Glasgow Caledonian University

Department of Law, Economics, Accounting and Risk

An attempt to attune VPIN to a high frequency (micro) vista to study its risk related utility

by

J.J. Moore

Table of Contents

ACKNOWLEDGEMENTS

I would like to take this opportunity to sincerely thank the following people for their assistance and guidance that greatly assisted the completion of this research dissertation: Professor John Hood for his structural and content guidance, and Judith Farrell for delivering the gentle, but decisive push I needed to get the work done.

ABSTRACT

With the advent of High Frequency Trading, trading algorithms, flash crashes, polarised competing market paradigms and the constantly evolving nature of today's electronic market place, this study set out to test the VPIN metric (Easley, López de Prado & O'Hara, 2011a) in a micro environment to ascertain its practical utility to identify and manage the risk of participation in today's machine based marketplace. Using tick-by-tick historic data from the recent ES 06-15 futures contract and a commercially available trading platform, this study deployed a micro configured VPIN algorithm over the historic data provided to examine the direct and post hoc effect of a VPIN event equal to or above its 95% CDF value. The results are interesting and provide strong evidence to support the argument that a VPIN event is significantly different from a Non VPIN event in terms of its immediate *market dynamics*. Moreover, the evidence suggests that all but one of these observable differences appear to be ephemeral, disappearing at the close of the originating period, and it is only significant increased levels of volume that appear to persist post hoc within the bounds of random price movements. However, the nature and extent of this post hoc VPIN originated volume

persistence is currently unknown and further research would be well placed to examine this phenomenon in more detail.

DEFINITION OF TERMS

Dark pools: Alternative exchanges where trades are matched and where the size and price of the orders are not revealed to other participants (FT.com, 2015).

GARCH: A model to estimate volatility in financial markets (Investopedia, 2015a).

***Intraday*:** A single trading session represented by one calendar day (Admati & Pfleiderer, 1988).

***Liquidity*:** The buy-side and sell-side resting orders outside of the current market levels (SEC & CFTC, 2010).

Market dynamics: The changing prices and volumes that result from the continual changes in both supply and demand (Investopedia, 2015b).

***Market microstructure*:** The study of mechanisms and structures used to trade financial securities (Vishwanath & Krishnamurti, 2009).

***Order-book*:** The electronic collection of the outstanding limit orders for a financial instrument (Kane, Lui & Nguyen, 2011).

***Order-flow*:** The cumulative flow of classified transactions where each transaction is classified by execution aggression; ‘at the ask’ denotes a buy and ‘at the bid’ denotes a sell (Evans & Lyons, 2004).

Passive order: An order that does not cross the market thus the originator has no direct control on the timing of its execution (Easley, López de Prado & O'Hara, 2011a).

Position or price - takers: Buying and selling transactions that are assumed to have no effect on the market (Mandelbrot & Hudson, 2005).

Price discovery: The process of determining the prices of assets in the marketplace through the interactions of buyers and sellers over time (NASDAQ, 2015).

Spread(s): The difference between the immediate buying and selling price of a particular financial instrument (Bodie, Kane & Marcus, 2008).

1.0 INTRODUCTION

Financial market participation is a very risky business and any would-be entrant is usually met with a full risk warning unambiguously stating in bold text that – ***"investing and or trading carries substantial risk and is not for every investor. Only money that we are told can be lost without jeopardizing ones financial security or life style should be used for trading"***. In addition to this, and not to be forgotten, **"*past performance is not necessarily indicative of future results*"**. That all sounds fair, above-board and many would retort: I know what I am able to risk and I am well versed in the paradigms of "modern financial theory" - so let me play the game. Well before we start, what the abovementioned risk warning does not state is that maybe the risks are not quite the same as the ones you remember or were taught. Consider that markets may now be very, very risky; riskier than the standard financial models state (Mandelbrot & Hudson, 2005). In addition to this, there are new risks emerging that could totally liquidate all of your risk capital in a flash and without warning. To that end and under this context, this research study explores and undertakes a critical analysis of relevant market paradigms, market structure evolution and participant behaviour and discusses, tests and evaluates the current academic research aimed at identifying and responding to such risks.

1.1 BACKGROUND LITERATURE

Shortly after the Greek parliament vote in favour of a thirty billion Euro austerity package - CNBC reporters speculate over the fate of global financial markets as riots break out in Athens. The Dow Jones Index is trading down 260 points with unusually high volatility and thinning *liquidity*. All ten sectors are down and six out of the last nine days have been down days (CNBC, 2010). According to the Joint SEC-CFTC Advisory Committee report (SEC & CFTC, 2010), against this gloomy backdrop a large fundamental trader initiated a program to sell a total of 75,000 E- Mini contracts (valued at approximately $4.1 billion) as a hedge to an existing equity position. This sell pressure was initially absorbed by High Frequency Traders (HFT), fundamental buyers and cross-market arbitrageurs who in turn transferred this selling pressure to the equity markets. Essentially, HFT initially acted like buyers for this selling program, but after a short while the HFT began to aggressively sell their net long positions. As a result, the sell algorithm used by the large trader responded to the increased volume by increasing the rate at which it was feeding the orders into the market. What happened next is best described in terms of two *liquidity* crises – one at the broad index level in the E-Mini, the other with respect to individual stocks. The events that followed are now known as the “2010 Flash Crash” whereby many US based equity products experienced an astonishingly rapid decline and recovery in value (SEC & CFTC, 2010). As the Dow Jones dropped by 998.5 points and then back up in minutes to around 400 points down for the day, CNBC’s Mad Money’s Jim Cramer declares that “the machines broke” (CNBC, 2010).

Anderson & Noss (2013) reiterate that events such as the 2010 Flash Crash again call into question the Efficient Market Hypothesis (EMH) paradigm and the ability of the Gaussian distribution to capture the likelihood of such so called “rare events”. Moreover, they point out that concepts that follow the assumed Gaussian “normality” of markets are still used widely today by risk managers and regulators in their evaluation of models such as Value-at-Risk. Further to this, they revisit the Fractal Market Hypothesis (FMH) and propose a preliminary quantitative model and show that it was able to better match some of the observed properties of financial market prices, particularly the cause of instability. The FMH highlights the role of market *liquidity* and the heterogeneity of investors’ interpretation of information as determinants of market stability. These are described as fractal structures that give rise to a sort of robustness whereby, under normal market conditions, the differing interpretation of information by, and behaviour of, investors with different time horizons combine to ensure market *liquidity* and orderly price movements. However, this fractal structure also implies a certain type of fragility in non-normal market conditions that can cause this *liquidity* to evaporate, producing panic selling and associated market crashes. Haldane (2011) suggests that the lack of *liquidity* during the 2010 Flash Crash was a consequence of this type of fragility and proposed that longer-term investors withdrew from the market when they came to doubt the veracity of price information caused by the stress of the situation and the interaction of investors (HFT) who viewed the market at a higher frequency than themselves. A recent news report took a look inside the esoteric world of HFT and reported to the American public that in today’s market structure the majority of trades executed are not placed by people, but by high speed computers loaded with predefined algorithms designed to scan the markets to find and execute opportunities that exist for only a fraction of a second. The reporter speculates that on the one hand this practice might detract from the traditional capital raising process, but on the other hand it appears that HFT might be good for the market; in terms of reducing transaction costs and providing *liquidity* (CBS NEWS, 2011).

In terms of specific academic research into the 2010 Flash Crash, Easley, López de Prado & O'Hara (2011a) present evidence that a "toxicity" in *order-flow* and the resultant order imbalance, as captured by their VPIN (Volume Synchronized Probability of Informed Trading) metric, reached increasingly high levels in the hours before the 2010 Flash Crash and this toxicity contributed to the withdrawal of many *liquidity* providers from the market. In addition to this, Easley, López de Prado & O'Hara (2011a) show that VPIN appears to predict short-term toxicity-induced volatility, particularly as it relates to large price movements, and a Monte Carlo study of their estimates appears to be robust for all theoretically possible combinations of parameters. In a following paper, Easley, López de Prado & O'Hara (2011b) state that the Joint SEC & CFTC Advisory Committee report (SEC & CFTC, 2010) identified conjunctural factors as the initial explanation for the 2010 Flash Crash and acknowledge that, while such factors may have played a role, their analysis suggests that the 2010 Flash Crash is better understood as a *liquidity* event (specifically a *liquidity* induced crash) arising from structural features of the new high frequency world of trading. Furthermore, rather than banning HFT firms, they suggest the VPIN metric will help market participants to recognise and manage the risk of trading in this new market structure.

1.2 MOTIVATIONS

According to Haldane (2011) a pre-emptive method for early warnings of systemic fault-lines and stresses in markets before they crash using recent transaction data to determine the risk before the fact would be a big prize indeed. In his speech he cites Easley, López de Prado & O'Hara's (2011b) suggestion that measures of *order-flow* imbalance may provide early warning signs of *liquidity* voids and price dislocations. Moreover, he states that this philosophy follows closely in the fractal footprints of Mandelbrot & Taylor (1967) insofar as any persistent *order-flow* imbalance could potentially cause *liquidity* problems down the line. To that end, it would be of obvious benefit if VPIN could be used in real-time with confidence to indicate the increased likelihood of future crashes that could currently manifest without any prior warning.

However, the problem with any proposed study of VPIN and its possible utility is that *liquidity* induced crashes are (thankfully) a relatively rare event at a macro level (Easley, López de Prado & O'Hara, 2011a). Notwithstanding this, and applying deductive logic to assist us with this particular problem - if we accept the premise that VPIN is able to provide an early warning of toxicity and *liquidity* induced volatility (Easley, López de Prado & O'Hara, 2011b) which the SEC-CFTC official study points out as the primary cause behind the 2010 Flash Crash (Easley, López de Prado & O'Hara, 2011a), and the further premise that fractal structures exist in economic time series data, as described by the Fractal Market Hypothesis (Peters, 1991), then it should follow that we should be able to identify and study instances of similar micro *liquidity* induced crashes from a higher frequency vista and deploy a fractal version of VPIN to measure its effectiveness. In other words, we should be able to detect *intraday* micro

Flash Crashes that are characterised by relatively quick, large drops and recoveries in price following VPIN events, and compare these events with those of randomly generated counterpart events to determine if any significant differences are immediately observable or persistent over the ensuing exogenous short period(s) of time.

1.3 AIMS AND OBJECTIVES

Therefore, using a suitable proxy for the market and the analysis and evaluation of randomly selected samples of historic data, the aim of this study is to determine if VPIN can be used by market participants as an effective predictive real-time risk management tool, or as a measurement of short term market risk during the *intraday price discovery* process.

The following sub-problems will be used to address the aim of this research paper:

Sub-problem 1: Explore and critique the relevant existing market paradigms relating to market risk and market stability.

Sub-problem 2: Describe and assess to what extent technology and specifically HFT shapes today's market structure and stability.

Sub-problem 3: Explore and discuss market participant behaviour pertaining to informed trading, adverse selection, toxic *order-flow* and the link between market confidence and the provision of *liquidity*.

Sub-problem 4: Define and evaluate the origins, development and use of VPIN as a response to such emergent risk(s).

Sub-problem 5: Conclude and present findings and justify this study's adapted procedure to estimate VPIN at a fractal level.

Sub-problem 6: Collect, organise and perform the algorithmic analysis and calculation outlined herein to discover the study results.

Sub-problem 7: Compare, comment upon, state and determine the result of the statistical analysis.

Sub-problem 8: Reject or support, in turn, each of the hypotheses presented below and summarise conclusions / recommendations.

1.4 THE HYPOTHESES

For VPIN to offer any practical utility in a real-time environment it must be able to, using only *order flow* data, effectively denote or predict a change of onward short term *market dynamics* so that market participants can adjust their understanding of risk and their participation appetite. It therefore follows that if VPIN events are to be reliable then there must be an observable significant difference between a "VPIN" event and a "Non VPIN" event. To that end, the following hypotheses will be deployed to test for such differences for each individual experiment undertaken in this study.

I. The first null hypothesis H_{0a}: proposes that there will be no difference in future short term *market dynamics* between a VPIN event and a Non VPIN event at the close of an event bar or a randomly selected post hoc exogenous time period(s).

II. The second null hypothesis H_{0b}: proposes that there will be a random non directional relationship between the sign of a VPIN event and the future short term price direction at the close of an event bar or a randomly selected post hoc exogenous time period(s).

III. H_{aa} and H_{ab}: support the alternative hypotheses respectively and the presence of a significant difference and a non-random directional relationship.

1.5 OUTLINE METHODOLOGY

A quantitative research methodology is proposed for this study. A simple random process will be used to select from the secondary data and the primary data will be derived, organised, analysed and evaluated using descriptive / inferential statistics techniques and relationship / non randomness analysis. The composition of the sample data sets will be such that they facilitate statistical inference, and finally the quantitative results will be presented to enable a direct comparison of each statistical distribution in turn, and measurements of association taken to either reject or support each hypothesis proposed herein.

In terms of delimitations, this study is delimited to the “observation” and “effect” phenomena of trade imbalance and does not consider trade intensity. In addition, this study is delimited to the most recent error free data available to the researcher from the historic ES 06-15 futures contract; considered to be a suitable market index proxy (CME, 2003) spanning its most active sessions (by volume).

1.6 CONCLUSIONS

The study will now proceed with a review of the relevant literature. The conclusions drawn from each section will provide a foundation and guide the structure and approach adopted in the methodology to follow; during its construction through to the testing necessary to discharge the aim of this study.

2.0 LITERATURE REVIEW

2.1 INTRODUCTION

Building on and considering the context and background literature above, we start the literature review by extracting what appear to be four key areas that will need to be unpicked during this review; areas that seem to interact and give rise to the risk of *liquidity* issues and ergo increase the likelihood of the specific type of market crashes that the researcher will attempt to investigate and discuss during this study. Essentially, these four areas encompass 1] the enduring paradigms used to explain how markets work and how participants should go about quantifying the risk of engagement. 2] The evolution / revolution of technology and the emergence of HFT and a machine based market place. 3] The ensuing behaviour of market users and the potential consequences of their individual and collective behaviour in a machine based market place, and finally 4] the current and specific academic research that attempts to explain and identify the risk of Flash Crash type events. We conclude the literature review with a brief summary that leads us to a natural reiteration of the argument to justify the micro / fractal nature of this study.

2.2 PARADIGMS, MARKET RISK AND STABILITY

Mandelbrot & Hudson (2005) use the term “orthodoxy” to describe what business schools now call “modern finance”. The fundamental concept: prices are not predictable, but their fluctuations can be described by the mathematical laws of chance and this asserts that price movements can be understood as if they follow a normal, Gaussian distribution. Therefore, risk is measurable and manageable using the variance or standard deviation as a proxy for your risk and your reward is the Gaussian distribution mean, known as the expected return. This uniform order is an apparent contradictory position if we consider the results derived from one of the first applications of computers in economics in the 1950s. According to Bodie, Kane & Marcus (2008) economic time series' were analysed by early computers to trace key economic variables in an attempt to predict the progress of the economy through boom and bust periods. Maurice Kendell, (1953) broadly speaking, concluded that there appears to be no hope of being able to predict movements in price without knowledge of extraneous information and that prices appeared to wander randomly. Bodie, Kane & Marcus (2008) suggest that Kendell's results at first appeared to confirm the "irrationality" of markets. However, economists came to revise their interpretation of Kendell's study and suggested that random price movements indicated a well-functioning or "efficient market", not an irrational one.

Fama (1965) describes an efficient market as a market where large numbers of rational participants actively compete to predict future market values and where important current information is almost freely available to all. This competition results in a situation where, at any point in time, actual prices already reflect the effects of information based both on events that have already occurred and on events which, as of now, the market expects to take place in the future. Moreover, Fama (1965) posits that the vagueness of new information in itself creates randomness, and a market where successive price changes are independent is by definition a "random-walk" market. In this context, the word "random" in the random walk theory denotes the presence of a Gaussian, normal or "Bell Curve" distribution which has become one of the major foundations of financial market theory (Mandelbrot & Hudson, 2005). An early version of the random-walk theory was proposed by Louis Bachelier in 1900 (Davis & Etheridge, 2006). In his work, that studied the French bond market, Bachelier concluded that the influences that determine the movements of the exchange are innumerable and deemed it impossible to hope for mathematical predictability. Despite the fundamental importance of Bachelier's process in itself, which has come to be known as “Brownian Motion” (Mandelbrot, 1963), it appears that Bachelier’s work was largely ignored at the time of its writing (Mandelbrot & Hudson, 2005). However, it formally introduces the Gaussian distribution assumption in price movement, and under this assumption you advance and arrive at the edifice that collectively forms today’s modern financial theory. Mandelbrot & Hudson (2005) point to the assumption of normality in price changes, and variance and standard deviation as good proxies for risk in the theories that describe how an investor should rationally select a risk adjusted portfolio of stocks; first described in the work of Markowitz (1952) and following this, some twelve years later in the work of Sharpe (1964), Lintner (1965) and Mossin (1966) in the development of the Capital Asset pricing model used to provide a benchmark rate of return for evaluating investments or providing an educated guess on the expected return of an asset not yet traded in the marketplace (Bodie, Kane & Marcus, 2008).

Notwithstanding the above, Mandelbrot (1963) originally pointed out that it is obvious, at the time of his writing, that the data accumulated since the nineteen hundreds by empirical economists demonstrates that price changes are usually too leptokurtic (higher “peaked" with “fatter tails”) to be samples from normal Gaussian populations. In their book, Mandelbrot & Hudson (2005) argue that their study of the daily index movements of the Dow Jones Industrial Average (1916 to 2003) shows that price changes do not appear to spread out on graph paper like a simple bell curve. Mandelbrot states that there are too many big changes evident over what the EMH predicts. For example, during the time studied above, there should be fifty-eight days when the Dow moved more than 3.4 percent; in fact, there were 1,001. Theory predicts six days of index swings beyond 4.5 percent; in fact, there were 366. Finally, an index swing of more than seven percent should come once every 300,000 years; in fact, the twentieth century saw forty-eight such days. Moreover, in his research on the movement of cotton prices, Mandelbrot (1963) showed that the changes in price were more like a mixture of sand, pebbles, rocks, and boulders and not simply a granulated heap as would be expected in a continuous normal distribution. To Mandelbrot this mixture appeared to exhibit what he termed “roughness“. Further, Mandelbrot (1963) mused over a bit of market folk-wisdom that advocates that all charts look alike, and without any identifying legends it is impossible to tell if the price chart covers eighteen minutes, eighteen months, or eighteen years. This insight and curiosity led Mandelbrot to look closer at cotton price movements over differing timeframes, and he found that they had a similar (but, not exact) statistical structure that was approximately persistent at each level of examination. In other words, at each timescale he saw a structure of “self-similarity”, i.e. the same proportion of big changes to small, the same fat tails and the same odds of another big change coming (Mandelbrot & Hudson, 2005). According to Anderson & Noss (2013) theorists have yet to agree on an exact mathematical definition of fractals, but there is a broad consensus that this self-similarity, whether exact or qualitative (given the whole has the same shape as one or more of the parts), symbolizes the defining characteristic. To Mandelbrot, the roughness of price changes and the inherent mathematics of self-similarity observed over different time-scales during his cotton price movement research denoted that fractals lie at the very heart of finance. In turn, this led to the development of a new tool to model the financial time/price series; Fractal Brownian Motion (Mandelbrot & Hudson, 2005).

Building on this body of work, further research aimed to address the shortcomings discussed above for the EMH can be found in the Fractal Market Hypothesis (FMH) that was first proposed and formalised by Peters (1991). The notion of mathematical self-similarity (fractal price movement) with the addition of chaos theory underpins the structure of this model. Peters' (1991) fundamental concept posits that the heterogeneity of investors, with respect to their investment horizons, facilitates financial trading due to information having a different effect on different investors; either because they obtain this information at different times, or because some property of their own preferences means they interpret the information itself differently. According to Anderson & Noss (2013) the FMH stresses the acknowledgement, role and importance of heterogeneity in terms of the provision of liquidity and the impact (relatedly) of information that is worryingly missing from the EMH. Moreover, the FMH provides a better match to some of the observed price movement characteristics apparent in the market when compared to the EMH, particularly during times of stress, and suggests that differing investment horizons when persistent create a "special" sort of fractal financial stability in the market. However, Peters (1991) describes this type of stability as at risk of failure if an exogenous event causes shorter term investors to sell off in a panic, and longer term investors to doubt the validity of the information on which they base their behaviour. To that end the result would be overwhelmingly negative short-term *market dynamics*.

2.3 MARKET STRUCTURE, HFT & STABILITY

At the time of their trading Talk publication, Rosenblatt (2008) estimated that HFT accounts for approximately two-thirds or more of U.S. equity volume. This enormous figure appears to be confirmed by research carried out by Iati (2009) that reports that HFT firms, which represent approximately only 2% of the nearly 20,000 trading firms operating in the U.S. markets, have accounted for over 73% of all U.S. equity trading volume. However, according to Easley, López de Prado & O'Hara (2011b), this increased share of HFT activity has not been accompanied by an increase in absolute volume. On the contrary, since 2009 they report that overall equity and futures volumes have fallen, in part due to the lack of participation of retail investors that followed the market downturn in 2008.

A CFTC (2010) report prior to the May 6th 2010 Flash Crash proclaims HFT as one of the most significant market structure developments in recent years. The report acknowledges that the term HFT is relatively new and is not yet clearly defined or understood, and that this absence of a clear definition complicates any proper examination of market structural issues. According to the CFTC (2010) report, HFT appears to be typically the domain of professional traders acting in a proprietary capacity that engage in various strategies that result in the generation of a large number of trades on a daily basis. In a recent market structure review paper the SEC (2014) suggests that, rather than focusing on any specific attempt at a single, precise definition of HFT, the attention should be on particular strategies and tools that may be used by such firms and should inquire as to whether any of the strategies and tools used raised concerns (structural / participant based) that need to be addressed. The SEC (2014) paper identifies four types of short-term trading strategies deployed by HFT firms: 1] passive market making, 2] aggressive arbitrage, 3], aggressive directional and 4] aggressive structural. In short, market making primarily involves the submission of non-marketable resting orders that provide *liquidity* to the marketplace at specified prices where the practitioner's profits are derived from buying at the bid and selling at the offer and capturing any rebates offered for this supply of *liquidity* by an exchange / market. Arbitrage generally seeks to capture miss-pricing between related / normally correlated products and does not depend on directional price moves. Directional strategies generally involve establishing a long or short position in anticipation of a price move up or down, and finally structural refers to the use of tools or speed by HFT firms that target trade with market participants operating from venues that offer execution opportunity at stale prices.

In terms of the impact of HFT on Institutional Investors, Tong (2015) asserts that traditional institutional investors have expressed serious concerns that HFT firms and the strategies discussed above may greatly adversely impact on their trading profits. These types of investors (mutual funds, pensions, insurance firms, and hedge funds) account for over 50% of the public equity ownership in the U.S. (French, 2008) and their participation generates huge volumes of trading. As a consequence of this activity they incur trading costs that ultimately define their overall performance (therefore, for larger investors - trading cost is often viewed as an important yardstick for measuring the quality of *liquidity* in a given financial market). As a result, regulators have serious reservations about the current equity market structure, whether markets that function with such high levels of HFT activity meet the Institutional Investor's need to trade efficiently and fairly, and if any perceived inefficiency / unfairness has led some asset managers to engage with off-exchange trading venues, such as *dark pools*. Tong (2015) examines the effect of HFT strategies on Institutional investors' and provides evidence that suggests that the actual *liquidity* resulting from HFT is at best described as ephemeral when HFT firms deploy passive market making strategies and, in addition to this, his evidence suggests that aggressive HFT strategies appear to provide *liquidity* at a premium. Tong (2015) provides statistically significant evidence that a one standard deviation increase in HFT activity intensity results in an increase in institutional trading costs by up to a third. However, he concludes his study by suggesting that the actual final cost impact of such HFT activity on institutional trading depends greatly on the "skill" of the institutional trader and his / her ability to identify and alleviate the potential adverse impact of HFT.

In an attempt to consider all relevant academic work regarding HFT to date the SEC (2014) recognize the apparent diverse effect of HFT strategies when participants try to assess market quality. In general, the data reviewed in the paper suggests that passive HFT strategies appear to have beneficial effects on market quality, such as reducing *spreads* and reducing *intraday* volatility on average. In contrast, aggressive HFT strategies raise more potential issues, with both positive and negative consequences. On the positive side, aggressive HFT strategies can improve certain dimensions of *price discovery* (speeding up efficiency), at least across very short time-frames. On the negative side, aggressive HFT activity can also impose costs on other market participants and contribute to extreme volatility events.

In terms of the May 6th 2010 Flash Crash, Kirilenko et al. (2014) argue that HFT firms did not cause the crash but contributed to the extraordinary market volatility experienced during the day. In short, they discuss and conclude that a large institutional sell program caused a large order imbalance which was amplified by HFT firms engaging in their typical practices, and that under calm market conditions this trading activity accelerates price movement and adds to trading volume, but does not result in an overall directional price move. However, during times of market stress typical HFT practices can exaggerate a directional price move and significantly increase volatility. Higher levels of volatility in turn induce HFTs to act even faster, creating a vicious cycle that results in a spike in trading volume - setting the stage for a Flash Crash type event as *liquidity* eventually disappears.

2.4 LIQUIDITY, CONFIDENCE AND BEHAVIOUR

Governor Kevin Warsh (2007) of the Federal Reserve Board describes *liquidity* as a sort of measurement of investor confidence and states that *liquidity* exists at its highest possible quality when investors are confident in their own ability to quantify their perceived risks. Therefore, *liquidity* appears to represent the life blood of a market. However, given that the quality of the *liquidity* provided in a majority (circa 75%) of today's electronic market place is the direct result of HFT participation (Rosenblatt, 2008), it would be prudent to be cognisant that a large amount of overall market *liquidity* could in fact be ephemeral as asserted by Tong (2015). In their work on Flow Toxicity and Liquidity in a High Frequency World, Easley, López de Prado & O'Hara

(2011a) hypothesise that HFT firms typically practice the provision of *liquidity* to *position-takers* by placing *passive orders* at various levels of the electronic order book and assume that HFT firms deploying this strategy, in general, do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades. As a result, and under this assertion, the ability of HFT to make money depends greatly on limiting positional risk which in turn is affected by the ability to identify and control "adverse selection" in the execution of passive *orders*. In this context, Jeria & Sofianos (2008) define adverse selection as the tendency for *passive orders* to fill quickly when they should fill slowly, and fill slowly (or not at all) when they should fill quickly. In general, adverse selection in financial markets theory (also known as negative selection) is said to signify the interaction between an uninformed trader and an informed trader whereby the uninformed trader is more likely to see prices move immediately against them following execution (Saraiya & Mittal, 2009). Moreover, Easley, López de Prado & O'Hara (2011a) define order flow as "toxic" when the practice of adverse selection is active and heightened in a market and passive *liquidity* providers (HFT or traditional market makers) are unaware (ergo uninformed) that they are providing *liquidity* at a loss.

In terms of how professionals or academics go about identifying informed trading, O'Hara (2014) acknowledges that this has always been a fundamental issue in the development of *market microstructure* models and is usually approached by an attempt to unpick and categorise market transactional data. The notion that informed traders leave "footprints" in markets is well established, and it is the reason why *market microstructure* models ascribe such significance to trade data. Indeed every trade must have a buyer and a seller, but *market microstructure* models have traditionally been interested in the active or aggressive side of a transaction because of its assumed signal denotation within the underlying information. In other words, that “BUYS” (aggressive buying from the offer) were signals of good news (and prices should rise) and “SELLS” (aggressive selling to the bid) were signals of bad news (and prices should drop), ergo directional. Easley and O'Hara (1992) suggest one of the keys to understanding the buying and selling interaction is recognizing the adverse selection problem that arises when some traders are informed and others are not and that this interaction can create persistent *order-flow* imbalance that can lead to broader market issues. They logically hypothesise that if informed traders act competitively then certain regularities should persist in their behaviour. For example, the informed would all trade on the same side of the market and their activity would lead to an imbalance in buy / sell volume. A second regularity is that informed traders prefer to trade larger amounts at any given price. And finally, a third regularity is that the informed will continue to trade until prices have adjusted to the new equilibrium (or fair) value.

However, O’Hara (2014) in a related piece of work that followed argues that the continuous evolution of HFT now complicates any attempt to draw meaningful inferences from any market data or previous notions on informed trader identification. O’Hara (2014) suggests that algorithmic trading dictates that it is orders, and not actual trades that now echo the “informed” trader’s true intentions. Algorithms chop up orders, and only some proportion of these orders ultimately turn into actual trades. In some ways O’Hara’s (2014) work suggests that the high frequency era is a great period for empirical researchers involved in market studies given that trading is almost wholly electronic, computerized and that there is a wealth of trading data. However, O'Hara (2014) argues that when practitioners attempt to use this data to match and classify trading activity, any such attempts will be undermined by a range of sequencing and latency problems. O'Hara (2014) suggests that it is tempting to believe that these issues can all be solved by better data sets. Nonetheless, she concludes that this thinking is at best described as naïve and argues that data sets simply cannot keep up with the high frequency world because HFT keeps evolving in advance of any suitable or relevant analysis procedure.

Notwithstanding this, O'Hara (2014) points to her earlier collaborative work (Easley, López de Prado & O'Hara, 2011a) that suggests that time is not a meaningful concept in a computer-driven low latency world and reiterates the argument for using a "volume clock" in the analysis of toxicity risk in high frequency markets. In addition to this, Easley, López de Prado & O'Hara (2011a) argue that persistent adverse selection denotes market toxicity, and market toxicity infers the presence of informed trading that manifests as *order-flow* imbalance that can be estimated when a trading session is divided into meaningful trading periods (non-time based) over which trade imbalances can be evidenced to have economic impact on the provision and quality of *liquidity*. In addition to this, it is important to note that this imbalance is not signed by way of importance, as either buying or selling dominance per se, rather it is the significance attached to a period's imbalance when it is compared against the cumulative distribution function that is noteworthy. Their VPIN (volume synchronized probability of informed trade) tool attempts to estimate the probability of informed trading based on volume imbalance and trade intensity; in the following section we will examine this tool in more detail.

2.5 THE VPIN RESPONSE

Easley, Engle, O'Hara and Wu (2008) argue that a fundamental insight of current *market microstructure* literature is that the order arrival process is informative for subsequent price moves (non-directional) in general, and *order-flow* toxicity in particular. As discussed in the previous sections of this literature review, Easley, López de Prado & O'Hara (2011a) hypothesise that *order-flow* is regarded as toxic when it adversely selects *liquidity* providers or indeed all types of traders who may be unaware (thus uninformed) that they are providing *liquidity* at a loss. Easley, Kiefer, O'Hara & Paperman (1996), under the original estimation approach, defined the idea that flow toxicity can be expressed by the Probability of Informed Trading (PIN) metric. They modelled the estimation of simulated unobservable market parameters and used three different types of Poisson distributions to fit daily buy and sell order executions. This work was extended thereafter by Easley, Engle, O'Hara & Wu (2008) to include the *GARCH* specification to better model the time-varying arrival information rate of informed and uninformed traders. The problem with this approach is that it relied on the somewhat artificial generation of non-observable parameters and numerical methods to describe the *order-flow,* and the process was updated in clock time which does not adequately simulate the observation of volatility clustering in a high frequency environment or the characteristics of actual price change first described by Mandelbrot (1963). To that end, Easley, López de Prado & O'Hara (2011a) presented a new procedure to estimate the Probability of Informed Trading based on volume imbalance and trade intensity, known as the VPIN informed trading metric. VPIN does not require any estimation of non-observable parameters or the application of distribution fitting / numeric methods. The metric is updated in volume-time rather than clock-time using volume buckets to estimate statistically significant trade imbalances (*order-flow* toxicity), and these features are suggested to improve the metrics' overall suitability and predictive power in a high frequency environment. In addition, Easley, López de Prado & O'Hara (2011a) provide simulation evidence via their Monte Carlo experiments that show that VPIN estimates remain accurate for all theoretically possible combinations of parameters and that the VPIN metric does indeed predict short-term, *liquidity*- induced volatility.

The actual procedure to calculate VPIN is not at all onerous. In essence, and as described by Easley, López de Prado & O'Hara (2011a), it involves sampling sequential trades, classified as either a buy or sell, into equal size volume buckets of exogenously defined size over an assumed continuous period. Thereafter, following an exogenously defined volume period, the VPIN metric is calculated and updated following the completion of each new volume period whilst dropping the 1st period in the calculation (akin to a moving average calculation). The VPIN estimate requires the selection of (V), the amount of volume that will be in every period (defined as a bucket size) and (n) the number of buckets used to approximate the expected trade imbalance and intensity. Easley, López de Prado & O'Hara (2011a) use a bucket size (V) equal to 1/50 of the average daily market volume and a bucket number (n) equal to 50. The VPIN metric is then calculated over fifty buckets which in this example and on a day of average volume would correspond to finding a daily VPIN estimate value. The originators of

VPIN argue that sampling by volume buckets allows the trading session to be split into trading periods over which trade imbalances are said to have a meaningful economic impact on *liquidity*.

Notwithstanding this, the problem with VPIN is the very same issue that was briefly discussed in the previous section that underpins and somewhat limits the value of all *market microstructure* research. This is the method used to effectively / accurately assign a trade execution in the market as either a buyer or seller demanded activity. There have been many attempts to formalise a process of buyer or seller induced trade identification. The simplest version, known as the tick test rule is defined by Rosenthal (2008) as a buy initiated transaction if, and only if, the most recent execution is above the previous executed price and as a sell initiated transaction under the opposite premise. Moreover, where an execution is observed at the same price then the execution is classified under the most recent instance of either buy (higher) or sell (lower) activity. According to Perlin, Brooks & Dufour (2011) the tick test rule algorithm has historically been popular with academics because of its economy (only transaction data is required) and simplicity (no intermediary calculations are required). In an attempt to use *order-book* data to advance the accuracy of the tick test rule, Lee & Ready (1991) undertook a different approach when they analysed NYSE trades to infer the initiator of trade by comparing trade prices to quote averages at the time of execution to identify buyer initiated activity when trades were executed at the ask quote, and sell initiated activity when trades were executed at the bid quote. However, they also noted that at the time of writing quotes appeared to lag, so any resulting classification calculations should adapt to consider this difficulty. As pointed out by O'Hara (2014), to a certain degree these sorts of problems persist in today's market due to the speed of execution and participant / general exchange location and latency issues. In addition, the Lee & Ready (1991) method introduces a problem termed as "indeterminacy" (when a traded price does not match any bid or offer, but falls between). However, practitioners using this method to classify trading have traditionally employed the tick test rule to cover off this shortfall when it occurs. More recently Easley, Lopez de Prado & O'Hara (2012) propose a bulk-volume classification algorithm that replaces the need for a discrete tick-by-tick classification process via the use of a continuous classification of standardized price changes. However, according to Chakrabarty, Pascual & Shkilko (2013) this approach solves as many problems as it creates. Clearly, it saves processing time when compared against the collection and manipulation of tick-by-tick data. However, their results suggest that this innate efficiency comes at a significant cost in terms of accuracy; as defined by the term "misclassification". Nevertheless, as pointed out by Easley, Lopez de Prado & O'Hara (2012) it is futile to expect perfect classification of trade data, and studies and results vary in terms of exact or acceptable levels of misclassification, but all academics appear to collectively agree that misclassification is ever-present under any approach. In addition, during their test of various classification approaches, they find that a tick test rule algorithm works reasonably well in the e-mini S&P 500 futures contract, and under their test it correctly classifies 86.43% of the data. Moreover, in their recent paper that compares the bulk-volume classification algorithm performance against the tick test rule, Chakrabarty, Pascual & Shkilko (2013) suggest and present evidence that the tick test rule produces the most accurate estimates of order imbalances and of *order flow* toxicity and therefore provides better VPIN estimates.

In terms of the 2010 Flash Crash, Easley, López de Prado & O'Hara (2011b) show that their VPIN metric for the E-mini S&P 500 futures contract was abnormally high at least one week before the crash, and that this situation further deteriorated several hours before the actual event where the VPIN metric loitered in the 10% to 5% tail of its distribution. According to the CFTC-SEC (2010) official time line, by 14:30 the VPIN metric reached its highest level in the history of the E-mini S&P, and at 14:32 the crash began. Easley, López de Prado & O'Hara (2011b) go on to argue that when *order-flow* toxicity increases, market participants facing significant losses curtail their risks by reducing, or even completely liquidating, their positions. The consequent market il*liquidity* can then have disastrous repercussions for all market participants and suggests that a VPIN contract would serve the dual goal of offering market participants an objective measurement of *order-flow* toxicity, or as a risk management tool to hedge against the risk of being adversely selected possibly avoiding the next Flash Crash.

2.6 CONCLUSIONS

It is indeed a neat and tidy proposition to be able to define market risk and return in the form of two numbers borne out of randomness. However, empirical data and experience suggest that real markets hide a dangerously different concoction to that of assumed uniformity, and adverse events appear to be much more frequent than standard financial models predict. In short, it appears that the standard models massively underestimate this risk and interaction is by observation; clearly more complicated than uniform randomness. As a result, alternative approaches have emerged that acknowledge the increased probability of such events, and advocate the consideration of the quality of *liquidity* provided by the heterogeneity of market participants that in turn forms the fractal market structure that provides both strength (*liquidity*) during uneventful times and creates weakness (il*liquidity*) during times of massive stress.

In terms of HFT, such firms form only a small proportion of market participants, but appear to generate the lion's share of volume. However, as this share has gone up, absolute volume and overall participation has apparently gone down. A clear definition for HFT remains elusory and is now thought to be better understood by addressing any concerns about the methods deployed under such a generalised term and any adverse impact of such activity. The current consensus suggest the impact of HFT is both good (increased *liquidity*, tighter *spreads*) and bad (short lived *liquidity,* higher trading costs for other professional large size traders) and for the most part determined by context (the type of participant you are) and the market condition (normal or stressed) you find yourself in. Given this ambiguity, in an already risky situation, it is not a surprise to find that market participants, big (via *dark pools*) or small (incremental retail withdraw) may be considering business elsewhere or deferring for now. Moreover, the overall aggregated impact of HFT seems to be changing as quickly as the technology / strategic deployment of HFT techniques evolve. Nonetheless, intuition suggests that during times of unparalleled uncertainty, and given that machines are programed by people and are therefore unconscious - pulling the plug when risk / losses escalates is often the only feasible eventuality. As a result, *liquidity*-induced volatility could erupt and in essence make it impossible and potentially disastrous to trade.

It also appears that the provision of *liquidity* and the quality of that *liquidity* are indeed very important variables that determine the confidence levels of participants and the underlying perception of market risk and therefore the risk of active or passive participation. Furthermore, if we consider the assumption that HFT firms engage to some extent in passive market making strategies, as the evidence in this literature review suggests, it would also be obvious recourse to stop and withdraw from this activity if it became loss making due to informed traders adversely selecting such a provision. Moreover, given the market share of HFT participation, a large part of market *liquidity* provision could indeed be ephemeral under such conditions, and it would follow that confidence and risk perceptions would additionally be only short-lived. Unfortunately, the identification of informed trading has been historically problematic, and current research suggests that HFT activity in itself, along with other algorithmic computerised trading, adds to the cloud of uncertainty as to whether informed trading is indeed extant in a market. Nevertheless, the longstanding inference that important information resides hidden within order-flow data, and the inferred link between adverse selection, informed trading and market toxicity (resulting in *order-flow* imbalance), coupled with a non-dependency on order execution / submission speed, and or time as a collaborative factor suggests that the VPIN tool may offer some utility in the probability based identification of such a presence and any coordinated pre-emptive response. Furthermore, this claim is supported by the observed high level of performance of the metric during a real-life (macro) event and backed-up by statistical study and Monte Carlo simulations.

In conclusion, the review of the research herein suggests that, despite the generally accepted ideas of longstanding paradigms, and the ambiguity surrounding HFT and potentially hindering participant behaviour, VPIN may offer market practitioners a tool to help manage their particular risks. Moreover, VPIN appeared to function well as the events of May 6th 2010 unfolded. Notwithstanding this, such events remain unrepeated and make active testing of VPIN (macro events) unwieldly and time dependent. To that end, we are led to scale down the frequency at which we attempt to observe the utility of VPIN to identify or predict *liquidity*-induced volatility in this study. Therefore, it is a micro, rather than macro view and approach that this study prescribes. This supposition follows on from a rational appreciation of the generalised deductions we can extract from this (background) literature review, namely: 1] that markets under stress appear to be best understood as fractal in structure (Peters, 1991) and exhibit instances of self-similarity (Mandelbrot, 1963). 2] HFT activity is significant and present in most markets (Rosenblatt, 2008) and either persistent or transitory (Tong, 2015). 3] The importance of heterogeneity in the provision of *liquidity* (Anderson & Noss, 2013). 4] VPIN measurements have historically predicted *liquidity* induced volatility at a macro level and VPIN parameters are evidenced to be robust for all theoretically possible combinations (Easley, López de Prado & O'Hara, 2011a). 5] VPIN levels equal or above the 5% tail of its CDF appeared to immediately precede the Flash Crash (Easley, López de Prado & O'Hara, 2011a) and 6] the tick test classification rule appears to be reasonably accurate for the ES futures contract (Easley, Lopez de Prado & O'Hara, 2012) and evidence suggests that the tick test rule classification produces the most accurate VPIN estimates for the ES contract when compared with other methods of trade classification (Chakrabarty, Pascual & Shkilko, 2013). It would therefore follow that, given this inherent flexibility, a proportional scaled down observation frequency, built upon the common factors and assumptions described above, ought to logically permit and facilitate a useful “relative” observation origin and “effect” comparison (micro: direct and post hoc and micro to macro). This study will now proceed with the detailed construction of the scaled down (micro) VPIN metric and the quantitative methodology used to discharge the aim of this study.

3.0 METHODOLOGY

3.1 INTRODUCTION

The methodology description to follow outlines, in appropriate sections, a comprehensive explanation of the approach, data, treatment and methods used during this study.

3.2 BACKGROUND TO THE APPROACH

In terms of VPIN algorithmic construction, according to Easley, López de Prado & O'Hara (2011a) the method and parameters selected by practitioners to classify buying / selling volume and the parametrisation (sample and bucket size) of the VPIN metric will affect the resulting estimated absolute value of VPIN. Therefore, this necessitates the comparison of relative levels of VPIN as captured by its cumulative distribution function (CDF) rather than the recognition of any individual or collective absolute value of VPIN. Where relevant, and to aid the direct comparison of this study's results with the body of literature discussed herein, the researcher maintained the use of the original VPIN parameters first described by Easley, López de Prado & O'Hara (2011a). In addition to this, due to its simplicity and robust equivocal historic accuracy (Chakrabarty, Pascual & Shkilko, 2013), this study used the simple tick-test algorithm to assign buy or sell classifications.

As argued above, in terms of a vista for observation the only change made for this particular study was an appropriate proportional rescaling of the overall average volume value used to determine volume per bucket size. In so doing the approach followed the general logic and method ascribed by Easley, López de Prado & O'Hara (2011a) where they apportioned average daily volume as the basis for their parametrisation of the VPIN metric used for the 2010 Flash Crash study (macro environment). Following this principal, this study (micro environment) apportioned a proportional scaled down fractal and HFT sensitive one minute average volume value (*OMAVV*) as the basis of its parameterisation.

3.3 SECONDARY DATA

The secondary tick-by-tick population data that was used in this study to provide the OMAVV data bars was harvested from the ES 06-15 futures contract historic source spanning fifty seven days (see appendix a) of its most active sessions (20th March 2015 through to the 10th June 2015). Using a simple bar number, time and volume extraction algorithm coded by the researcher in C#.Net (see appendix i) the modified secondary data was collected (regular open to the regular close) with the addition of ten minutes added to each close to account for any onward time border excursions during the experiment. Daylight saving time was also considered and adjusted for to capture a consistent observation period (six hours and forty minutes) throughout the duration of the study. Furthermore, it is noteworthy to add that all required calculations and manipulations carried out on the secondary data set were derived from simple addition, subtraction, averaging and standard Microsoft Excel filtering and sorting procedures.

3.4 PRIMARY DATA

The primary data required for this quantitative research study was selected from the abovementioned secondary data via the use of a simple random sample selection process (with replacement) that utilised Microsoft Excel's RANDBETWEEN function and a simple unique identification "KEY" reference to obtain any corresponding values. In all cases of statistical description and inference, the primary data was presented as a randomly derived fifty value single sample for each test and onward conclusion presented.

3.5 DATA ORGANISATION AND TREATMENT METHODS

The first step in the construction of the VPIN metric was to obtain a scaled down volume sample parameter. A random sample of raw volume was selected from the secondary data set (see appendix a) to provide the OMAVV. The VPIN algorithm (Kinlay, 2011) equation (Fig. 3.1) was coded by the researcher in C#.Net (see appendix ii) and configured in accordance with the original non volume specific parameterisation criteria used for the 2010 Flash Crash by Easley, López de Prado & O'Hara (2011a). To that end, a single bucket was configured to capture 1/50th of the OMAVV (denoted as V in the formula), over a rolling window of fifty buckets (*n)* were *V(t)b* were classified as *"BUYS" and V(t)s* were classified as *"SELLS"* using the tick rule, during a given (*n*) period. Under this parameter set the VPIN algorithm was run on the secondary data set to furnish the raw VPIN population CDF data found in appendix (b). A random sample was selected from the raw VPIN CDF data to provide the 95% CDF statistic that would be used in the tests and analysis to follow. Appendix (c) provides the sample data and the resulting statistics used to configure the VPIN metric deployed in this study.

$$VPIN = \sum_{t=0}^{n} \frac{|V(t)b - V(t)s|}{nV}$$

Fig.3.1

Thereafter, a 95% CDF configured VPIN algorithm was run again on the secondary data set (with the addition of the raw data calculation algorithm) using the abovementioned (V) and (n) parameters to furnish the test data found in appendix (d). This data is made up notably of one minute bar periods with a bar number reference, average price per period: $\frac{(high+low+close+open)}{4}$, volume per period, price range per period, VPIN value per period, VPIN sign per period, and the immediate change (delta) in average price from the previous period. In addition to this a unique identification "KEY" was applied to each Microsoft Excel spreadsheet row to facilitate absolute reference to individual values. Finally, utilising Microsoft Excel (RANDBETWEEN function and cumulative calculation capacity) two exogenous time periods were selected at random (one minute and five minutes - in harmony with the fractal micro nature of this study) and for each period the cumulative range, volume and average delta values were calculated and projected to provide suitable post hoc "effect" test data.

Moreover, using Microsoft Excel (filter and sort functions) the test data found in appendix (d) was organised into two (providing sample independence) distinct population sources: a VPIN event - "event^{a}" and a Non VPIN event - "event^{b}". Event^{a}

was defined as an individual observation of VPIN equal to or above 95% of its CDF value (appendix e). All other instances were defined as event^{b} (appendix f). Any repeat of an event^{a} within an OMAVV bar was ignored so to only capture the origin of the VPIN event itself on an OMAVV bar-by-bar basis, and to logically simplify but not nullify any onward directional relationship or non-randomness test procedure. A random sample of event^{a} and event^{b} were selected from the modified secondary data set (appendix e and appendix f) to provide the sampled range, volume, bar and absolute delta data for the event^{a} and event^{b} direct "effect" analysis (i.e. the effect on the bar itself) and range, volume and absolute delta for the selected exogenous time period's "effect" post hoc analysis.

In terms of deriving suitable directional relationship data for analysis - the secondary data set VPIN data found in appendix (e) was additionally manipulated using Microsoft Excel (filter and sort functions) to provide two further independent event observation populations (appendix l). These populations were differentiated by their given VPIN sign, where (using the tick test) and integer value of 2 indicated a buying bias (defined as event^{a2}) and an integer value of 3 indicated a selling bias (defined as event^{a3}). A random sample of event^{a2} and event^{a3} were selected from appendix l to provide Microsoft Excel with lookup reference values for the selected exogenous time periods. The non-absolute average delta direct "effect" analysis (i.e. the effect on the bar itself) and non-absolute average delta value for each exogenous period referenced the "effect" post hoc were recorded in appendix (d).

3.6 STATISTICAL METHODS

This study utilised Microsoft Excel and the commercially available ModelRisk Excel add-in to determine the descriptive statistical points of central tendency, dispersion, relative frequencies (graphical presentation used in appendix a through m) and for all hypothesis testing. Moreover, as the study's sample sizes were greater than thirty, the sample means were used to estimate the population means in all cases. In addition to this, and given that the nature of any difference between event^{a} and event^{b} was unknown prior to this study, and that the null hypothesis H_{0a} assumed no such difference in population mean values - this in turn warranted the use of a two tailed Z-Test (two-sample assuming equal variances) inferential statistical technique to test for evidence of a significant difference between the two means of each independent sample. A significance factor of 5% alpha was used in all hypothesis testing and the test statistic was compared to the equivalent critical two-tail value; an absolute test statistic value greater than the equivalent critical value would indicate evidence of a significant difference. Moreover, the two-sample test assuming equal variances probability value was also compared against the critical 5 % probability value and a test probability statistic value of less than the critical value was also used to indicate evidence of a significant difference.

Furthermore, to look for any evidence of a potential directional bias or relationship between two variables within a given sample of event^{a2} or event^{a3}, a scatter plot was run for each sampled period of event^{a2} and event^{a3} and a linear trend line was fitted showing the R^2 value and line equation to facilitate a visual examination for evidence of a distinct directional relationship within the test data. In addition to this, a Chi

Square test was used to test for any evidence of differences between the observed value and the expected value in the data set sample. A Chi Square probability value of less than the critical 5 % value was used to indicate evidence of a significant difference between observed and expected values. Finally, a Runs test (Bradley, 1968) was used to test each sampled period of $event^{a2}$ and $event^{a3}$ for evidence of randomness. Given the large-sample size used in this study (n^1 and n^2 greater than 10) the test statistic was compared to the standard normal table. An absolute test statistic of greater than 5% and or a probability test result value of less than 5% would indicate evidence of non-randomness. The aforementioned two tests for directional bias and randomness required the discrete conversion of the actual values observed into a binary sequence. Therefore, average delta values greater than zero were denoted as “1” and those values observed equal or less than zero were denoted as “0”. As no difference was expected, given the null hypothesis H_{0b}, the expected values for each denoted value was 25 or, 50% of the sampled value.

3.7 LIMITATIONS

The principal limitation of this study relates to the availability of an error free secondary data set. The secondary data set used in this study was not randomly selected, rather it was selected because it represented the most relevant, recent and error free data set available to the researcher at the time of writing and as a result unknown biases may be present. In addition to this, the random nature of the exogenous “effect” period selection may have by chance precluded the direct observation of important post hoc “effects”. Finally, the selection and use of the tick-test trade classification algorithm may no longer accurately represent the signed nature of *order-flow* in HFT environments and as such could have affected the results obtained. Notwithstanding this, the study now moves on to a summarised presentation of the results from the testing described above.

4.0 RESULTS

4.1 INTRODUCTION

The full test results, including the sample data used, test calculations and an in-depth pictorial and tabular presentation of all results derived from this study can be found in appendix (d), appendix (g) and appendix (m). However, for ease of review, the essential summarised findings are presented below to support the arguments constructed in the study conclusion section to follow.

4.2 THE DIRECT VPIN EFFECT

[Test a] Bar comparison range test: $event^a$ had a mean of 1.21 and a variance of 0.42 and $event^b$ had a mean of 0.76 and a variance of 0.13. The two tailed t-test results for $event^a$ and $event^b$ yielded a t-statistic of 4.2735 and a p value of 4.46366E-05.

[Test b] Bar comparison volume test: $event^a$ had a mean of 5477.48 and a variance of 38698216.91 and $event^b$ had a mean of 2022.24 and a variance of 3049235.08. The two tailed t-test results for $event^a$ and $event^b$ yielded a t-statistic of 3.7813 and a p value of 0.0002682.

[Test c] Bar comparison bar position test: $event^a$ had a mean of 211.72 and a variance of 19345.92 and $event^b$ had a mean of 188.96 and a variance of 12848.61. The two tailed t-test results for $event^a$ and $event^b$ yielded a t-statistic of 0.8969 and a p value of 0.3719.

[Test d] Bar comparison absolute average delta test: $event^a$ had a mean of 0.5012 and a variance of 0.1233 and $event^b$ had a mean of 0.2512 and a variance of 0.0364. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 4.4228 and a p value of 2.52468E-05.

The results for test (a), (b) and (d) above yielded statistically significant t and p values and these test results provide evidence to reject the null hypothesis H_{0a} and support the alternative hypothesis h_{aa} described herein. However, the result for test (c) did not yield statistically significant t and p values and this particular test result does not provide adequate evidence to reject the null hypothesis H_{0a}.

4.3 THE POST HOC VPIN EFFECT

[Test e] One minute post hoc comparison range test: event[a] had a mean of 0.86 and a variance of 0.2964 and event[b] had a mean of 0.77 and a variance of 0.2266. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 0.9287 and a p value of 0.3552.

[Test f] One minute post hoc comparison volume test: event[a] had a mean of 4486.72 and a variance of 42142471.35 and event[b] had a mean of 2149.56 and a variance of 4675657.109. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 2.4152 and a p value of 0.0175.

[Test g] One minute post hoc comparison absolute average delta test: event[a] had a mean of 0.3562 and a variance of 0.13922 and event[b] had a mean of 0.2587 and a variance of 0.0567. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 1.5572 and a p value of 0.1226.

[Test h] Five minute post hoc comparison range test: event[a] had a mean of 1.98 and a variance of 1.2597 and event[b] had a mean of 1.75 and a variance of 0.7078. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 1.134 and a p value of 0.2594.

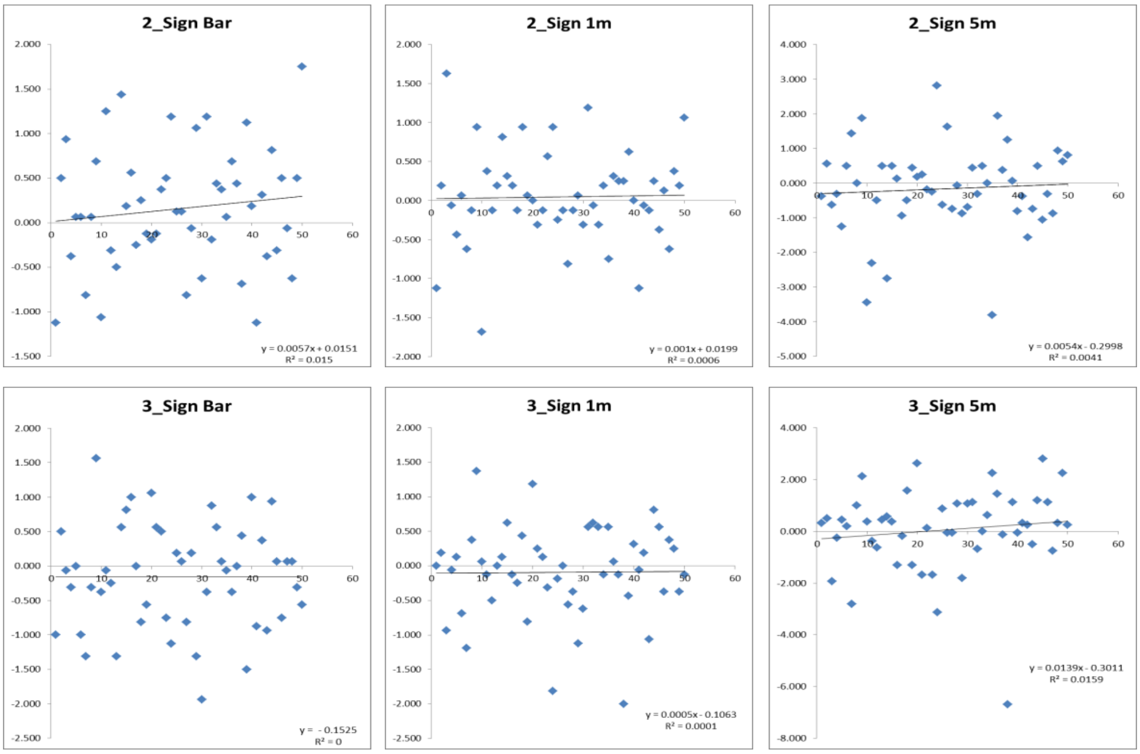
[Test i] Five minute post hoc comparison volume test: event[a] had a mean of 22964.38 and a variance of 360800374 and event[b] had a mean of 12439.56 and a variance of 103730611.70. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 3.4529 and a p value of 0.0008202.

[Test j] Five minute after post hoc comparison absolute average delta test: event[a] had a mean of 0.816 and a variance of 0.680 and event[b] had a mean of 0.781 and a variance of 0.4715. The two tailed t-test results for event[a] and event[b] yielded a t-statistic of 0.2305 and a p value of 0.8181.

The results for test (f) and test (i) above yielded statistically significant t and p values and these test results provide evidence to reject the null hypothesis H_{0a} and support the alternative hypothesis h_{aa} described herein. However, the results for tests (e), (g), (h) and (j) did not yield statistically significant t and p values and these particular test results do not provide adequate evidence to reject the null hypothesis H_{0a}.

4.4 DIRECTIONAL BIAS AND RANDOMNESS

The disbursement of directional average price movement per period sample is given below in the individual scatter plots.



A basic visual examination of the scatter plots above shows similar small R^2 values and an almost horizontal line slope for all periods observed.

[Test k] The p value results yielded from the Chi Square test are given below in Table I.

Table I	2_Sign 1m	2_Sign 5m	3_Sign 1m	3_Sign 5m	2_Sign Bar	3_Sign Bar
Expected [1]	25	25	25	25	25	25
Expected [0]	25	25	25	25	25	25
Actual [1]	25	22	22	29	30	21
Actual [0]	25	28	28	21	20	29
p value	1	0.396	0.396	0.257	0.157	0.257

[Test l] The t and p values yielded from the runs test are given below in Table II.

Table II	2_Sign 1m	2_Sign 5m	3_Sign 1m	3_Sign 5m	2_Sign Bar	3_Sign Bar
Expected [1]	25	25	25	25	25	25
Expected [0]	25	25	25	25	25	25
Actual [1]	25	22	22	29	30	21
Actual [0]	25	28	28	21	20	29
t Stat	0.5715	0.1044	0.1856	1.0681	1.4896	0.1056
p Value	0.5676	0.9168	0.8527	0.2854	0.1363	0.9158

The results of tests (k) and (l) above did not yield a statistically significant p value for the Chi Square test or t or p values for the Runs test and these results provide no evidence to reject the null hypothesis H_{0b} described herein.

In summary, the results above provide significant evidence to support the notion that a VPIN metric attuned to a micro vista appears on the average to exhibit an immediate larger than normal response in terms of *market dynamics* without predicting the sign of the price change itself. In short, this result is analogous to the idea and result characteristics of toxicity-induced volatility as originally described by Easley, López de Prado & O'Hara (2011a) for a VPIN metric attuned to the macro environment. Moreover, the evidence of a persistent volume increase in the exogenous periods examined warrants further discussion, and in the section to follow the researcher will close this study by drawing reasonable inference, conclusions and make recommendations derived from the results presented above.

5.0 STUDY CONCLUSIONS

The results above are mixed but nonetheless interesting. The direct bar related tests (a), (b) and (d) present strong evidence to support the argument that a VPIN bar is indeed wholly different in terms of its *market dynamics*, characterised by a significant increase in size (range, average absolute delta and volume) when compared to a Non VPIN bar within the originating one minute period. In addition to this, the likely location of a VPIN bar given in test (c) appears to be no different than that of a Non VPIN bar, and as such this evidence suggests that a VPIN bar could happen anywhere, and no immediately obvious significant predictable pattern appears present in the data.

Furthermore, if we exclude the volume tests (f) and (i) the results of the range and absolute average delta tests (e), (g), (h) and (j) suggest that the persistence of any such significant differences observed in tests (a) and (d) disappear over a post hoc event period of one and five minutes, and as such no evidence was found to support a difference between a VPIN originated event or a Non VPIN originated event in these cases. However, this is not the case when we examine the volume tests (f) and (i) as these tests present strong evidence to support the argument that the originating VPIN bar's increase in volume persists (test b) when examined at the close of post hoc one and five minute periods.

In terms of directional bias, no evidence of predicable onward directional price movement was found from a visual inspection of the scatter plots or from tests (k) and (l) carried out to determine directional relationships and non-randomness. It would therefore appear that the particular sign of the originating VPIN event has no effect on the onward price direction either on the bar itself or at the close of post hoc one and five minute periods. Moreover, the test result recorded in test (l) provides strong evidence to support the assertion that the relationship between signed events and the resulting direct and post hoc directional price movement is in fact random.

Given the above, and more generally, it could be argued that the change in *market dynamics* evidenced and discussed above for tests (a), (b) and (d) represent an increase in short term market risk if volatility is viewed as maintaining orthodoxy (i.e. that volatility is a good proxy for risk), as described by Mandelbrot & Hudson (2005). Notwithstanding this, outside of the persistence of volume this observed change in *market dynamics* appears to vanish following the close of the originating bar. Moreover, given that the evidence suggests that volume appears to persist in test (f) and (i) then, according to Tong's (2015) assertion, the persistence of this activity would be judged as less likely to be HFT driven. However, this assertion should be treated with caution as the extent of the volume persistence remains unknown (see recommendations to follow) and this manifestation could be for all intents and purposes the relative ephemeral "boom" (increase in *liquidity*) before the "bust" (il*liquidity*). Furthermore, evidence from tests (k) and (l) suggest that a VPIN event's originating signed bias does not appear to convey any hidden information as to future price direction and market participants should therefore not use the originating VPIN sign bias, derived from buying and selling trade classification, as an inference for post hoc directional intent. To that end, it appears that any attempts to unearth signal denotation within underlying trade information, as described by O'Hara (2014), to form the basis of directional inference appear fruitless and as such, directional randomness, as described by Fama (1965), Kendell (1953) and Davis & Etheridge (2006), should be expected.

In addition to this, given the context (micro environment), and if we offer the abovementioned collective on the average consequence of a micro VPIN event as a post hoc VPIN micro event forecast of short term *market dynamics* it could be argued that a passive strategy (SEC, 2014) should thrive at the advent / during such conditions, as long as such volume and non-directional price randomness persists. On the contrary, other strategies discussed herein such as an aggressive directional bias (SEC, 2014) would logically not be immediately well suited to the *market dynamics* of a post micro VPIN event. Furthermore, given that a micro VPIN event denotes an immediate self-similar reaction, in terms of *market dynamics* when compared to a macro environment VPIN event (as described by Easley, López de Prado & O'Hara (2011a)), it could be argued that the usefulness of VPIN as a predictive real-time risk management tool, or as a measurement of short term market risk is in effect subject to the observer's fractal perspective. To that end, it follows that it is the relevant scaling of the VPIN metric attuned to an agent's particular fractal vista and strategic intent; where arguably the greatest specific VPIN related information utility resides. This logical end would appear to support the FMH work of Peters (1991) and the importance of persistent levels of investor heterogeneity within a given market's structure. In turn, if reliable, this presence should beget investor confidence (via a perceived greater understanding of risk) and facilitate the highest quality level of *liquidity*, as described by Warsh (2007). Moreover, it also follows that an appreciation and maintained "careful eye" on the actions of participants with different perspectives (i.e. contextual) might help individual agents adjust their particular risk appetite and strategic intent to avoid instances of fractal fragility faults described by Anderson & Noss (2013) that can cogently result in homogeneous mass panic events, as depicted by Haldane (2011).

Finally, it must be stressed that the actual degree to which greater volume persists as evidenced in test (f) and test (i) following a micro VPIN event remains unknown as it fell outside of the scope of this experimental study.

5.1 RECOMMENDED FURTHER STUDY

Given the above conclusions and end note, any further study should attempt to ascertain the actual on the average duration persistence of the observed increased levels of volume, and if indeed such a persistence does in turn result in a predicable micro *liquidity* induced volatility event (ergo ill*iquidity* crash) at some subsequent point in time.

6.0 REFERENCES

Admati, A., & Pfleiderer, P. (1988) *A Theory of Intraday Patterns: Volume and Price Variability.* Review of Financial Studies. Society for Financial Studies vol. 1(1), pages 3-40.

Anderson, N., & Noss, J. (2013) *The Fractal Market Hypothesis and its implications for the stability of financial markets*. The Bank of England. Paper No. 23.

Bodie, Z., Kane, A. & Marcus, A. (2008) *Investments.* McGraw-Hill. New York.

Bradley, J. (1968) *Distribution-Free Statistical Tests*. Management Science Vol. 16, No. 1, Chapter 12, Theory Series. September. pp. 141-143.

CBS NEWS. (2011) *Wall Street: The speed traders*. [online]. Available from: http://www.cbsnews.com/videos/wall-street-the-speed-traders-50105949 [Accessed 25th November 2015].

CFTC. (2010) *Proposed Rules.* Federal Register, Vol. 75, No. 13. January. pp. 33198-33202.

Chakrabarty, B., Pascual, R. & Shkilko, A. (2013) *Trade Classification Algorithms: A Horse Race between the Bulk-based and the Tick-based Rules*. Working Paper, SSRN, December.

CME. (2003) *Using the S&P 500 Stock Index Futures and Options*. CME Open Interest Paper 49 pdf.

CNBC. (2010) *Flash Crash*. [online]. Available from: http://video.cnbc.com/gallery/?play=1&video=1487044534 [Accessed 11th November 2014].

Davis, M., & Etheridge, A. (2006) *Louis Bachelier's Theory of Speculation*. Princeton: Princeton University Press. New Jersey.

Easley, D., Engle, R., O'Hara, M., & Wu, L. (2008) *Time-Varying Arrival Rates of Informed and Uninformed Trades.* The Journal of Financial Econometrics. pp. 171-207.

Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. (1996) *Liquidity, Information, and Infrequently Traded Stocks.* The Journal of Finance. Volume 51. September. pp. 1405-1436.

Easley, D., M. López de Prado., & M. O'Hara. (2011a) *Flow Toxicity and Liquidity in a High Frequency World*. Review of Financial Studies *25*, 1457–1493.

Easley, D., M. López de Prado., & M. O'Hara. (2011b) *The Microstructure of the Flash Crash*. The Journal of Portfolio Management. Vol. 37, No. 2, Winter, 118-128.

Easley, D., M. López de Prado., & M. O'Hara. (2012) *Bulk classification of trading activity*. Johnson School Research Paper Series #8-2012.

Evans, M., & Lyons, R. (2004) *Frequently Asked Questions About the Micro Approach to FX*. [online]. Available from: faculty.haas.berkeley.edu/lyons.

Fama, E. (1965) *Random Walks In Stock Market Prices*. Financial Analysts Journal 21 (5): 55–59.

French, K. (2008) *Presidential address: The cost of active investing*. The Journal of Finance 63.

FT.com. (2015) *Dark pools definition*. [online]. Available from: http://lexicon.ft.com/Term?term=dark-pools [Accessed 4th October 2014].

Haldane, A. (2011) *The race to zero*. [online]. Available from: www.bankofengland.co.uk/publications/Documents/speeches/2011/speech509.pdf.

Iati, R. (2009) *High-Frequency Trading Technology*. TABB Group.

Investopedia. (2015a) *GARCH definition.* [online]. Available from: http://www.investopedia.com/terms/g/generalalizedautogregressiveconditionalheteroskedasticity.asp. [Accessed 5th May 2015].

Investopedia. (2015b) *Market Dynamics definition.* [online]. Available from: http://www.investopedia.com/terms/m/market-dynamics.asp [Accessed 5th April 2015].

Jeria, D., & G. Sofianos. (2008) *Passive orders and natural adverse selection*. Street Smart, Issue33, September 4th.

Kane, D., Lui, A., & Nguyen, K. (2011) *Analyzing an Electronic Limit Order Book*. The R Journal, Vol. 3, No. 1. June. pp. 64-68.

Kendell, M. (1953) *The Analysis of Economic Time Series, Part I: Prices*. The Journal of the Royal Statistical Society 96.

Kinlay, J. (2011) *Measuring Toxic Flow for Trading & Risk Management*. [online]. Available from: http://jonathankinlay.com/index.php/2011/09/measuring-toxic-flow-for-trading-risk-management/ [Accessed 11th May 2015].

Kirilenko, A., Kyle, A., Samadi, M., & Tuzun, T. (2014) *The Flash Crash: The Impact of High Frequency Trading on an Electronic Market*. [online]. Available from SSRN: http://ssrn.com/abstract=1686004 or http://dx.doi.org/10.2139/ssrn.1686004.

Lee, C., & Ready, M. (1991) *Inferring Trade Direction From Intraday Data*. The Journal of Finance 46:2(1991), 733–746.

Lintner, J. (1965) *The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital budgets*. The Review of Economics and Statistics. Vol.47. No 1. February. pp.13-37.

Mandelbrot, B. (1963) *The Variation of Certain Speculative Prices*. The Journal of Business, Vol. 36, No. 4. October. pp. 394-419.

Mandelbrot, B., & Hudson, R. (2005) *The (Mis)behaviour of Markets: A Fractal View of Risk, Ruin and Reward*. Profile Books LTD.

Mandelbrot, B., & Taylor, H. (1967) *On the Distribution of Stock Price Differences*. Operations Research, Vol. 15(6), pp. 1057–1062.

Markowitz, H. (1952) *Portfolio Selection.* The Journal of Finance, vol. 7, no. 1 (March):77–91.

Mossin, J. (1966) *Equilibrium in a Capital Asset Market.* Econometrica. Vol. 34. No. 4. October. pp. 768-783.

NASDAQ. (2015) *Price discovery process definition.* [online]. Available from: http://www.nasdaq.com/investing/glossary/p/price-discovery-process. [Accessed 12th April 2015].

O'Hara, M. (2014) *High Frequency Market Microstructure*. April 2014. Johnson Graduate School of Management, Cornell University.

Perlin, M., Brooks, C., & Dufour, A. (2011) *On the Performance of the Tick Test*. Working Paper. Available from: http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2012-Barcelona/papers/EFMA2012_0020_fullpaper.pdf.

Peters, E. (1991) *Fractal Market Hypothesis*. New York: John Wiley & Sons

Rosenblatt. (2008) *Trading Talk - Market Structure analysis & Trading strategy* [online]. Available from: http://www.rblt.com/uploads/documents/newsletter/Trading%20Talk%20081408_Security.pdf [Accessed 16th April 2015].

Rosenthal, D. (2008) *Modelling Trade Direction.* University of Illinois at Chicago - Department of Finance.

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

SEC. (2014) *Equity Market Structure Literature Review - Part II: High Frequency Trading*. [online].Available from: https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf [Accessed 16th April 2015].

SEC & CFTC. (2010) *Findings Regarding the Market Events of May 6, 2010*. Joint Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC) Advisory Committee on Emerging Regulatory Issues.

Sharpe, W, F. (1964) *Capital asset prices: A theory of market equilibrium under conditions of risk*. The journal of finance. Vol. XIX. No.3. September.

Tong, L. (2015) *A Blessing or a Curse? The Impact of High Frequency Trading on Institutional Investors*. European Finance Association Annual Meetings 2014 Paper Series. [online]. Available from SSRN: http://ssrn.com/abstract=2330053 or http://dx.doi.org/10.2139/ssrn.2330053 [Accessed 12th May 2015].

Vishwanath, R. & Krishnamurti C. (2009) *Investment Management: A Modern Guide 13 to Security Analysis and Stock Selection*. DOI: 10.1007/978-3-540-88802-4 2, 2009 Springer-Verlag Berlin Heidelberg.

Warsh, K. (2007) *Market Liquidity: Definitions and Implications*. Speech to the Institute of International Bankers Annual Washington Conference in Washington, D.C.

7.0 APPENDIX

Appendix (i) Volume_Msc.pdf (see electronic submission)

Appendix (ii) Vpin_Msc.pdf (see electronic submission)

Appendix (a) Raw Volume (see electronic submission)

Appendix (b) Raw Vpin (see electronic submission)

Appendix (c) Configuration (see electronic submission)

Appendix (d) Test Data (see electronic submission)

Appendix (e) Vpin Data (see electronic submission)

Appendix (f) Non Vpin Data (see electronic submission)

Appendix (g) Hypothesis Test (see electronic submission)

Appendix (h) Bar Sample Charts (see electronic submission)

Appendix (i) Sample Tables (see electronic submission)

Appendix (j) Vpin Filter (see electronic submission)

Appendix (k) Non Vpin Filter (see electronic submission)

Appendix (l) Direction (see electronic submission)

Appendix (m) Runs Test (see electronic submission)