998409 [-723.018]	-1.002260 [-541.046]
C	-0.003821	-0.008111	[-628.044]	-0.004130	0.006969
N	1043	545	1091	1043	584

Error Correction:

Independent Variable

	Dependent Variable									
	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$
ECT 1	-0.116704 [-0.58900]	0.479599 [2.42067]	0.068489 [0.19880]	0.806825 [2.28296]	-0.024187 [-0.17222]	0.625099 [4.20429]	-0.137947 [-0.69349]	0.462012 [2.32229]	0.210402 [0.60171]	0.983775 [2.74734]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	-0.049615 [-0.26708]	0.217697 [1.17195]	-0.296728 [-0.96813]	-0.167372 [-0.53233]	-0.212438 [-1.65785]	0.054733 [0.40346]	-0.038039 [-0.20454]	0.227367 [1.22237]	-0.391873 [-1.26479]	-0.287928 [-0.90747]
$\mathrm{Dln\ O}(-2)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	-0.090842 [-0.62517]	-0.005392 [-0.03711]	0.004650 [0.01997]	0.053145 [0.22247]	-0.111164 [-1.14437]	0.040394 [0.39280]	-0.086164 [-0.59288]	-0.001436 [-0.00988]	-0.042949 [-0.18381]	-0.005340 [-0.02232]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.047453 [-0.25653]	-0.320688 [-1.73376]	0.153360 [0.50798]	0.019174 [0.06191]	0.076236 [0.61210]	-0.196582 [-1.49090]	-0.059322 [-0.32031]	-0.330615 [-1.78492]	0.248320 [0.81275]	0.140207 [0.44812]
$\mathrm{Dln\ O}(-2)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	0.028076 [0.19625]	-0.056501 [-0.39497]	-0.087271 [-0.38141]	-0.123935 [-0.52801]	0.045023 [0.48218]	-0.092057 [-0.93128]	0.022716 [0.15874]	-0.061017 [-0.42632]	-0.039733 [-0.17295]	-0.065002 [-0.27629]
C	0.000220 [0.52332]	0.000323 [0.76743]	0.001710 [0.43523]	0.002920 [0.72443]	0.000312 [0.66868]	0.000304 [0.61549]	0.000522 [1.03453]	0.000567 [1.12407]	0.002231 [0.52918]	0.004697 [1.08799]
% DPs	- -	- -	-0.022547 [-0.49898]	-0.036716 [-0.79210]	- -	- -	`- -	- -	-0.027152 [-0.58694]	-0.050019 [-1.05585]
VIX	- -	- -	- -	- -	-0.000838 [-0.98185]	-0.000760 [-0.84142]	- -	- -	-0.000244 [-0.17476]	-0.001231 [-0.86260]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}*$ VIX	- -	- -	- -	- -	- -	- -	-0.000257 [-1.08466]	-0.000208 [-0.87768]	- -	- -
R-squared	0.013674	0.038577	0.027815	0.045056	0.023567	0.079759	0.014801	0.039260	0.028359	0.048900

Table 19: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

WMT, BIDS	Equation 2	Equation 3	Equation 4	Equation 5	Equation 6
Cointegrating Equation:					
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	1.000000	1.000000	1.000000	1.000000	1.000000
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.996414 [-601.263]	-0.997634 [-458.646]	-0.996507 [-586.173]	-0.996505 [-584.944]	-0.998496 [-424.943]
C	-0.016338	-0.011447	-0.015979	-0.015983	-0.008029
N	1172	506	1172	1172	506

Error Correction:

Dependent Variable (columns); Independent Variable (rows)

	Equation 2: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	Equation 2: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	Equation 3: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	Equation 3: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	Equation 4: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	Equation 4: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	Equation 5: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	Equation 5: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	Equation 6: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	Equation 6: $\mathrm{Dln\ B}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$
ECT 1	-0.329733 [-1.73595]	0.411290 [2.14641]	0.081245 [0.18318]	0.905822 [2.04400]	-0.315839 [-1.66540]	0.424937 [2.22091]	-0.316364 [-1.66793]	0.424425 [2.21794]	0.156122 [0.35384]	0.980379 [2.22555]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	0.366020 [2.20107]	0.513782 [3.06264]	0.431158 [1.14635]	0.488947 [1.30106]	0.354158 [2.13295]	0.502117 [2.99738]	0.354664 [2.13572]	0.502612 [2.99996]	0.366170 [0.97755]	0.423329 [1.13196]
$\mathrm{Dln\ B}(-2)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	0.081749 [0.63799]	0.146013 [1.12956]	0.286526 [1.00543]	0.339818 [1.19341]	0.077314 [0.60446]	0.141665 [1.09782]	0.077483 [0.60570]	0.141830 [1.09895]	0.257402 [0.90739]	0.310569 [1.09656]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.444566 [-2.69877]	-0.583517 [-3.51134]	-0.610331 [-1.64602]	-0.648919 [-1.75153]	-0.435928 [-2.65074]	-0.574991 [-3.46551]	-0.436258 [-2.65238]	-0.575314 [-3.46700]	-0.555916 [-1.50548]	-0.593902 [-1.61092]
$\mathrm{Dln\ B}(-2)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-0.152995 [-1.22949]	-0.204766 [-1.63115]	-0.378432 [-1.36210]	-0.421348 [-1.51781]	-0.151201 [-1.21732]	-0.202986 [-1.61984]	-0.151232 [-1.21740]	-0.203017 [-1.61986]	-0.358832 [-1.29743]	-0.401721 [-1.45483]
C	-0.000148 [-0.35323]	-0.000149 [-0.35300]	-0.012254 [-1.11496]	-0.014443 [-1.31515]	0.000434 [0.89149]	0.000424 [0.86250]	0.000416 [0.85329]	0.000406 [0.82561]	-0.012272 [-1.12190]	-0.014437 [-1.32189]
% DPs	- -	- -	-0.004845 [-1.10778]	-0.005721 [-1.30909]	- -	- -	`- -	- -	-0.005668 [-1.29819]	-0.006569 [-1.50712]
VIX	- -	- -	- -	- -	-0.002198 [-2.31954]	-0.002163 [-2.26278]	- -	- -	-0.003684 [-2.36949]	-0.003843 [-2.47614]
$\mathrm{Dln\ B}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}*$ VIX	- -	- -	- -	- -	- -	- -	-0.000542 [-2.24576]	-0.000533 [-2.19154]	- -	- -
R-squared	0.017761	0.061082	0.052305	0.086851	0.022250	0.065220	0.021969	0.064967	0.062931	0.098491

Table 20: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

WMT, OFFERS	**Equation 2**	**Equation 3**	**Equation 4**	**Equation 5**	**Equation 6**
Cointegrating Equation:					
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000	1.000000	1.000000	1.000000	1.000000
$\text{Dln O}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-1.001826 [-441.037]	-1.000411 [-474.946]	-1.002105 [-432.335]	-0.997886 [-431.679]	-0.998325 [-451.259]
C	0.009371	0.003683	0.010464	-0.010492	-0.004596
N	1144	510	1144	1144	510

Error Correction:

Dependent Variable (columns); Independent Variable (rows)

	Eq. 2: $\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	Eq. 2: $\text{Dln O}^{\text{WMT}}_{\text{nonNYSE}}$	Eq. 3: $\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	Eq. 3: $\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	Eq. 4: $\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	Eq. 4: $\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	Eq. 5: $\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	Eq. 5: $\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$	Eq. 6: $\text{Dln O}^{\text{WMT}}_{\text{NYSE}}$	Eq. 6: $\text{Dln O}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	0.048043 [0.29634]	0.607608 [3.72859]	-0.329484 [-0.74518]	0.504443 [1.14058]	0.029505 [0.18211]	0.591026 [3.62745]	-0.591719 [-3.62462]	-0.029004 [-0.17867]	-0.595320 [-1.34118]	0.257993 [0.57971]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}$	-0.015411 [-0.09820]	0.176864 [1.12124]	0.037492 [0.09942]	-0.001452 [-0.00385]	-0.009703 [-0.06194]	0.181845 [1.15448]	-0.265826 [-1.71374]	-0.071830 [-0.46568]	0.202919 [0.53880]	0.152389 [0.40358]
$\text{Dln O}(-2)^{\text{WMT}}_{\text{NYSE}}$	0.127791 [0.99880]	0.156690 [1.21835]	0.350445 [1.17972]	0.339079 [1.14117]	0.129092 [1.01107]	0.157758 [1.22870]	-0.247492 [-1.97125]	-0.211883 [-1.69712]	0.440310 [1.48856]	0.423042 [1.42645]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-0.062478 [-0.40436]	-0.257604 [-1.65865]	-0.199079 [-0.53027]	-0.166697 [-0.44390]	-0.071555 [-0.46382]	-0.265558 [-1.71175]	0.182132 [1.15647]	-0.009401 [-0.06003]	-0.374806 [-0.99924]	-0.330369 [-0.87847]
$\text{Dln O}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	-0.207681 [-1.66003]	-0.243856 [-1.93914]	-0.463999 [-1.57020]	-0.461393 [-1.56097]	-0.211773 [-1.69594]	-0.247387 [-1.97011]	0.157857 [1.22965]	0.129198 [1.01207]	-0.562925 [-1.91233]	-0.553882 [-1.87671]
C	-7.39E-05 [-0.17325]	-8.36E-05 [-0.19487]	-0.014628 [-1.34141]	-0.012525 [-1.14825]	0.000559 [1.11874]	0.000489 [0.97388]	0.000510 [1.01510]	0.000580 [1.16108]	-0.014532 [-1.34113]	-0.012509 [-1.15140]
% DPs	- -	- -	-0.005828 [-1.34407]	-0.004993 [-1.15123]	- -	- -	- -	- -	-0.006708 [-1.55310]	-0.005787 [-1.33640]
VIX	- -	- -	- -	- -	-0.002329 [-2.42457]	-0.002109 [-2.18250]	- -	- -	-0.004172 [-2.65534]	-0.003636 [-2.30788]
$\text{Dln O}(-1)^{\text{WMT}}_{\text{NYSE}}*$ VIX	- -	- -	- -	- -	- -	- -	-0.000555 [-2.26199]	-0.000612 [-2.50627]	- -	- -
R-squared	0.014121	0.050905	0.040849	0.052298	0.019204	0.055002	0.055301	0.019550	0.054671	0.061552

Table 21: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

BAC, BIDS	Equation 2		Equation 3		Equation 4		Equation 5		Equation 6	
Cointegrating Equation:										
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000		1.000000		1.000000		1.000000		1.000000	
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-0.997603 [-1864.98]		-0.994593 [-733.923]		-0.998147 [-1451.56]		-0.997817 [-1652.34]		-0.994772 [-559.179]	
C	-0.011471		-0.019315		-0.009633		-0.010747		-0.018809	
N	1209		567		1209		1209		567	
Error Correction:										
	Dependent Variable									
Independent Variable	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nonNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	-0.270131 [-1.31178]	0.554488 [2.62859]	-0.441814 [-1.37238]	0.425899 [1.29051]	-0.241051 [-1.16652]	0.588455 [2.78121]	-0.247630 [-1.20136]	0.579076 [2.74331]	-0.430816 [-1.33639]	0.437459 [1.32389]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	-0.220812 [-1.24422]	-0.129277 [-0.71111]	-0.166520 [-0.60685]	-0.107066 [-0.38061]	-0.242270 [-1.36041]	-0.154939 [-0.84971]	-0.238887 [-1.34472]	-0.149154 [-0.81987]	-0.175774 [-0.63925]	-0.116983 [-0.41507]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{NYSE}}$	-0.152992 [-1.16428]	-0.163060 [-1.21137]	-0.126396 [-0.62743]	-0.158353 [-0.76679]	-0.164423 [-1.24857]	-0.176880 [-1.31180]	-0.163366 [-1.24260]	-0.174494 [-1.29604]	-0.131652 [-0.65235]	-0.164005 [-0.79285]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	0.246547 [1.39391]	0.160505 [0.88586]	0.193521 [0.70711]	0.140122 [0.49944]	0.267764 [1.50852]	0.186023 [1.02353]	0.264464 [1.49361]	0.180292 [0.99429]	0.202602 [0.73880]	0.149892 [0.53326]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	0.207320 [1.58096]	0.209153 [1.55699]	0.181118 [0.89877]	0.204003 [0.98751]	0.218749 [1.66422]	0.223158 [1.65811]	0.217951 [1.66093]	0.220970 [1.64435]	0.186402 [0.92330]	0.209743 [1.01357]
C	-0.000868 [-0.67530]	-0.000830 [-0.63051]	-0.002082 [-0.12996]	-0.006844 [-0.41674]	-0.000299 [-0.19950]	1.59E-05 [0.01040]	0.000185 [0.12478]	0.000374 [0.24693]	-0.000339 [-0.01986]	-0.004572 [-0.26128]
% DPs	- -	- -	0.003685 [0.01989]	0.059296 [0.31225]	- -	- -	- -	- -	-0.006264 [-0.03317]	0.044737 [0.23110]
VIX	- -	- -	- -	- -	-0.002183 [-0.74134]	-0.003245 [-1.07597]	- -	- -	-0.001780 [-0.32301]	-0.002050 [-0.36297]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$* VIX	- -	- -	- -	- -	- -	- -	-0.001495 [-1.43115]	-0.001710 [-1.59888]	- -	- -
R-squared	0.017681	0.012365	0.022309	0.008334	0.017995	0.013568	0.019299	0.014564	0.022388	0.008667

Table 22: Tables 15-22 contain the coefficients and error correction terms (ECT) for the 4 selected stocks, estimated by the VEC Model. The relations are estimated for equation 2: ln(quote), equation 3: influenced by dark pool volume, equation 4: influenced by volatility, equation 5: influenced by VIX combined with a NYSE lag and equation 6: influenced by both dark pool volume and VIX. T-values are provided between brackets.

BAC, OFFERS	Equation 2	Equation 3	Equation 4	Equation 5	Equation 6
Cointegrating Equation:					
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	1.000000	1.000000	1.000000	1.000000	1.000000
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	-1.001247 [-2845.97]	-0.996268 [-1245.48]	-0.999941 [-2243.90]	-1.000644 [-2546.33]	-0.998266 [-962.966]
C	0.007155	-0.013691	-0.003138	0.005117	-0.008011
N	1144	550	1144	1144	550

Error Correction:

Dependent Variable

Independent Variable	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{NYSE}}$	$\mathrm{Dln\ O}^{\mathrm{WMT}}_{\mathrm{nNYSE}}$
ECT 1	0.034139 [0.09394]	0.806365 [2.23321]	-1.244058 [-1.94331]	-0.472298 [-0.73125]	-0.671624 [-1.84789]	0.126046 [0.34488]	-0.082819 [-0.22714]	0.704766 [1.94381]	-0.926535 [-1.45967]	-0.147258 [-0.23021]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	-0.361870 [-1.15177]	-0.174061 [-0.55757]	0.622378 [1.11546]	0.844836 [1.50080]	0.103165 [0.33044]	0.286364 [0.91215]	-0.283048 [-0.89967]	-0.105228 [-0.33636]	0.407440 [0.73323]	0.624029 [1.11439]
$\mathrm{Dln\ O}(-2)^{\mathrm{WMT}}_{\mathrm{NYSE}}$	0.040662 [0.17972]	0.083336 [0.37071]	0.118090 [0.29015]	0.149964 [0.36521]	-0.073530 [-0.32601]	-0.039770 [-0.17535]	0.081323 [0.35901]	0.117924 [0.52352]	0.017393 [0.04274]	0.045404 [0.11073]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	0.393986 [1.25796]	0.194535 [0.62513]	-0.598149 [-1.07094]	-0.809440 [-1.43645]	-0.082744 [-0.26412]	-0.254108 [-0.80664]	0.315179 [1.00518]	0.125543 [0.40265]	-0.382890 [-0.68805]	-0.588044 [-1.04860]
$\mathrm{Dln\ O}(-2)^{\mathrm{WMT}}_{\mathrm{nonNYSE}}$	0.011414 [0.05044]	-0.031427 [-0.13977]	-0.067705 [-0.16686]	-0.098980 [-0.24179]	0.125775 [0.55754]	0.092610 [0.40826]	-0.028420 [-0.12548]	-0.065561 [-0.29109]	0.033964 [0.08369]	0.007081 [0.01731]
C	-0.000707 [-0.52160]	-0.000707 [-0.52534]	0.000705 [0.04084]	-0.004397 [-0.25243]	-0.000502 [-0.31639]	7.10E-05 [0.04450]	0.000522 [0.33225]	0.000143 [0.09126]	-0.004787 [-0.27036]	-0.005105 [-0.28610]
% DPs	- -	- -	-0.025738 [-0.12724]	0.034384 [0.16848]	- -	- -	- -	- -	0.036936 [0.18453]	0.054174 [0.26857]
VIX	- -	- -	- -	- -	-0.000761 [-0.24474]	-0.002880 [-0.92134]	- -	- -	0.000404 [0.07150]	-0.001863 [-0.32731]
$\mathrm{Dln\ O}(-1)^{\mathrm{WMT}}_{\mathrm{NYSE}}*$ VIX	- -	- -	- -	- -	- -	- -	-0.001693 [-1.53854]	-0.001171 [-1.07029]	- -	- -
R-squared	0.007934	0.011883	0.013030	0.011007	0.010904	0.008686	0.009989	0.012372	0.010089	0.010370

Table 23: This table contains the coefficients and error correction terms (ECT) for the stocks IBM and KO estimated by the VEC Model. The relations are estimated for equation 7: influenced by an ECT for dark pool share in trade volume. T-values are provided between brackets.

	Equation 7: IBM BID	**Equation 7: IBM OFR**	**Equation 7: KO BID**	**Equation 7: KO OFR**
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000	1.000000	1.000000	1.000000
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	0.998593 [16855.69]	0.998771 [17831.95]	0.998271 25081.30	0.998347 [22984.09]
C	0	0	0	0
N	607	607	607	607
Error Correction:				

Independent Variable

	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nonNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	- -	- -	- -	- -	- -	- -	- -	- -
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	-0.314925 [-0.128068]	0.255120 [0.103924]	0.047707 [0.022268]	0.615230 [0.287417]	-0.229448 [-0.120853]	0.390587 [0.205551]	-0.305494 [-0.154199]	0.342034 [0.172668]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{NYSE}}$	-0.214486 [-0.101542]	0.101093 [0.048019]	-0.250465 [-0.111689]	0.073238 [0.032703]	0.113813 [0.057002]	0.414305 [0.207950]	-0.120149 [-0.055739]	0.150014 [0.069562]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	0.245993 [0.098367]	-0.307103 [0.902164]	-0.099769 [-0.047308]	-0.674436 [-0.320225]	0.094693 [0.050143]	-0.528114 [-0.279423]	0.203797 [0.099984]	-0.449483 [-0.220576]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	0.162697 [0.076255]	-0.158227 [-0.074423]	0.244611 [0.104047]	-0.079140 [-0.033698]	-0.227968 [-0.114425]	-0.514801 [-0.258928]	0.039073 [0.017654]	-0.223674 [-0.101081]
C	5.63E-05 [0.029205]	5.36E-05 [0.027856]	5.24E-05 [0.026886]	5.65E-05 [0.028984]	-8.18E-05 [-0.041894]	-7.98E-05 [-0.040938]	-8.54E-05 [-0.042946]	-8.31E-05 [-0.041687]
ECT 2 ln % DPs	-43.10126 [-0.093499]	-39.62321 [-0.085975]	1.697403 [0.004154]	21.54714 [0.052648]	-4.070410 [-0.007923]	-2.991684 [-0.005814]	86.23288 [0.158741]	75.47861 [0.138897]
Log lilelihood	5543.764		5584.776		5477.929		5493.721	

Table 24: This table contains the coefficients and error correction terms (ECT) for the stocks WMT and BAC estimated by the VEC Model. The relations are estimated for equation 7: influenced by an ECT for dark pool share in trade volume. T-values are provided between brackets.

	Equation 7: WMT BID		**Equation 7: WMT OFR**		**Equation 7: BAC BID**		**Equation 7: BAC OFR**	
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	1.000000		1.000000		1.000000		1.000000	
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	0.975729 [1445.883]		0.976319 [1432.409]		0.886295 [824.3091]		0.975193 [1989.392]	
C	0		0		0		0	
N	529		529		586		586	
Error Correction:								
	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nonNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{NYSE}}$	$\text{Dln B}^{\text{WMT}}_{\text{nNYSE}}$
ECT 1	- -	- -	- -	- -	- -	- -	- -	- -
$\text{Dln B}(-1)^{\text{WMT}}_{\text{NYSE}}$	0.616622 [4.223903]	1.232606 [8.637467]	-0.039196 [-0.315746]	0.471264 [3.504776]	-0.440211 [-8.060348]	0.326266 [5.647237]	-0.905555 [-17.40938]	-0.203490 [-3.657779]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{NYSE}}$	0.306262 [2.559560]	0.652997 [5.817032]	0.262539 [1.738926]	0.561194 [3.898607]	-0.257030 [-6.322372]	-0.013955 [-0.326766]	-0.127507 [-2.350392]	0.145203 [2.606270]
$\text{Dln B}(-1)^{\text{WMT}}_{\text{nonNYSE}}$	-0.778166 [-5.270627]	-1.366243 [-9.595530]	-0.124318 [-1.057851]	-0.637551 [-4.906394]	0.497385 [9.035035]	-0.262993 [-4.580822]	0.977138 [18.61506]	0.259935 [4.637843]
$\text{Dln B}(-2)^{\text{WMT}}_{\text{nonNYSE}}$	-0.406492 [-3.738318]	-0.737684 [-7.180913]	-0.390247 [-2.605199]	-0.694868 [-4.880048]	0.303207 [7.544418]	0.042850 [1.000739]	0.173809 [3.095999]	-0.097535 [-1.693928]
C	-2.52E-05 [-0.046922]	-2.12E-05 [-0.039882]	-6.95E-05 [-0.136883]	-4.74E-05 [-0.093931]	-0.000552 [-1.578498]	-0.000590 [-1.740752]	-0.000534 [-1.894037]	-0.000542 [-1.821510]
ECT 2 ln % DPs	-157.5583 [-7.658525]	-175.7653 [-8.335849]	[-87.89531] [-4.528218]	[-95.74257] [-4.392477]	50.38214 [1.593612]	177.2313 [7.774267]	3.687481 [0.131240	6.696968 [0.229575]
Log lilelihood	3691.749		3690.131		2551.585		3150.093	

Independent Variable

Table 25: This table contains estimations of a Bivariate GARCH model for the stock IBM. The relations are estimated for bids and offers, directly on the quote as well as on ln(quote), Dln(quote) and on the residuals of the VEC model. P-values are provided between brackets.

Ticker: IBM, BIDS	$\ln B^{IBM}$	$Dln B^{IBM}$	$\ln B^{IBM}_{VEC\,M\,resid}$	$Dln B^{IBM}_{VEC\,M\,resid}$
μ_1	4.536584	0.000536	0.000262	2.61E-06
	[0.0000]	[0.0508]	[0.0285]	[0.9938]
μ_2	4.533174	0.000417	0.000322	-0.000172
	[0.0000]	[0.1556]	[0.0000]	[0.6204]
ω_{11}	0.005645	0.003485	0.004091	0.004583
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.005953	0.003683	0.003833	0.004383
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.001251	0.001245	0.000976	0.001113
	[0.0000]	[0.0000]	[0.0004]	[0.0000]
α_{11}	0.526817	0.368178	0.343677	0.372688
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.526200	0.377669	0.316848	0.354800
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.843580	0.897879	0.884838	0.874101
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.843653	0.888844	0.900691	0.884702
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	6719.538	9609.390	10006.75	9803.969
Ticker: IBM, OFFERS	$\ln O^{IBM}$	$Dln O^{IBM}$	$\ln O^{IBM}_{VEC\,M\,resid}$	$Dln O^{IBM}_{VEC\,M\,resid}$
μ_1	4.536584	0.000728	0.000193	0.000332
	[0.0000]	[0.0046]	[0.3904]	[0.2868]
μ_2	4.533174	0.000770	6.31E-05	0.000230
	[0.0000]	[0.0022]	[0.7824]	[0.4498]
ω_{11}	0.005645	0.003212	0.002528	0.002804
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.005953	0.002743	0.002259	0.002982
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.001251	0.000360	0.000478	0.000584
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{11}	0.526817	0.387573	0.357135	0.323865
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.526200	0.354145	0.329777	0.343394
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.843580	0.901881	0.915800	0.927759
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.843653	0.921816	0.928763	0.918005
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	6719.538	9815.038	10370.39	10177.23

Table 26: This table contains estimations of a Bivariate GARCH model for the stock KO. The relations are estimated for bids and offers, directly on the quote as well as on ln(quote), Dln(quote) and on the residuals of the VEC model. P-values are provided between brackets.

Ticker: KO, BIDS	$\ln \mathrm{B}^{\mathrm{KO}}$	$\mathrm{Dln\ B}^{\mathrm{KO}}$	$\ln \mathrm{B}^{\mathrm{KO}}_{\mathrm{VEC\ M\ resid}}$	$\mathrm{Dln\ B}^{\mathrm{KO}}_{\mathrm{VEC\ M\ resid}}$
μ_1	3.880444	0.000278	0.000142	-0.000110
	[0.0000]	[0.2245]	[0.4639]	[0.5724]
μ_2	3.879871	3.46E-05	0.000359	-0.000377
	[0.0000]	[0.8762]	[0.0903]	[0.0751]
ω_{11}	0.003322	0.001334	0.001744	0.001654
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.003156	0.001526	0.002210	0.002062
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.000251	-1.11E-06	0.000924	0.000487
	[0.0000]	[0.9999]	[0.0004]	[0.0000]
α_{11}	0.494798	0.320900	0.418915	0.409412
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.485282	0.372377	0.432983	0.442752
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.871688	0.950251	0.902734	0.921848
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.877285	0.935329	0.890423	0.906719
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	7107.429	9884.541	10348.55	10209.87

Ticker: KO, OFFERS	$\ln \mathrm{O}^{\mathrm{KO}}$	$\mathrm{Dln\ O}^{\mathrm{KO}}$	$\ln \mathrm{O}^{\mathrm{KO}}_{\mathrm{VEC\ M\ resid}}$	$\mathrm{Dln\ O}^{\mathrm{KO}}_{\mathrm{VEC\ M\ resid}}$
μ_1	3.869455	7.27E-05	0.000157	8.59E-05
	[0.0000]	[0.7163]	[0.4311]	[0.6336]
μ_2	3.869799	0.000224	-5.43E-05	0.000213
	[0.0000]	[0.2531]	[0.7936]	[0.2625]
ω_{11}	0.003004	0.001956	0.001144	0.001886
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.002937	0.001660	0.000972	0.001775
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.000211	0.000716	0.000503	0.000793
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{11}	0.499874	0.420714	0.306165	0.444563
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.507477	0.420638	0.303052	0.442450
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.871979	0.902464	0.947384	0.894965
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.868009	0.906054	0.949451	0.896751
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	7512.010	10242.01	10667.29	10485.44

Table 27: This table contains estimations of a Bivariate GARCH model for the stock WMT. The relations are estimated for bids and offers, directly on the quote as well as on ln(quote), Dln(quote) and on the residuals of the VEC model. P-values are provided between brackets.

Ticker: WMT, BIDS	$\ln B^{WMT}$	$\mathrm{Dln}\, B^{WMT}$	$\ln B^{WMT}_{VEC\,M\,resid}$	$\mathrm{Dln}\, B^{WMT}_{VEC\,M\,resid}$
μ_1	3.878225	-5.19E-05	0.000141	0.000370
	[0.0000]	[0.8740]	[0.6775]	[0.2425]
μ_2	3.876086	-0.000207	0.000206	0.000149
	[0.0000]	[0.5466]	[0.5563]	[0.6602]
ω_{11}	0.006814	0.006312	0.002805	0.003936
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.006861	0.006355	0.003308	0.003500
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.001504	0.001655	0.000548	0.000159
	[0.0000]	[0.0000]	[0.0000]	0.3175
α_{11}	0.605099	0.512970	0.257903	0.403438
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.601795	0.473942	0.277933	0.348675
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.776255	0.762263	0.943997	0.884520
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.778305	0.774302	0.929187	0.912967
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	7776.154	9230.375	9516.793	9362.728

Ticker: WMT, OFFERS	$\ln O^{WMT}$	$\mathrm{Dln}\, O^{WMT}$	$\ln O^{WMT}_{VEC\,M\,resid}$	$\mathrm{Dln}\, O^{WMT}_{VEC\,M\,resid}$
μ_1	3.879265	-0.000347	-1.22E-05	0.000340
	[0.0000]	[0.2889]	[0.9698]	[0.3252]
μ_2	3.880704	-0.000253	-9.54E-05	0.000209
	[0.0000]	[0.4396]	[0.7698]	[0.5394]
ω_{11}	0.011026	0.003745	0.002231	0.002862
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.011199	0.003708	0.002477	0.002903
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.002264	-0.001019	0.000428	0.000768
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{11}	0.787198	0.328610	0.195619	0.302925
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.793425	0.345039	0.224621	0.312920
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.556779	0.908770	0.964774	0.932652
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.548456	0.902763	0.954596	0.927154
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	7765.656	9285.957	9571.911	9430.644

Table 28: This table contains estimations of a Bivariate GARCH model for the stock BAC. The relations are estimated for bids and offers, directly on the quote as well as on ln(quote), Dln(quote) and on the residuals of the VEC model. P-values are provided between brackets.

Ticker: BAC, BIDS	$\ln B^{BAC}$	$D\ln B^{BAC}$	$\ln B^{BAC}_{VEC\,M\,resid}$	$D\ln B^{BAC}_{VEC\,M\,resid}$
μ_1	3.828315	9.21E-05	0.001039	0.001083
	[0.0000]	[0.7504]	[0.0005]	[0.0005]
μ_2	3.824744	-0.000285	0.000670	0.000192
	[0.0000]	[0.4623]	[0.0379]	[0.5656]
ω_{11}	0.006601	0.001866	0.001715	0.003135
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.006079	0.002688	0.002268	0.002316
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.002452	0.001986	0.001027	0.001659
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{11}	0.777172	0.458854	0.386644	0.654168
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.774618	0.504611	0.435979	0.620737
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.679330	0.902666	0.930941	0.839217
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.682610	0.882696	0.912231	0.859938
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	5482.776	8168.494	8419.226	8168.780

Ticker: BAC, OFFERS	$\ln O^{BAC}$	$D\ln O^{BAC}$	$\ln O^{BAC}_{VEC\,M\,resid}$	$D\ln O^{BAC}_{VEC\,M\,resid}$
μ_1	3.832453	-0.000457	0.000361	0.000352
	[0.0000]	[0.0987]	[0.2431]	[0.2225]
μ_2	3.837208	-0.000298	0.000698	0.000509
	[0.0000]	[0.3254]	[0.0163]	[0.1013]
ω_{11}	0.004689	0.001550	0.000965	0.001283
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{22}	0.004656	0.001096	0.000941	0.958595
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
ω_{12}	0.001654	0.000618	0.000516	0.284893
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{11}	0.595880	0.343407	0.238328	0.000434
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
α_{22}	0.598108	0.317256	0.258024	0.001020
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{11}	0.814441	0.940683	0.969927	0.964923
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
β_{22}	0.812741	0.951305	0.963813	0.266324
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Log Likelihood	5606.886	8352.998	8800.232	8561.941

Table 29: Tables 29-32 contain the estimations for simple regressions of the dark pool market share on the covariance and correlation between the NYSE and non NYSE series estimated by the Biv GARCH model. The test is performed on the log difference as well as VEC M residuals. P-values are provided between brackets.

Covariance estimated by Biv GARCH M on:

Ticker: IBM, BIDS	$\mathbf{Dln\ B^{IBM}}$	$\mathbf{ln\ B^{IBM}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{IBM}_{VEC\ M\ resid}}$
Ln DP %	-0.000164	-0.000146	-0.000186
	[0.0159]	[0.0036]	[0.0070]
Constant	-9.36E-05	-0.000103	-0.000112
	[0.5786]	[0.4052]	[0.5096]
N	608	608	608
R^2	0.009556	0.013860	0.011932

Ticker: IBM, OFFERS	$\mathbf{Dln\ O^{IBM}}$	$\mathbf{ln\ O^{IBM}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{IBM}_{VEC\ M\ resid}}$
Ln DP %	-0.000251	-0.000218	-0.000231
	[0.0022]	[0.0008]	[0.0229]
Constant	-0.000274	-0.000263	-0.000187
	[0.1750]	[0.1005]	[0.4574]
N	608	608	608
R^2	0.015423	0.018570	0.008514

Correlation estimated by Biv GARCH M on:

Ticker: IBM, BIDS	$\mathbf{Dln\ B^{IBM}}$	$\mathbf{ln\ B^{IBM}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{IBM}_{VEC\ M\ resid}}$
Ln DP %	-0.017336	-0.013977	-0.013015
	[0.0085]	[0,0000]	[0,0000]
Constant	0.912974	0.936851	0.943987
	[0,0000]	[0,0000]	[0,0000]
N	608	608	608
R^2	0.011355	0.030140	0.030136

Ticker: IBM, OFFERS	$\mathbf{Dln\ O^{IBM}}$	$\mathbf{ln\ O^{IBM}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{IBM}_{VEC\ M\ resid}}$
Ln DP %	-0.003601	-0.011501	-0.009339
	[0.4868]	[0.0030]	[0.0014]
Constant	0.956457	0.947102	0.958320
	[0,0000]	[0,0000]	[0,0000]
N	608	608	608
R^2	0.000798	0.014431	0.016764

Table 30: Tables 29-32 contain the estimations for simple regressions of the dark pool market share on the covariance and correlation between the NYSE and non NYSE series estimated by the Biv GARCH model. The test is performed on the log difference as well as VEC M residuals. P-values are provided between brackets.

	Covariance estimated by Biv GARCH M on:		
Ticker: KO, BIDS	$\mathbf{Dln\ B^{KO}}$	$\mathbf{ln\ B^{KO}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{KO}_{VEC\ M\ resid}}$
Ln DP %	-0.034259	-0.000233	-0.000189
	[0,0000]	[0.0057]	[0.0055]
Constant	-0.069072	-0.000263	-0.000208
	[0,0000]	[0.2064]	[0.2175]
N	608	608	608
R^2	0.228797	0.012522	0.012632
Ticker: KO, OFFERS	$\mathbf{Dln\ O^{KO}}$	$\mathbf{ln\ O^{KO}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{KO}_{VEC\ M\ resid}}$
Ln DP %	-0.037727	-0.000260	-0.000250
	[0,0000]	[0.0024]	[0.0003]
Constant	-0.075726	-0.000363	-0.000365
	[0,0000]	[0.0865]	[0.0326]
N	608	608	608
R^2	0.213253	0.015101	0.021412

	Correlation estimated by Biv GARCH M on:		
Ticker: KO, BIDS	$\mathbf{Dln\ B^{KO}}$	$\mathbf{ln\ B^{KO}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{KO}_{VEC\ M\ resid}}$
Ln DP %	0.010373	-0.014641	-0.007695
	[0.1600]	[0.0059]	[0.1350]
Constant	0.989510	0.931101	0.957713
	[0.0000]	[0,0000]	[0,0000]
N	608	608	608
R^2	0.003254	0.012433	0.003683
Ticker: KO, OFFERS	$\mathbf{Dln\ O^{KO}}$	$\mathbf{ln\ O^{KO}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{KO}_{VEC\ M\ resid}}$
Ln DP %	0.036312	0.007231	-0.004423
	[0.0302]	[0.3760]	[0.6210]
Constant	1.041761	0.985655	0.960324
	[0,0000]	[0,0000]	[0,0000]
N	608	608	608
R^2	0.007728	0.001294	0.000404

Table 31: Tables 29-32 contain the estimations for simple regressions of the dark pool market share on the covariance and correlation between the NYSE and non NYSE series estimated by the Biv GARCH model. The test is performed on the log difference as well as VEC M residuals. P-values are provided between brackets.

Covariance estimated by Biv GARCH M on:

Ticker: WMT, BIDS	$\mathbf{Dln\ B^{WMT}}$	$\mathbf{ln\ B^{WMT}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{WMT}_{VEC\ M\ resid}}$
Ln DP %	-8.62E-05	-7.92E-05	-0.000206
	[0.2680]	[0.0957]	[0.0182]
Constant	6.28E-05	4.48E-05	-0.000162
	[0.7481]	[0.7072]	[0.4600]
N	529	529	529
R^2	0.002328	0.005259	0.010528
Ticker: WMT, OFFERS	$\mathbf{Dln\ O^{WMT}}$	$\mathbf{ln\ O^{WMT}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{WMT}_{VEC\ M\ resid}}$
Ln DP %	-0.000106	-5.45E-05	-0.000183
	[0.0884]	[0.1736]	[0.0168]
Constant	4.88E-06	9.52E-05	-0.000120
	[0.9750]	[0.3435]	[0.5317]
N	529	529	529
R^2	0.005500	0.003510	0.010792

Correlation estimated by Biv GARCH M on:

Ticker: WMT, BIDS	$\mathbf{Dln\ B^{WMT}}$	$\mathbf{ln\ B^{WMT}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{WMT}_{VEC\ M\ resid}}$
Ln DP %	-0.017060	-0.007297	-0.008507
	[0.0179]	[0.0158]	[0.0406]
Constant	0.925006	0.957552	0.955641
	[0.0000]	[0.0000]	[0.0000]
N	529	529	529
R^2	0.010599	0.011010	0.007932
Ticker: WMT, OFFERS	$\mathbf{Dln\ O^{WMT}}$	$\mathbf{ln\ O^{WMT}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{WMT}_{VEC\ M\ resid}}$
Ln DP %	-0.013967	0.003535	-0.011414
	[0.0045]	[0.1437]	[0.0003]
Constant	0.934443	0.984504	0.951120
	[0.0000]	[0.0000]	[0.0000]
N	529	529	529
R^2	0.015204	0.004052	0.024978

Table 32: Tables 29-32 contain the estimations for simple regressions of the dark pool market share on the covariance and correlation between the NYSE and non NYSE series estimated by the Biv GARCH model. The test is performed on the log difference as well as VEC M residuals. P-values are provided between brackets.

	Covariance estimated by Biv GARCH M on:		
Ticker: BAC, BIDS	$\mathbf{Dln\ B^{BAC}}$	$\mathbf{ln\ B^{BAC}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{BAC}_{VEC\ M\ resid}}$
Ln DP %	0.000327 [0.8295]	0.000404 [0.7715]	0.000598 [0.8358]
Constant	0.005373 [0.1554]	0.005460 [0.1148]	0.009260 [0.1971]
N	587	587	587
R^2	0.000079	0.000144	0.000074
Ticker: BAC, OFFERS	$\mathbf{Dln\ O^{BAC}}$	$\mathbf{ln\ O^{BAC}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{BAC}_{VEC\ M\ resid}}$
Ln DP %	0.000326 [0.7898]	0.000468 [0.6517]	0.000567 [0.7102]
Constant	0.004937 [0.1045]	0.004865 [0.0591]	0.006743 [0.0760]
N	587	587	587
R^2	0.000122	0.000349	0.000236

	Correlation estimated by Biv GARCH M on:		
Ticker: BAC, BIDS	$\mathbf{Dln\ B^{BAC}}$	$\mathbf{ln\ B^{BAC}_{VEC\ M\ resid}}$	$\mathbf{Dln\ B^{BAC}_{VEC\ M\ resid}}$
Ln DP %	-0.095108 [0,0000]	-0.100812 [0,0000]	-0.069542 [0,0000]
Constant	0.715921 [0,0000]	0.711153 [0,0000]	0.799382 [0,0000]
N	587	587	587
R^2	0.039484	0.064636	0.035206
Ticker: BAC, OFFERS	$\mathbf{Dln\ O^{BAC}}$	$\mathbf{ln\ O^{BAC}_{VEC\ M\ resid}}$	$\mathbf{Dln\ O^{BAC}_{VEC\ M\ resid}}$
Ln DP %	0.037110 [0.0007]	0.028281 [0,0000]	0.059038 [0,0000]
Constant	1.056229 [0.0000]	1.048021 [0,0000]	1.127745 [0,0000]
N	587	587	587
R^2	0.019546	0.053175	0.181952

Table 33: This table contains simple regression tests on the VIX index having explanatory power on the share dark pools have in US consolidated equity volume. Observations are monthly. P-values are mentioned between brackets).

	Ln DP %	**DP %**
Ln VIX	0.057684	-
	[0.7894]	-
VIX	-	-0.000381
	-	[0.1756]
Constant	-3.176471	0.101248
	[0.0001]	[0.0000]
N	42	35
R^2	0.001804	0.054864