Order toxicity and liquidity crisis: An academic point of view on Flash Crash

Discussant

Fulvio Corsi

University of Lugano and SFI

11 May 2011

Introduction

We review two papers on the causes of the Flash Crash by Easley, De Prado and O'Hara:

- "The Microstructure of Flash Crash" (Working Paper November 2010)
- "Flow Toxicity and Volatility in High Frequency World" (Working Paper February 2011)

Summary

1. Flash Crash caused by severe mismatch in liquidity: liquidity providers withdraw from the market or even turned into liquidity takers.
2. Liquidity dries up due to "toxic" (unbalanced) order flows.
3. Authors propose a measure of order toxicity, the VPIN metric.
4. They show that this VPIN measure anticipated the Flash Crash.

Summary

1. Flash Crash caused by severe mismatch in liquidity: liquidity providers withdraw from the market or even turned into liquidity takers.

2. Liquidity dries up due to “toxic” (unbalanced) order flows.

3. Authors propose a measure of order toxicity, the VPIN metric.

4. They show that this VPIN measure anticipated the Flash Crash.

Summary

1. Flash Crash caused by severe mismatch in liquidity: liquidity providers withdraw from the market or even turned into liquidity takers.
2. Liquidity dries up due to “toxic” (unbalanced) order flows.
3. Authors propose a measure of order toxicity, the VPIN metric.
4. They show that this VPIN measure anticipated the Flash Crash.

Summary

1. Flash Crash caused by severe mismatch in liquidity: liquidity providers withdraw from the market or even turned into liquidity takers.
2. Liquidity dries up due to “toxic” (unbalanced) order flows.
3. Authors propose a measure of order toxicity, the VPIN metric.
4. They show that this VPIN measure anticipated the Flash Crash.

Recent trends in market structure

- Since 2009, HF trading firms ($\approx 2\%$ of total $20,000$ US firms) accounted for over 70% of U.S. equity trading volume.
- Many of these HF firms are in the business of “liquidity provision”, i.e. acting as market maker (MM) to “position takers”.
- HF MM generally do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades.
- Their ability to do so depends on limiting their position risk by:
  - hold very small or zero inventory positions
  - have high inventory turnover (5 or more times a day)
  - control “adverse selection”
- $\Rightarrow$ Allow them to operate with very low capital, essentially using their speed of trading to control inventory risk.

Recent trends in market structure

- Since 2009, HF trading firms ($\approx 2\%$ of total $20,000$ US firms) accounted for over 70% of U.S. equity trading volume.
- Many of these HF firms are in the business of "liquidity provision", i.e. acting as market maker (MM) to "position takers".
- HF MM generally do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades.
- Their ability to do so depends on limiting their position risk by:
  - hold very small or zero inventory positions
  - have high inventory turnover (5 or more times a day)
  - control "adverse selection"
- $\Rightarrow$ Allow them to operate with very low capital, essentially using their speed of trading to control inventory risk.

Recent trends in market structure

- Since 2009, HF trading firms ($\approx 2\%$ of total $20,000$ US firms) accounted for over 70% of U.S. equity trading volume.

- Many of these HF firms are in the business of "liquidity provision", i.e. acting as market maker (MM) to "position takers".

- HF MM generally do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades.

- Their ability to do so depends on limiting their position risk by:
  - hold very small or zero inventory positions
  - have high inventory turnover (5 or more times a day)
  - control "adverse selection"

- $\Rightarrow$ Allow them to operate with very low capital, essentially using their speed of trading to control inventory risk.

Recent trends in market structure

- Since 2009, HF trading firms ($\approx 2\%$ of total $20,000$ US firms) accounted for over 70% of U.S. equity trading volume.
- Many of these HF firms are in the business of “liquidity provision”, i.e. acting as market maker (MM) to “position takers”.
- HF MM generally do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades.
- Their ability to do so depends on limiting their position risk by:
  - hold very small or zero inventory positions
  - have high inventory turnover (5 or more times a day)
  - control “adverse selection”
- $\Rightarrow$ Allow them to operate with very low capital, essentially using their speed of trading to control inventory risk.

Recent trends in market structure

- Since 2009, HF trading firms ($\approx 2\%$ of total $20,000$ US firms) accounted for over 70% of U.S. equity trading volume.
- Many of these HF firms are in the business of “liquidity provision”, i.e. acting as market maker (MM) to “position takers”.
- HF MM generally do not make directional bets, but rather strive to earn razor thin margins on large numbers of trades.
- Their ability to do so depends on limiting their position risk by:
  - hold very small or zero inventory positions
  - have high inventory turnover (5 or more times a day)
  - control “adverse selection”
- $\Rightarrow$ Allow them to operate with very low capital, essentially using their speed of trading to control inventory risk.

Market Microstructure Models

- Microstructure models view trading as a game between liquidity providers (or MM) and liquidity takers (or traders or position takers).
- MMs set the spread to be compensated for:
  - operational costs
  - inventory costs
  - adverse selection costs
- Adverse selection arises because some traders may have better information on the future price than MM.
- The Authors define toxicity "the expected loss from trading with better informed counterparties".

Market Microstructure Models

- Microstructure models view trading as a game between liquidity providers (or MM) and liquidity takers (or traders or position takers).
- MMs set the spread to be compensated for:
  - operational costs
  - inventory costs
  - adverse selection costs
- Adverse selection arises because some traders may have better information on the future price than MM.
- The Authors define toxicity “the expected loss from trading with better informed counterparties”.

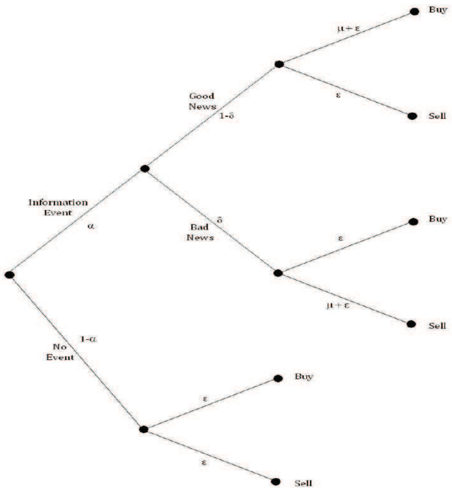
Market Microstructure Models

- Microstructure models view trading as a game between liquidity providers (or MM) and liquidity takers (or traders or position takers).
- MMs set the spread to be compensated for:
  - operational costs
  - inventory costs
  - adverse selection costs
- Adverse selection arises because some traders may have better information on the future price than MM.
- The Authors define toxicity “the expected loss from trading with better informed counterparties”.

Market Microstructure Models

- Microstructure models view trading as a game between liquidity providers (or MM) and liquidity takers (or traders or position takers).
- MMs set the spread to be compensated for:
  - operational costs
  - inventory costs
  - adverse selection costs
- Adverse selection arises because some traders may have better information on the future price than MM.
- The Authors define toxicity “the expected loss from trading with better informed counterparties”.

Sketch of a simple model of adverse selection



Market Microstructure Models

- If $\delta = 1/2$, it can be shown that the bid-ask spread simplified to

$$s = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}\left[\,\overline{S_i} - \underline{S_i}\,\right]$$

where $\overline{S_i}$ and $\underline{S_i}$ are price predictions of informed trades in case of good and bad news.

- The probability that a trade in a period is information-based (PIN) is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}$$

where $\alpha\mu + 2\epsilon$ is the arrival rate for all orders and $\alpha\mu$ is the arrival rate for information-based orders.

- PIN is thus a measure of the fraction of orders that arise from informed traders relative to the total order flow.

- MMs need to correctly estimate their PIN in order to identify the optimal spread s.

Market Microstructure Models

- If $\delta = 1/2$, it can be shown that the bid-ask spread simplified to

$$s = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \left[\overline{S_i} - \underline{S_i} \right]$$

  where $\overline{S_i}$ and $\underline{S_i}$ are price predictions of informed trades in case of good and bad news.

- The probability that a trade in a period is information-based (PIN) is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}$$

  where $\alpha\mu + 2\epsilon$ is the arrival rate for all orders and $\alpha\mu$ is the arrival rate for information-based orders.

- PIN is thus a measure of the fraction of orders that arise from informed traders relative to the total order flow.

- MMs need to correctly estimate their PIN in order to identify the optimal spread s.

Market Microstructure Models

- If $\delta = 1/2$, it can be shown that the bid-ask spread simplified to

$$s = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \left[\overline{S_i} - \underline{S_i} \right]$$

where $\overline{S_i}$ and $\underline{S_i}$ are price predictions of informed trades in case of good and bad news.

- The probability that a trade in a period is information-based (PIN) is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}$$

where $\alpha\mu + 2\epsilon$ is the arrival rate for all orders and $\alpha\mu$ is the arrival rate for information-based orders.

- PIN is thus a measure of the fraction of orders that arise from informed traders relative to the total order flow.

- MMs need to correctly estimate their PIN in order to identify the optimal spread s.

Market Microstructure Models

- If $\delta = 1/2$, it can be shown that the bid-ask spread simplified to

$$s = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \left[\, \overline{S_i} - \underline{S_i} \, \right]$$

  where $\overline{S_i}$ and $\underline{S_i}$ are price predictions of informed trades in case of good and bad news.

- The probability that a trade in a period is information-based (PIN) is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}$$

  where $\alpha\mu + 2\epsilon$ is the arrival rate for all orders and $\alpha\mu$ is the arrival rate for information-based orders.

- PIN is thus a measure of the fraction of orders that arise from informed traders relative to the total order flow.

- MMs need to correctly estimate their PIN in order to identify the optimal spread s.

PIN estimation: VPIN theory

- Standard approach to estimate the PIN is to employ maximum likelihood estimation to get the unobservable parameters α, μ, ϵ and then derive PIN from those estimates.
- The Authors propose a more direct volume-based approach observing that:

  the expected trade imbalance is:

  $$\mathbb{E}\left[\left|V_\tau^S - V_\tau^B\right|\right] \approx \alpha\mu$$

  where V_τ^S is the sell volume and V_τ^B is the buy volume.

  and the expected arrival rate of total trades $V = V_\tau^S + V_\tau^B$ is:

  $$\mathbb{E}\left[V\right] = \alpha\mu + 2\epsilon$$

- Hence, the *Volume-Synchronized Probability of Informed Trading* VPIN is

  $$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \approx \frac{\alpha\mu}{V} \approx \frac{\sum_{\tau=1}^{n} \left|V_\tau^S - V_\tau^B\right|}{nV} = VPIN$$

PIN estimation: VPIN theory

- Standard approach to estimate the PIN is to employ maximum likelihood estimation to get the unobservable parameters α, μ, ϵ and then derive PIN from those estimates.

- The Authors propose a more direct volume-based approach observing that:

  the expected trade imbalance is:

$$\mathbb{E}\left[\left|V_\tau^S - V_\tau^B\right|\right] \approx \alpha\mu$$

  where V_τ^S is the sell volume and V_τ^B is the buy volume.

  and the expected arrival rate of total trades $V = V_\tau^S + V_\tau^B$ is:

$$\mathbb{E}\left[V\right] = \alpha\mu + 2\epsilon$$

- Hence, the *Volume-Synchronized Probability of Informed Trading* VPIN is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \approx \frac{\alpha\mu}{V} \approx \frac{\sum_{\tau=1}^{n} \left|V_\tau^S - V_\tau^B\right|}{nV} = VPIN$$

PIN estimation: VPIN theory

- Standard approach to estimate the PIN is to employ maximum likelihood estimation to get the unobservable parameters α, μ, ϵ and then derive PIN from those estimates.

- The Authors propose a more direct volume-based approach observing that:

  the expected trade imbalance is:

$$\mathbb{E}\left[\left|V_\tau^S - V_\tau^B\right|\right] \approx \alpha\mu$$

  where V_τ^S is the sell volume and V_τ^B is the buy volume.

  and the expected arrival rate of total trades $V = V_\tau^S + V_\tau^B$ is:

$$\mathbb{E}\left[V\right] = \alpha\mu + 2\epsilon$$

- Hence, the *Volume-Synchronized Probability of Informed Trading* VPIN is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon} \approx \frac{\alpha\mu}{V} \approx \frac{\sum_{\tau=1}^{n}\left|V_\tau^S - V_\tau^B\right|}{nV} = VPIN$$

VPIN in practice

$$VPIN = \frac{\sum_{\tau=1}^{n} \left| V_\tau^S - V_\tau^B \right|}{nV}$$

- Sample the prices in "Volume-time", i.e. in intervals having equal amount of volume V. They choose $V = 1/50$ of the average daily volume and $n = 50 \Rightarrow$ "daily" VPIN (on average).

- Volume Classification (in buy V_τ^B and sell V_τ^S volume).

  Trade classification is always problematic: more so in the HF world of electronic order book where applying standard tick-based algos over individual transactions would be "futile".

  $\Rightarrow$ propose to aggregate trades over short time intervals Δ (e.g. 1-minute) and sign the aggregated volume in that time interval as the corresponding transaction:

  An aggregated transaction is buy if either

  i $P_i > P_{i-\Delta}$ or
  ii $P_i = P_{i-\Delta}$ and the transaction $i - \Delta$ was also a buy.

  Otherwise, the transaction is a sell.

VPIN in practice

$$VPIN = \frac{\sum_{\tau=1}^{n} \left| V_\tau^S - V_\tau^B \right|}{nV}$$

- Sample the prices in "Volume-time", i.e. in intervals having equal amount of volume V. They choose $V = 1/50$ of the average daily volume and $n = 50 \Rightarrow$ "daily" VPIN (on average).
- Volume Classification (in buy V_τ^B and sell V_τ^S volume).

  Trade classification is always problematic: more so in the HF world of electronic order book where applying standard tick-based algos over individual transactions would be "futile".

  $\Rightarrow$ propose to aggregate trades over short time intervals Δ (e.g. 1-minute) and sign the aggregated volume in that time interval as the corresponding transaction:

  An aggregated transaction is buy if either

  i $P_i > P_{i-\Delta}$ or
  ii $P_i = P_{i-\Delta}$ and the transaction $i - \Delta$ was also a buy.

  Otherwise, the transaction is a sell.

VPIN of E-mini S&P500 over 3 years

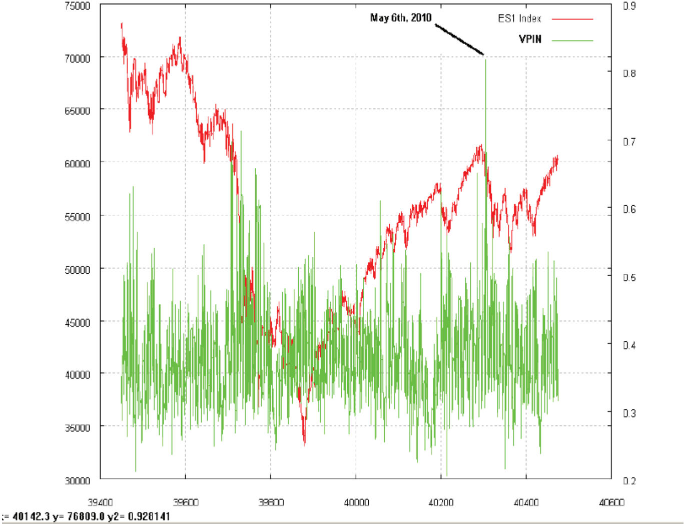


Figure 2 - VPIN metric between January 1st 2008 and October 30th 2010

VPIN: Historical PDF and CDF

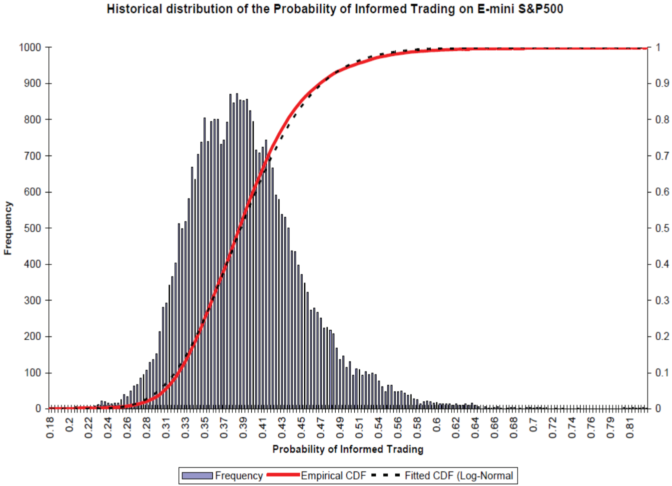


Figure 3 - The empirical CDF of the VPIN metric as fitted through a log-normal distribution

VPIN 1 week before the Flash Crash

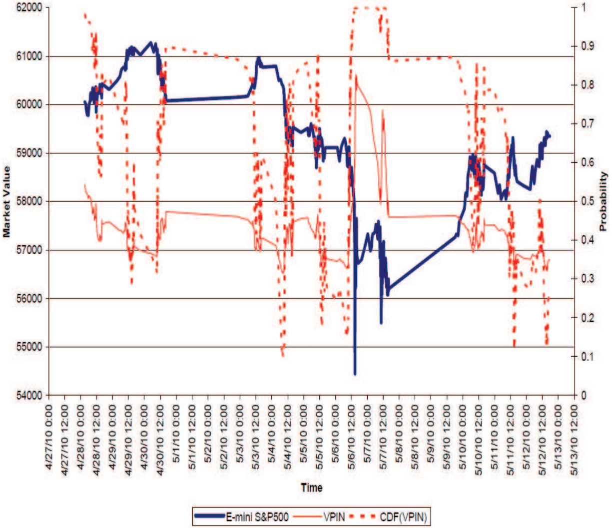


Figure 4 – E-mini S&P 500's VPIN metric one week before and after the flash crash

VPIN on the Flash Crash day

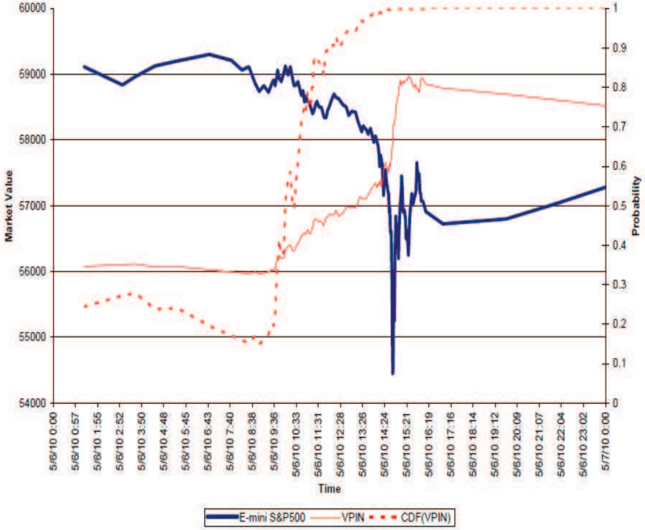


Figure 5 – E-mini S&P 500's VPIN metric on May 6th

VPIN vs VIX

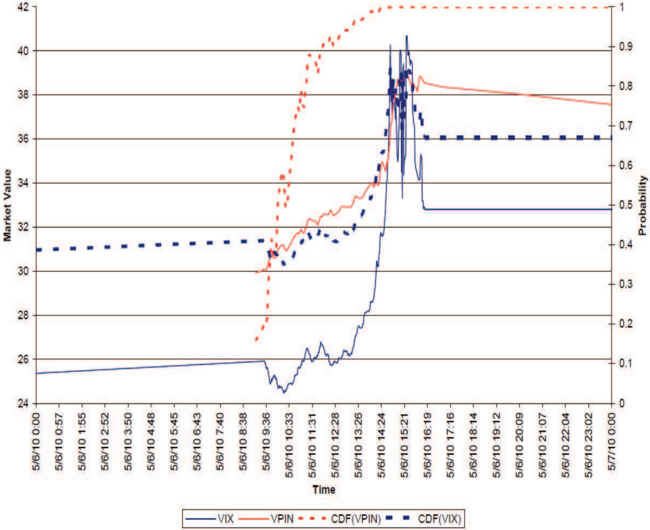


Figure 6 – VIX and E-mini's VPIN metric during the crash

Point of caution: Impact trade aggregation interval

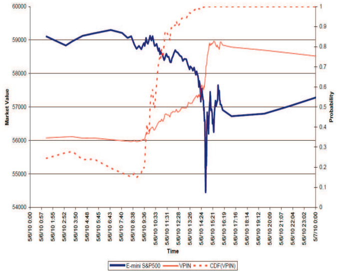


Figure 8(a) – VPIN estimated with One-Minute Time Bars

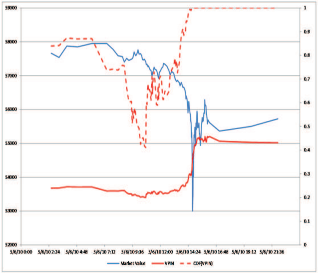


Figure 8(b) – VPIN Estimated with Ten-Second Time Bars

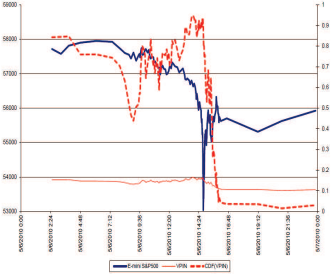


Figure 8(c) – VPIN Estimated with Trade-by-Trade Data

VPIN of EUR/USD and T-Note

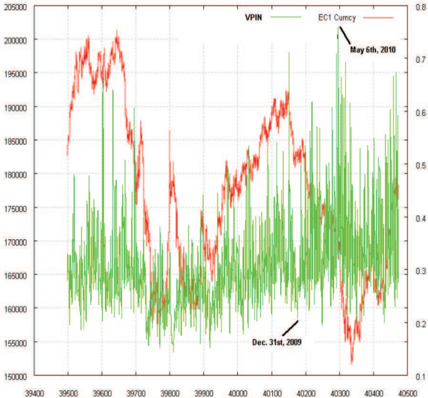


Figure 10 - VPIN for EUR/USD futures

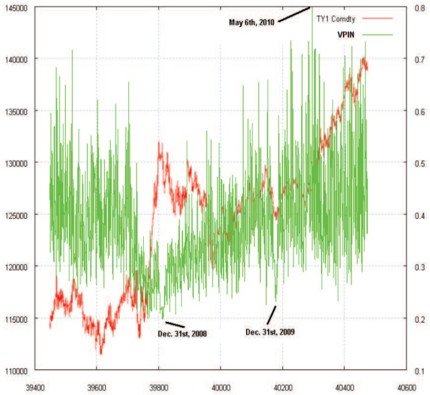


Figure 11 - VPIN for T-Note futures

Conclusions

- Flash Crash causes:
  - When flow toxicity unexpectedly rose (unusually unbalanced order flow as measured by VPIN) HF MMs face large losses.
  - Inventory may grow beyond their risk limits, forcing them to withdraw from the market.
  - If they keep accumulating losses, at some point they may capitulate, dumping their inventory to take the loss.

  Hence, extreme toxicity can transform liquidity providers into liquidity consumers.

- By measuring imbalance in order flow (toxicity) the proposed VPIN metric should predict liquidity crisis (as claimed for the Flash Crash).

- Authors proposed solution to liquidity crisis: Creating an exchange future with the VPIN metric as underlying.