THE MICROSTRUCTURE OF THE 'FLASH CRASH'
Flow toxicity, liquidity crashes and the Probability of Informed Trading

David Easley
dae3@cornell.edu

Marcos M. López de Prado
marcos.lopezdeprado@tudor.com

Maureen O'Hara
mo19@cornell.edu

November 19th, 2010

ABSTRACT

The 'flash crash' of May 6th 2010 was the second largest point swing (1,010.14 points) and the biggest one-day point decline (998.5 points) in the history of the Dow Jones Industrial Average. For a few minutes, $1 trillion in market value vanished. In this paper, we argue that the 'flash crash' is the result of the new dynamics at play in the current market structure. We highlight the role played by order toxicity in affecting liquidity provision, and we show that a measure of this toxicity, the Volume-Synchronized Probability of Informed Trading (VPIN)*, captures the increasing toxicity of the order flow in the hours and days prior to collapse. Since the 'flash crash' might have been avoided had liquidity providers remained in the marketplace, a solution is proposed in the form of a 'VPIN contract' which would allow them to dynamically monitor and manage their risks.

Keywords: Flash crash, liquidity, flow toxicity, market microstructure, probability of informed trading, VPIN.

JEL codes: C02, D52, D53, G14.

1.Introduction

The events of May 6th continue to reverberate through the financial markets. The "flash crash" featured the biggest one-day point decline (998.5 points) in the history of the Dow Jones Industrial Average. Futures were also affected, with the price of the E-mini S&P 500 futures collapsing by 5% between 2:30pm[1] and 2:45pm, on top of the 2.97% it had already retreated intraday. This price drop was accompanied by an unusually large volume of transactions (see **Figure 1**). Between 2:30pm and 3:00pm, in excess of 1.1 million contracts were exchanged in E-mini S&P 500 June 2010 futures alone.[2] Across both futures and equity markets, "there was a complete evaporation of liquidity in the marketplace".[3]
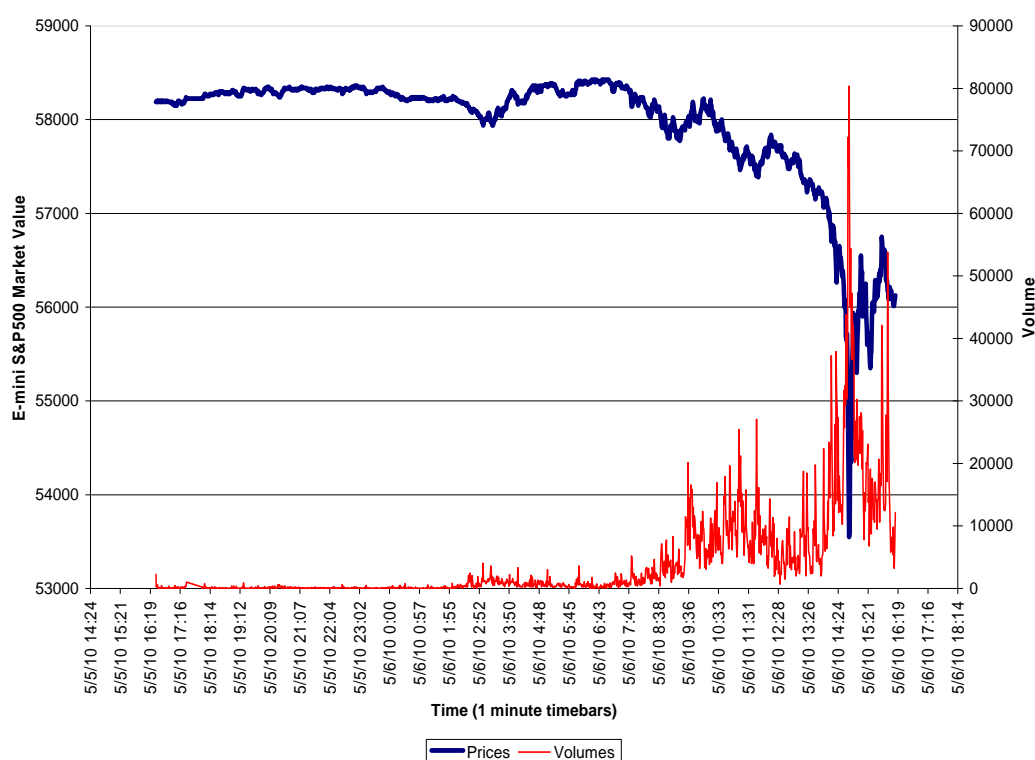

Figure 1 – Futures price and volume during the crash.

Observers were quick to offer explanations for the flash crash:

1. Minutes after the crash, there was speculation that a "fat-finger trade" in 'Procter & Gamble' had triggered a cascade of stop loss orders. This explanation was short-lived, as E-mini S&P500 tick data demonstrates that the market was already down by the time 'Procter & Gamble' stock plummeted.[4]

---

[1] All times are Eastern Daylight Time.

[2] The 'flash crash' occurred despite the CME's stop logic protocol working as expected. The CFTC and SEC report states: *"Starting at 2:45:28pm, CME's Globex stop logic functionality initiated a brief pause in trading in the E-mini S&P 500 futures. This functionality is initiated when the last transaction price would have triggered a series of stop loss orders that, if executed, would have resulted in a cascade in prices outside a predetermined 'no bust' range (6 points in either direction in the case of E-minis)"*. See *"Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues"*, May 18th, 2010, pg. 6-7.

[3] See testimony of Chris Nagy, TD Ameritrade Holding Corp's Managing Director of Order Routing, cited in *"Panel urges big thinking in 'flash crash' response"*. Reuters, August 11th, 2010.

[4] Phillips, Matt. *"SEC's Schapiro: Here's My Timeline of the Flash Crash"*. The Wall Street Journal. May 20th, 2010.

2

2. Technical reporting difficulties at NYSE and ARCA, as well as delays in the consolidated tape, were alleged to have contributed to the market's free fall.[5]
3. Some analysts blamed currency movements, in particular changes in the U.S. Dollar/Japanese Yen exchange rate.[6]
4. The Wall Street Journal suggested a large purchase of put options by the hedge fund 'Universa Investments' may have been the primary cause.[7]
5. Similar speculation centered on a sale of 75,000 E-Mini contracts by Waddell & Reed as causing the futures market to dislocate.[8]
6. Nanex argued that a predatory practice called "*quote stuffing*" forced competitors to slow down their operations in order to catch up with the overwhelming amount of data to be processed by their algorithms.

The CFTC-SEC Staff Report on the market events of May 6 identifies automated execution of a large sell order in the E-mini contract as precipitating the actual crash. What then followed was "two liquidity crises – one at the broad index level in the E-mini, the other with respect to individual stocks."[9] This generalized severe mismatch in liquidity was exacerbated by the withdrawal of liquidity by some electronic market makers and by uncertainty about, or delays in, market data affecting the actions of market participants.

This paper presents additional evidence to support this liquidity explanation. Our analysis shows that the liquidity problem was slowly developing in the hours and days before the collapse. Just prior to the inauspicious trade, volume was high and unbalanced, but liquidity was low. We present evidence that during this period order flow was becoming increasingly toxic for market makers. In a high frequency world, order flow toxicity can cause market makers to leave the market, setting the stage for episodic illiquidity. In other research (see Easley, López de Prado and O'Hara (2010)), we develop a technique that allows us to measure this order flow toxicity. In this paper, we use this new measure to address two flash crash-related questions of particular relevance for portfolio managers: Is this anomaly likely to occur in future? And if so, are there any tools to monitor in real-time the likelihood of it occurring again? In the next section, we argue that the answer to the first question is yes. The rest of the paper will be dedicated to answering the second question.

2. New Trends in Market Structure

Since 2009 high-frequency trading firms, which represent approximately 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.[10] The CFTC, citing research by Rosenblatt

---

[5] See AI5000's article of August 24th 2010 by Joe Flood: *"NYSE Confirms Price Reporting Delays That Contributed to the Flash Crash"*.

[6] *"The Yen Did It?"*, Seekingalpha.com, May 7th 2010

[7] *"Did a Big Bet Help Trigger 'Black Swan' Stock Swoon?"* The Wall Street Journal, May 11, 2010.

[8] The CME Group disputes this as a cause, noting that the order for 75,000 contracts was entered in relatively small quantities and in a manner designed to dynamically adapt to market liquidity by participating in a target percentage of 9% of the volume executed in the market. The order was completed in approximately twenty minutes, with more than half of the participant's volume executed as the market rallied – not as the market declined. See *"Statement on the Joint CFTC/SEC Report Regarding the Events of May 6"*.

[9] See *"CFTC-SEC Findings Regarding the Market Events of May 6, 2010"* of September 30, 2010, p. 3

[10] See Iati (2009).

Securities, estimates that high-frequency trading constitutes approximately 35% of U.S. futures markets volume, and that its share is expected to grow to 60% by the end of 2010.[11] This increased share of high-frequency traders has not been accompanied by an increase in absolute volume. On the contrary, since 2009 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.

Many of these high frequency firms are in the business of *liquidity provision*, acting as *market makers* to *position takers*.[12] Liquidity provision is a complex process, as position takers may know more about the future direction of prices than do market makers. But position takers also need liquidity, or someone to take the other side of their trade, and market makers can profit by earning the spread provided they can control their position risk. Most liquidity providers do not seek to make a directional bet, but instead participate on both sides of the book in an attempt to maximize the turnover of their inventory. Indeed, the typical high frequency market maker turns over his or her inventory 5 or more times a day, explaining how high frequency firms have come to have such a high share of trading volume. These market makers also seek to hold very small or even zero inventory positions at the end of the session.[13] This short holding period, combined with very small inventories, allows market makers to operate intraday with very low capital, essentially using their speed of trading to control their position risk.

Providing liquidity in a high frequency environment introduces new risks for market makers. When order flows are essentially balanced, high frequency market makers have the potential to earn razor thin margins on massive numbers of trades. When order flows become unbalanced, however, market makers face the prospect of losses due to adverse selection. The market makers' estimate of the toxicity (the expected loss from trading with better informed counterparties) of the flow directed to them by position takers now becomes a crucial factor in determining their participation. If they believe that this toxicity is too high, they will liquidate their positions and leave the market.

In summary, we see three forces at play in the current market structure:
- **Concentration of liquidity provision** into a small number of highly specialized firms.
- **Reduced participation of retail investors** resulting in increased toxicity of the flow received by market makers.
- **High sensitivity of liquidity providers to intraday losses**, as a result of the liquidity providers' low capitalization, high turnover, increased competition and small profit target.

---

[11] *"CFTC proposals for new rules governing co-location of HFT servers and exchange servers"*: http://www.cftc.gov/ucm/groups/public/@lrfederalregister/documents/file/2010-13613a.pdf

[12] Kirilenko, Kyle, Samadi and Tuzun (2010) use transaction level data sorted by the type of trader to make the point that high-frequency traders typically act as market makers, but that during the flash crash their trading exacerbated the crash. Those authors differentiate between different types of market makers, but we use the more expansive definition to include all market makers using high frequency trading strategies.

[13] The CFTC-SEC report (pg. 15) found that "net holdings of HFTs fluctuated around zero so rapidly that they rarely ever held more than 3000 contracts long or short on that day."

These forces, combined with the ability of high frequency traders to vanish quickly from the market, portend episodes of sudden illiquidity.

3. Liquidity on May 6th: Market Makers vs. Position Takers

Although May 6th 2010 was the third-highest-volume-day in the history of E-mini S&P 500 futures, there is consensus in categorizing it as an extremely illiquid day. Indeed, the CFTC-SEC report stresses that "high trading volume is not necessarily a reliable indicator of market liquidity."[14] That volume and liquidity need not be congruent is a reflection of the delicate symbiosis between *market makers* and *position takers* in a high frequency world.

In the following analysis, we evaluate whether market makers may have withdrawn from the marketplace during the events of May 6th 2010 as a result of an accumulation of losses and/or extraordinary flow toxicity, inflicted by position takers, in the preceding days and hours. To investigate this hypothesis, we apply a measure of order toxicity (the VPIN metric) developed in Easley, López de Prado and O'Hara (2010) to show how order flow became increasingly toxic over the day. We also show that movements in this VPIN metric measure foreshadowed the actual crash, providing as it were an early-warning of liquidity problems.

In a high frequency framework, both time and information have different meanings than in more standard microstructure models. Because trades take place in mili-seconds, trade time rather than clock time is the relevant metric to use in sampling the information set. Trade time can be measured by volume increments, and the VPIN metric is calibrated using pre-set volume buckets. Similarly, because market makers hold positions for very short periods, information events can reflect asset-related news and/or portfolio-related news. For example, in a futures setting, information that induces traders to all hedge in one direction can portend future movements in futures prices, and thus prove toxic to market makers on the other side of those trades. Easley, López de Prado and O'Hara (2010) develop a microstructure model to capture these high frequency dimensions, and this model produces the VPIN metric we use here to measure toxicity.

For any time period, the VPIN metric is the ratio of average unbalanced volume to total volume in that period. Heuristically, the VPIN metric measures the fraction of volume-weighted trade that arises from informed traders as the informed tend to trade on one-side of the market and so their activity leads to unbalanced volume (either more buy volume than sells volume or the reverse). In periods in which there is a lot of information-based trade, the VPIN metric will be large. During these periods, market makers are on the wrong side of the trade with the informed (i.e. buying when prices are moving down, and conversely), and so they will accumulate or lose inventory on the wrong side of the market. As market prices move, market makers will make losses on their positions. If these losses accumulate, we would expect market makers to undo their positions, thus adding to the imbalance in trade and potentially leading to a crash.

---

[14] CFTC-SEC Findings Regarding the Market Events of May 6, 2010, see pg. 3.

Electronic copy available at: https://ssrn.com/abstract=1695041

4. Estimating Order Toxicity: The VPIN metric

The methodology for estimating the VPIN metric is developed in Easley, López de Prado and O'Hara (2010), and we refer the interested reader to that paper. In this paper we focus on the VPIN metrics for the E-mini S&P 500 futures contract for the time period between January 1st 2008 and October 30th 2010. We compute the average daily volume for this contract and then estimate a VPIN value for each time period in which one-fiftieth of this volume is traded. This procedure results in an average of fifty VPIN metric values per day, but on very active days the VPIN metric will be updated much more frequently than on less active days.

The following chart shows the evolution of E-mini S&P 500 (red line, expressed in terms of market value) and the VPIN metric (green line). Two features of the data are striking. First, the VPIN metric is generally a stable process. Second, the VPIN metric reached its highest level for this sample on May 6th 2010, providing quantitative support to the qualitative assertion that *"liquidity evaporated"*.
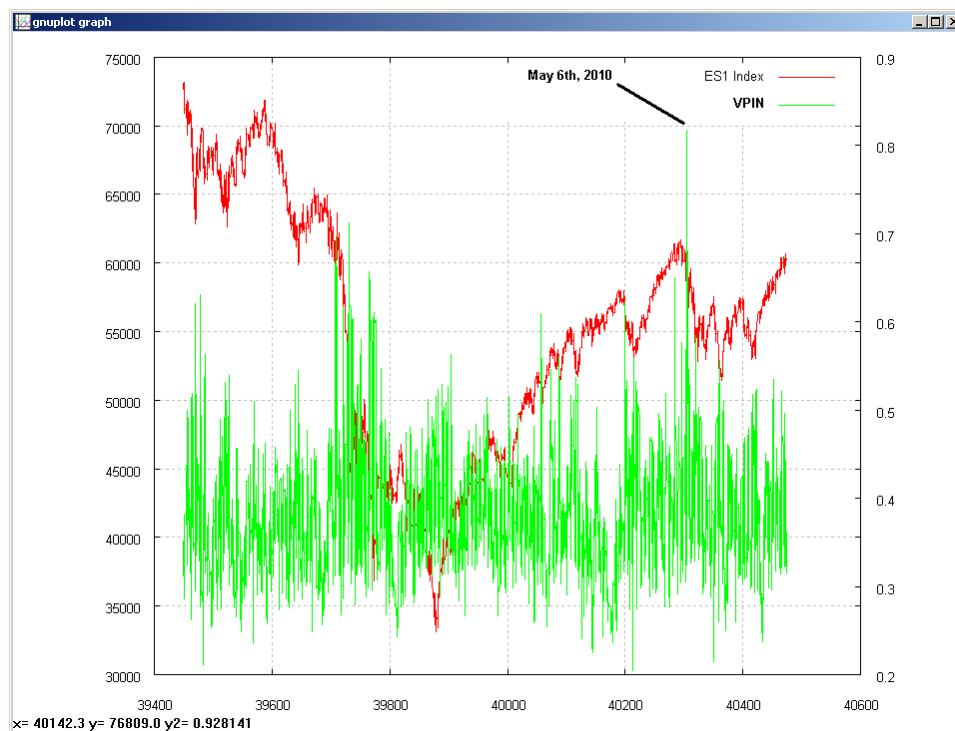

Figure 2 - VPIN metric between January 1st 2008 and October 30th 2010

Before analyzing the VPIN metric in the crash period, it is useful to consider more carefully some statistical properties of the VPIN metric. **Figure 3** plots the empirical distribution of the VPIN metric estimates for the entire sample period. This distribution can be closely approximated by a log-normal, and we will use the cumulative distribution function of the fitted log-normal to provide a measure of how unusual a particular level of the VPIN metric is relative to what is normal for the E-mini S&P 500 futures contract. For example, 80% of the VPIN metric estimates are below 0.44 (i.e. $CDF(0.44) = 0.8$), so VPIN metric estimates of more than 0.44 occur in only about 20% of the estimates.
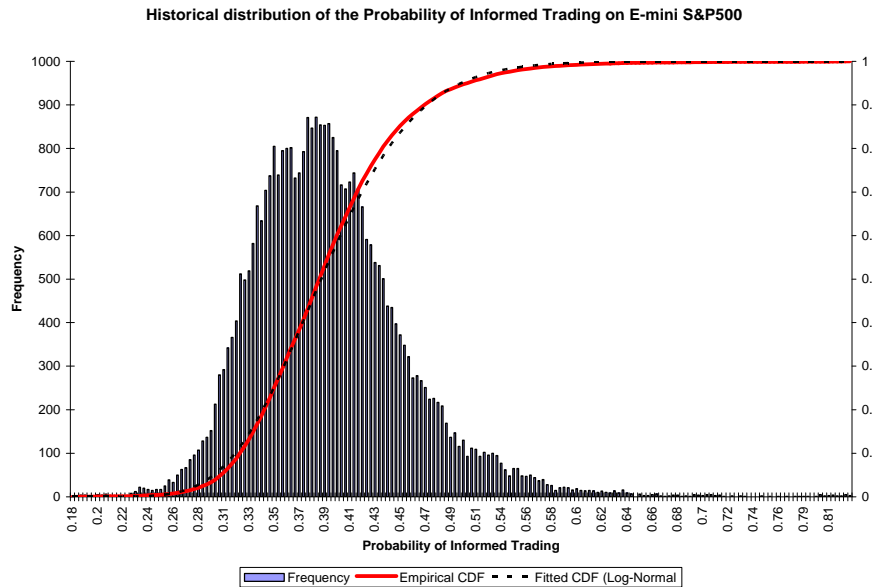
6

Historical distribution of the Probability of Informed Trading on E-mini S&P500

Figure 3 - The empirical CDF of the VPIN metric
as fitted through a log-normal distribution

5. Measuring Order Flow Toxicity before the Crash

We now turn to the behavior of the VPIN metric in the hours and days prior to the crash. The events of May 6th have been categorized as a liquidity-induced crash, so the VPIN metric should have reflected the increasing toxicity of order flow in the market and its consequent effect on liquidity providers. We focus on the behavior of the E-mini future VPIN metric, but we note that other futures contracts were also affected by events on May 6th and VPIN results for those contracts are similar.[15]

Our first observation is that the VPIN metric for the E-mini S&P 500 future was abnormally high at least one week before the flash crash. **Figure 4** shows the value of the E-mini S&P 500, the value of the VPIN metric, and, for each estimated value of the VPIN metric, the fraction of the empirical distribution that is less than this value of the VPIN metric, CDF(VPIN). This measure of the likelihood of the VPIN metric being less than or equal to the current value is volatile, but it was generally unusually high during the week before the flash crash. Such behavior must have placed market makers on the alert, as the toxicity of flow directed to them was gradually becoming more unpredictable.

---

[15] See Easley, et al (2010) for analysis of VPIN behavior in currency futures, interest rate futures, metal futures and energy futures.
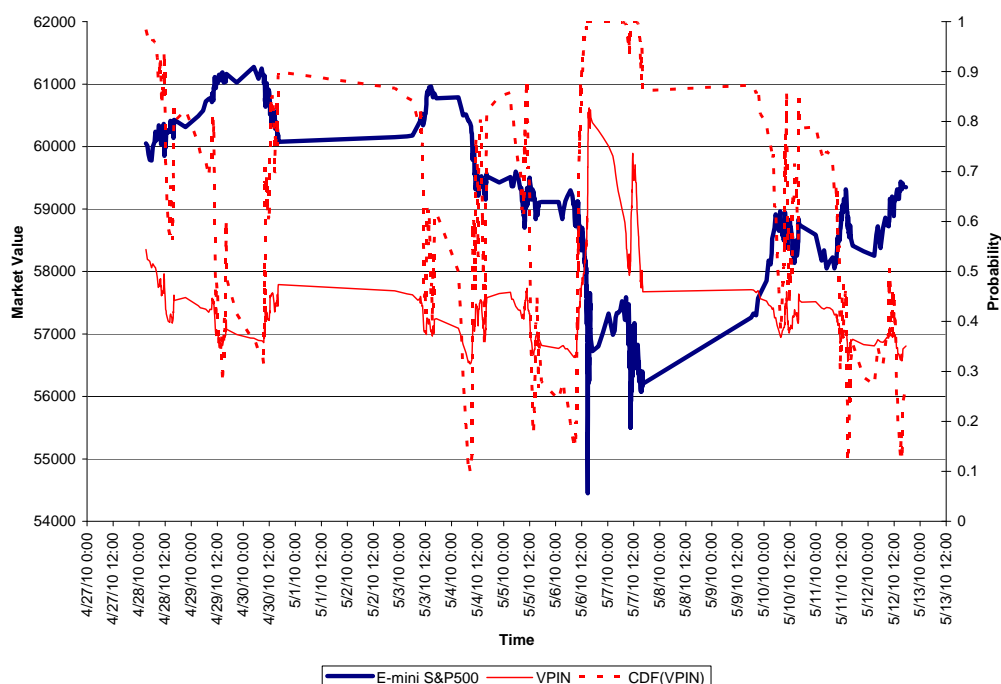
Figure 4 – E-mini S&P 500's VPIN metric one week before and after the flash crash

Our second observation is that this situation worsened (from the point of view of liquidity providers) several hours before the crash. This is illustrated in **Figure 5** which shows that by 11:55am on May 6th, the realized value of the VPIN metric was in the 10% tail of the distribution (it exceeded a 90% CDF(VPIN) critical value). By 1:08pm, the realized value of the VPIN metric was in the 5% tail of the distribution (over a 95% CDF(VPIN)). By 2:30pm the VPIN metric reached its highest level in the history of the E-mini S&P500. At 2:32pm the crash begins according to the CFTC-SEC Report time line.
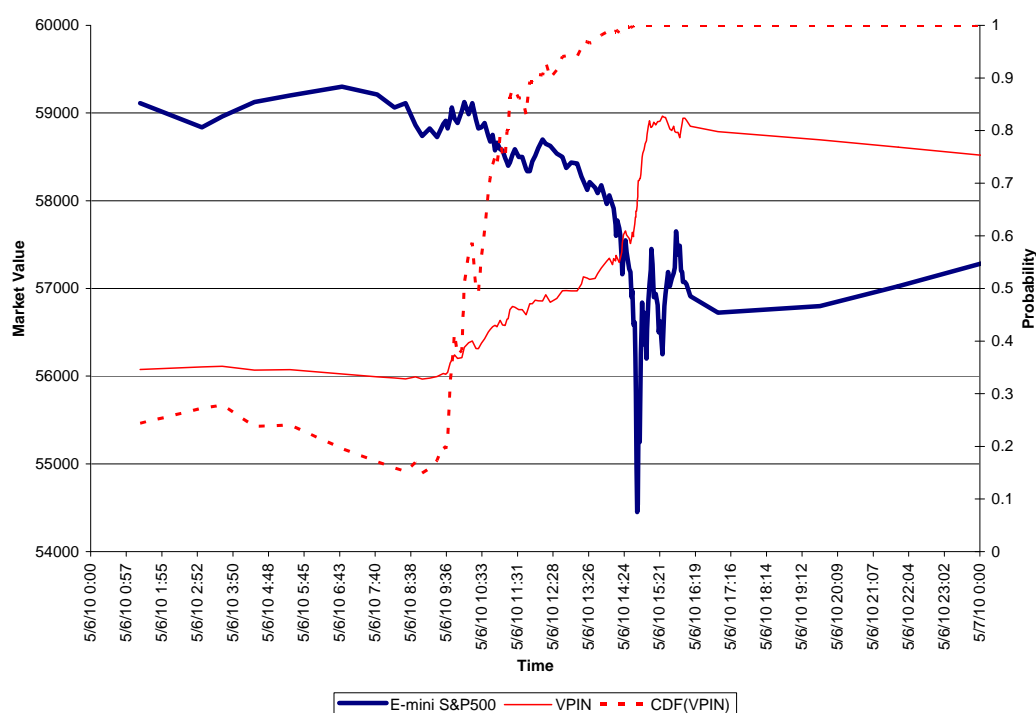

Figure 5 – E-mini S&P 500's VPIN metric on May 6th

8

As market makers were being overwhelmed by toxic flow (measured in terms of unusually high levels of the VPIN metric), many high frequency firms decided to withdraw from the market[16]. According to the CFTC-SEC report, *"HFTs, therefore, initially provided liquidity to the market. However, between 2:41 and 2:44p.m., HFTs aggressively sold about 2,000 E-Mini contracts in order to reduce their temporary long positions"*.[17] The report also notes that the activity of the HF firms from 2:00 through 2:45pm *"is consistent with some HFT firms reducing or pausing trading during that time"*.[18]

Large liquidity providers experienced severe losses and some eventually had to stop trading. According to filings with the Securities and Exchange Commission, Goldman Sachs had 10 days of trading losses, including three days of more than $100 million in trading losses. These have been partly attributed to the events surrounding May 6th, 2010.[19] Morgan Stanley and Bank of America reported similar trading losses during the same period. This picture contrasts with that of firms taking liquidity from the market, like TD Ameritrade, which saw an unusually high level of activity on May 6th 2010. *"Like everyone else, that day was a big day for us"*, CEO Fred Tomczyk said at the Sandler O'Neill Conference. *"We had a lot of trades. We don't talk too publicly about the number, but that day was a record day for us."*[20]

6. VPIN vs. VIX

The movement in VPIN foreshadowed the price movement in the E-mini contract. One might conjecture that measures of price volatility, such as the VIX, might also have played a similar role on May 6th. But that is simply not the case: VIX and VPIN exhibited very different behavior on the day of the crash. Of particular importance, VIX lagged E-mini S&P 500 future's VPIN metric before, during and after the event.

Figure 6 demonstrates the behavior of the VIX, the VPIN, and the E-mini future price on the day of the crash. Following an initial dip at the beginning of the trading day, VIX did experience a run up, from an open of 25.92 at 9:30am to its highest value of the day, 40.69 at 15:28pm. Unlike what we observed with the E-mini VPIN metric, these levels of VIX are far from being their highest in recent history (the VIX reached 89.53 on October 24, 2008).

While the VPIN metric exhibited a smooth, gradual increase during the day of the flash crash, reaching levels consistent with a 90% CDF(VPIN) by 11:55am, VIX did not hit comparable levels until after the market had collapsed to its lowest level. So rather than anticipating the crash, VIX was impacted by the crash. Indeed, VIX itself became extremely volatile between 2:46pm and 4pm, retreating to 33.32 at around 3:16pm. Such behavior is not surprising: VIX is an investable product, and it was subject to the same price and liquidity imperfections as the rest of the investment universe.

---

[16] *"Speedy New Traders Make Waves Far From Wall St."*. The New York Times, May 17th, 2010.
[17] Op Cit, pg. 14.
[18] Ibid, pg. 48
[19] *"Goldman had 10 days of trading losses in Q2"*. Reuters, August 9th, 2010
[20] *"'Flash crash' a record day at TD Ameritrade: CEO"*. Reuters, June 3rd, 2010.
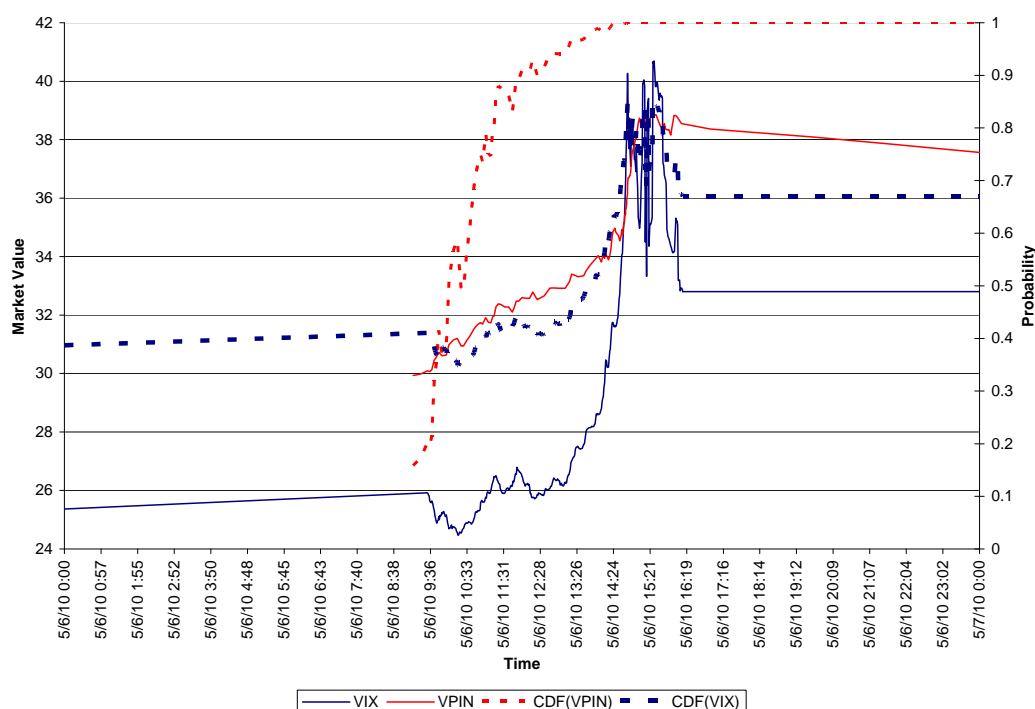
Figure 6 – VIX and E-mini's VPIN metric during the crash

7. VPIN as a Measure of the Risk of Liquidity-Induced Crashes

High frequency traders generate well over 60% of the volume in equity futures markets. But we have estimated that during the volume burst that triggered the crash, nearly 80% of the flow in the E-mini future was toxic. So although high frequency traders usually provide liquidity, the CFTC-SEC report suggests they turned to consuming liquidity during the crash, in effect producing toxic order flow. This behavior, in turn, exacerbated the developing liquidity crisis.

To understand why toxicity of order flow can induce such behavior from market makers, let us return to the role that information plays in affecting liquidity in the market. Easley and O'Hara (1992) sets out the mechanism by which informed traders extract wealth from liquidity providers. For example, if a liquidity provider trades against a buy order he loses the difference between the ask price and the expected value of the contract if the buy is from an informed trader. On the other hand, he gains the difference between the ask price and the expected value of the contract if the buy is from an uninformed trader. This loss and gain, weighted by the probabilities of the trade arising from an informed trader or an uninformed trader just balance due to the intense competition between liquidity providers.

$$\underbrace{\left(A - E\left[S_T \middle| U, Buy\right]\right)\Pr{ob}\left(U \middle| Buy\right)}_{Gain \text{ from an uninformed trader}} = -\underbrace{\left(A - E\left[S_T \middle| I, Buy\right]\right)\Pr{ob}\left(I \middle| Buy\right)}_{Loss \text{ to an informed trader}}$$

If flow toxicity unexpectedly rises (a greater than expected fraction of trades arises from informed traders), market makers face losses. Their inventory may grow beyond their risk limits, in which case they are forced to withdraw from the side of the market that is being adversely selected. Their withdrawal generates further weakness on that side of the market and their inventories keep accumulating additional losses. At some

10

point they capitulate, dumping their inventory and taking the loss. In other words, *extreme toxicity has the ability to transform liquidity providers into liquidity consumers*. This is likely to be particularly true in the context of equity and equity futures markets, which have been the most impacted by the new trends in market microstructure discussed in Section 2.

Over the short horizon high frequency liquidity providers deem relevant, the toxicity of orders matters because it can signal potentially adverse future movement of returns. **Table 1** shows the distribution of absolute returns between two consecutive volume buckets, conditional on the previous level of VPIN for the E-mini S&P500. That these are probabilities of absolute returns conditional on VPIN can be appreciated from the fact that each row adds up to 100%. VPIN ranges have been chosen to contain each 5% of the total 37163 observations from January 1st 2008 to October 30th 2010.

	0.25%	0.50%	0.75%	1.00%	1.25%	1.50%	1.75%	2.00%	>2.00%
0.285	83.53%	13.19%	2.37%	0.59%	0.27%	0.00%	0.00%	0.05%	0.00%
0.301	82.51%	14.32%	2.21%	0.59%	0.11%	0.11%	0.11%	0.00%	0.05%
0.312	83.42%	13.83%	2.10%	0.48%	0.05%	0.00%	0.11%	0.00%	0.00%
0.320	81.32%	14.64%	2.74%	0.81%	0.32%	0.05%	0.11%	0.00%	0.00%
0.327	81.33%	15.55%	2.10%	0.54%	0.22%	0.16%	0.05%	0.00%	0.05%
0.334	80.25%	16.15%	2.64%	0.65%	0.16%	0.16%	0.00%	0.00%	0.00%
0.341	79.87%	16.31%	2.69%	0.86%	0.16%	0.11%	0.00%	0.00%	0.00%
0.348	78.90%	16.42%	2.80%	0.97%	0.54%	0.32%	0.00%	0.00%	0.05%
0.354	81.27%	14.91%	2.80%	0.75%	0.11%	0.11%	0.00%	0.00%	0.05%
0.360	79.40%	15.76%	3.50%	0.65%	0.38%	0.16%	0.05%	0.05%	0.05%
0.367	80.03%	15.45%	3.18%	0.81%	0.38%	0.05%	0.11%	0.00%	0.00%
0.374	80.41%	15.45%	3.12%	0.81%	0.11%	0.11%	0.00%	0.00%	0.00%
0.382	78.63%	15.82%	3.50%	1.02%	0.43%	0.38%	0.11%	0.05%	0.05%
0.390	78.26%	16.42%	3.77%	0.91%	0.43%	0.11%	0.05%	0.00%	0.05%
0.399	74.77%	19.53%	3.77%	1.34%	0.32%	0.16%	0.11%	0.00%	0.00%
0.410	73.36%	20.13%	4.47%	1.02%	0.59%	0.16%	0.16%	0.05%	0.05%
0.423	74.43%	19.21%	4.09%	1.08%	0.59%	0.38%	0.05%	0.00%	0.16%
0.439	70.51%	20.83%	5.76%	1.72%	0.65%	0.27%	0.11%	0.05%	0.11%
0.466	65.39%	22.07%	7.37%	2.58%	1.56%	0.43%	0.27%	0.22%	0.11%
0.685	59.53%	22.82%	9.10%	4.84%	1.67%	1.02%	0.48%	0.32%	0.22%

Table 1 –Absolute Returns conditional on prior VPIN for E-mini S&P500

Table 1 shows that below average VPIN levels are followed by absolute returns below or equal to 0.25% in about 80% of the cases. As VPIN levels increase, there is a transfer of probability from lower returns to higher returns. For absolute returns greater than 0.25%, the highest concentration of probability occurs in the 5% top range of prior VPIN readings.

Table 1 provides information about the distribution of absolute returns following a certain toxicity level. It shows VPIN is useful as a *predictive* measure of absolute returns. An alternative use of VPIN is as a *warning* measure. **Table 2**, which shows the distribution of prior VPIN levels conditional on the following absolute returns, provides information about this aspect of the relationship between VPIN and absolute returns. As these are probabilities of VPIN conditional on absolute returns, each column in Table 2 sums to one.

11

	0.25%	0.50%	0.75%	1.00%	1.25%	1.50%	1.75%	2.00%	>2.00%
0.285	5.40%	3.89%	3.20%	2.57%	2.98%	0.00%	0.00%	6.67%	0.00%
0.301	5.33%	4.23%	2.98%	2.57%	1.19%	2.53%	5.71%	0.00%	5.26%
0.312	5.39%	4.08%	2.83%	2.10%	0.60%	0.00%	5.71%	0.00%	0.00%
0.320	5.26%	4.32%	3.71%	3.50%	3.57%	1.27%	5.71%	0.00%	0.00%
0.327	5.26%	4.59%	2.83%	2.34%	2.38%	3.80%	2.86%	0.00%	5.26%
0.334	5.19%	4.77%	3.56%	2.80%	1.79%	3.80%	0.00%	0.00%	0.00%
0.341	5.16%	4.81%	3.63%	3.74%	1.79%	2.53%	0.00%	0.00%	0.00%
0.348	5.10%	4.85%	3.78%	4.21%	5.95%	7.59%	0.00%	0.00%	5.26%
0.354	5.25%	4.40%	3.78%	3.27%	1.19%	2.53%	0.00%	0.00%	5.26%
0.360	5.13%	4.65%	4.72%	2.80%	4.17%	3.80%	2.86%	6.67%	5.26%
0.367	5.17%	4.56%	4.29%	3.50%	4.17%	1.27%	5.71%	0.00%	0.00%
0.374	5.20%	4.56%	4.22%	3.50%	1.19%	2.53%	0.00%	0.00%	0.00%
0.382	5.08%	4.67%	4.72%	4.44%	4.76%	8.86%	5.71%	6.67%	5.26%
0.390	5.06%	4.85%	5.09%	3.97%	4.76%	2.53%	2.86%	0.00%	5.26%
0.399	4.84%	5.77%	5.09%	5.84%	3.57%	3.80%	5.71%	0.00%	0.00%
0.410	4.74%	5.94%	6.03%	4.44%	6.55%	3.80%	8.57%	6.67%	5.26%
0.423	4.81%	5.67%	5.52%	4.67%	6.55%	8.86%	2.86%	0.00%	15.79%
0.439	4.56%	6.15%	7.78%	7.48%	7.14%	6.33%	5.71%	6.67%	10.53%
0.466	4.23%	6.51%	9.96%	11.21%	17.26%	10.13%	14.29%	26.67%	10.53%
0.685	3.85%	6.74%	12.28%	21.03%	18.45%	24.05%	25.71%	40.00%	21.05%

Table 2 –Prior VPIN conditional on absolute returns for E-mini S&P500

For example, the VPIN metric was greater than 0.41 prior to 80% of all absolute returns between 1.75% and 2%. So although Table 1 tells us that a VPIN greater than 0.41 does not necessarily imply a crash, Table 2 shows that, if the crash occurs, it is likely that the prior VPIN level was elevated.

These results lead to a dual interpretation for the VPIN metric:
- At relatively normal levels, it is a measurement of flow toxicity.
- At abnormally high levels, it can also be understood as indicating the likelihood that market makers turn into liquidity consumers, or that position takers may join and reinforce a brewing market imbalance. In markets dominated by high frequency liquidity providers, such as equity futures, this could lead to them destroying the market they were making.

This second interpretation makes the VPIN metric an interesting measure to use in monitoring the risk of a liquidity crash. Its accuracy is difficult to evaluate as liquidity-induced crashes are so rare that there is no sample on which we can perform a quantitative study. Nevertheless, both anecdotal evidence and the CFTC-SEC Report confirm that some liquidity providers turned into liquidity consumers during the liquidity crash[21].

We can also use our model to provide empirical insight into this phenomenon. During periods of unusually high VPIN metric values, we would expect *some* increase in stocks' volatility as a result of liquidity providers withdrawing from the marketplace. To check for this effect, we computed the correlation between the VPIN metric and the absolute value of the subsequent price changes. Note that the VPIN metric *is not* designed to forecast volatility; it is based on volume information rather than price

---

[21] For example, The New York Times reports: *"But on the afternoon of May 6, as the stock market began to plunge in the "flash crash," someone here walked up to one of those computers and typed the command HF STOP: sell everything, and shutdown. Across the country, several of Tradeworx's counterparts did the same. In a blink, some of the most powerful players in the stock market today — high-frequency traders — went dark. The result sent chills through the financial world."* Cf. Julie Creswell's article from May 16th 2010, *"Speedy New Traders Make Waves Far From Wall St.".*

information. We found a positive and statistically significant correlation (0.1596) between the VPIN metric and future volatility of E-mini S&P500, suggesting that an increase in the VPIN metric does foreshadow an increase in volatility in that instrument.

8. Proposed Solution: The 'VPIN contract'

The 'flash crash' might have been avoided, or at least tempered, had liquidity providers remained in the marketplace. Not only did some withdraw, but arguably they became liquidity consumers by dumping their inventories, thus exacerbating the crash.

One approach to lessen the likelihood and magnitude of future flash crashes may be to offer market makers the tools they need to *measure* and *manage* their risk of being adversely selected:

- **Measurement**: Our measure of flow toxicity, the VPIN metric, could be used by market makers to anticipate a rise in volatility and estimate the risk of a liquidity-induced crash.
- **Management**: Creating an *exchange future with the VPIN metric as underlying* would make available a visible reading of flow toxicity and a venue in which liquidity providers could hedge the risk of being adversely selected.

A 'VPIN contract' could work as a hedging and speculation mechanism. It may make it less likely that liquidity providers would turn into liquidity consumers, because as they perceive an inventory growth they can dynamically and continuously hedge their risks, rather than trying to hold their position and possibly being forced to capitulate in a cascade.

Other potential uses of the VPIN metric are:

- **As a benchmark for execution brokers**, who could try to fill their customers' orders while beating the average flow toxicity of the session. Similarly, clients could use the VPIN metric as a tool to indicate under which conditions brokers should stop filling their orders, and to measure how effectively their brokers avoided adverse selection.
- **As a warning sign for market regulators**, who may decide to slow down or stop market activity as flow toxicity reaches levels comparable to those witnessed on May 6th 2010, thus preventing or mitigating the collapse.
- **An instrument for volatility arbitrage**, since the VPIN metric is useful in improving forecasts on volatility.

9. Conclusions

The CFTC-SEC Staff Report identifies conjunctural factors as its explanation of the 'Flash Crash'. While acknowledging such factors may have played a role, our analysis suggests that the 'Flash Crash' is better understood as a liquidity event arising from structural features of the new high frequency world of trading. In this high frequency world, liquidity provision is dominated by computerized market makers programmed to place buy and sell orders while avoiding taking significant inventory positions. When order flow toxicity increases, such market makers face significant losses and curtail their risks by reducing, or even liquidating, their

13

positions. The consequent market illiquidity can then have disastrous repercussions for market participants.

While some have called for banning high frequency trading, we believe a better solution lies in recognizing and managing the risks of trading in this new market structure. The creation of an exchange-traded 'VPIN contract' would serve the dual goal of offering market makers an objective *measurement* of flow toxicity, plus a risk *management* tool to hedge the risk of being adversely selected, with implications for execution brokers and market regulators. During periods of market stress, dynamic hedging of their VPIN metric exposure might allow high frequency market makers to remain in the marketplace providing liquidity, thus mitigating or possibly avoiding the next 'flash crash'.

BIBLIOGRAPHY

- Ané, T. and H. Geman (2000): *"Order flow, transaction clock and normality of asset returns"*, The Journal of Finance, 55: 2259–2284.
- CFTC-SEC, *Preliminary findings Regarding the Market Events of May 6, 2010*, May 18, 2010.
- CFTC-SEC, *Findings Regarding the Market Events of May 6, 2010*, September 30, 2010.
- CME Group, *Statement on the Joint CFTC/SEC Report Regarding the Events of May 6*, October 1, 2010.
- Easley, D. and M. O'Hara (1987): *"Price, Trade Size, and Information in Securities Markets"*, Journal of Financial Economics, 19.
- Easley, D. and M. O'Hara (1992): *"Time and the process of security price adjustment"*, Journal of Finance, 47, 576-605.
- Easley, D., Kiefer, N., O'Hara, M. and J. Paperman (1996): *"Liquidity, Information, and Infrequently Traded Stocks"*, Journal of Finance, September.
- Easley, D., R. F. Engle, M. O'Hara and L. Wu (2008): *"Time-Varying Arrival Rates of Informed and Uninformed Traders"*, Journal of Financial Econometrics.
- Easley, D., M. López de Prado and M. O'Hara (2010): *"Measuring flow toxicity in a high frequency world"*, SSRN, Working Paper Series, http://papers.ssrn.com/abstract=1695596.
- Glosten, L. R. and P. Milgrom (1985): *"Bid, ask and transaction prices in a specialist market with heterogeneously informed traders"*, Journal of Financial Economics, 14, 71-100.
- Iati, Robert (2009): *"High Frequency Trading Technology"*, TABB Group.
- Kirilenko, A., A. Kyle, M. Samadi and T. Tuzun (2010): " *The Flash Crash: The Impact of High Frequency Trading on an Electronic Market*", Working paper.
- O'Hara, M. and G. Oldfield (1986): *"The Microeconomics of Market Making"*, Journal of Financial and Quantitative Analysis, December.

DISCLAIMER