Spoofing Detection in Digital Asset Centralized Exchanges

Faverjon & Fabre

Sun Zu Lab

July 20, 2023

Outline

Introduction

The Why ?, The What ?

Why Detect Manipulation

- Core to Sun Zu Lab's mission of transparency.
- Enhancement of clients' alpha in three ways.

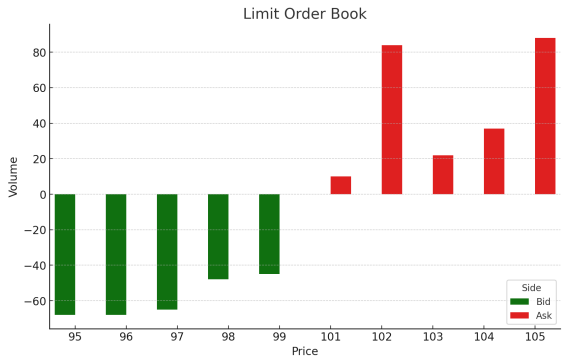
What we did

1. Implemented a function to warn investors of manipulating risk.
2. Pinned Orders suspected of being manipulative

Limit Order Book (LOB) & Conventions

Type of orders

(1) Limit Orders :
- Bid Orders : The market-maker proposes to buy a given amount at a given price. (We will always use the maker convention)
-Ask Order : The maker proposes to sell.
(2) Market Orders :
- Sell Trade : The order is an aggressive bid order, so the trade happens at Ask. The market-maker sells.
- Buy Trade : The order is an aggressive ask order, so the trade happens at Bid.



Types of spoofing and our focus

1. Spoofer Maker : Places an order at best bid (resp. ask) and wishes to get executed fast. He places a manipulating order on the ask (resp. bid) side. He modifies the imbalance, and increase the execution probability of his initial order.
2. Spoofer Taker : Wants to iniate a market order at bid (resp.ask) at the lowest (resp. highest) possible price. He places a manipulating order deep in the book, that results in the appearance of new limit order on the same size that he can execute. He then cancels his order.

- We focused on the the spoofer maker : manipulation in this case is much more profitable as the seller (resp. buyer) is able to avoid both crossing the spread and being applicated taker fees.
- In this presentation, we will focus on the maker spoofer willing to sell (i.e they manipulate at bid)

Our Assumption

Spoofer's behaviour

- The spoofer keeps his manipulating order until the next down best price movement : the market didn't react as expected, and fears the execution of his manipulating order.
- The spoofer puts a small order size at ask, so his fill ratio for this order is in $\{0,1\}$

Detection Algorithm

Notations

- $\phi_a^{\delta,Q}$: Fill ratio on ask side of an order placed δ bps from the best,and of size Q
- W^{maker} : Wealth resulting of the sell of one ATS of a token, by spoofing
- $W^{\text{post-wait}}$: Wealth resulting of the sell of one ATS of a token, with post-and-wait strategy.
- $\forall t \geq 0, p_t^b, p_t^a$: Best Bid Price, Best Ask Price
- $\forall t \geq 0, \Delta p_t^b := p_t^b - p_0^b$
- $h := \inf(t > 0, \Delta_t^b < 0)$
- f_b^+, f_a^+ : maker fees on bid and ask side.
- f_b^-, f_a^- : taker fees on bid and ask side
- $\mathbb{E}_x[.], \mathbb{P}_x$: Expectancy and Probability parametrized by imbalance depth and value.
- $\mathbb{E}_{x^-}[.], \mathbb{P}_{x^-}$: At the previous tick

The Equations

1. We compare the expected wealth resulting from the spoofing strategy with the one resulting from a post-and-wait.
2. Spoofing Wealth :

$$\begin{aligned}\mathbb{E}_x[W^{\text{maker}}] =& \mathbb{E}_x[\phi_b^{\delta,Q}]\, Q\,[-f_b^+\,(p_0^b-\delta)+f_a^-\,p_0^b]\\ &+\mathbb{E}_x[\phi_b^{\delta,Q}\,\Delta p_h^{\text{b}}]\, Q\, f_a^-\\ &+\mathbb{P}_x(\phi_a^{0,1}=0)\, f_a^-\,(p_0^b+\mathbb{E}_x[\Delta p_h^{\text{b}}\mid \phi_a^{0,1}=0])\\ &+\mathbb{P}_x(\phi_a^{0,1}=1)\, f_a^+\, p_0^a\end{aligned}$$

3. Post-and-Wait Wealth :

$$\begin{aligned}\mathbb{E}_{x^-}[W^{\text{post-wait}}] =& \mathbb{P}_{x^-}(\phi_h^{0,1}=0)\, f_a^-\,(p_b^0+\mathbb{E}_{x^-}[\Delta p_h^{\text{b}}\mid \phi_a^{0,1}=0])\\ &+\mathbb{P}_{x^-}(\phi_a^{0,1}=1)\, f_a^+\, p_0^a\end{aligned}$$

4. Spoofer's excess wealth function :
$W = \mathbb{E}_x[W^{\text{maker}}] - \mathbb{E}_{x^-}[W^{\text{post-wait}}]$

Methodology

Calculating ATS

For each pair and venue, we compute over a day of data the average trade size. It will be used as parameter in the calculation of other functions.

Calculating relevant depths

- For each venue and symbol, the book depth is different.
- We compute over a day the distances so that the cumulative bid volume over this distance amounts to a certain dollar value, between \$10,000 and \$1,000,000.

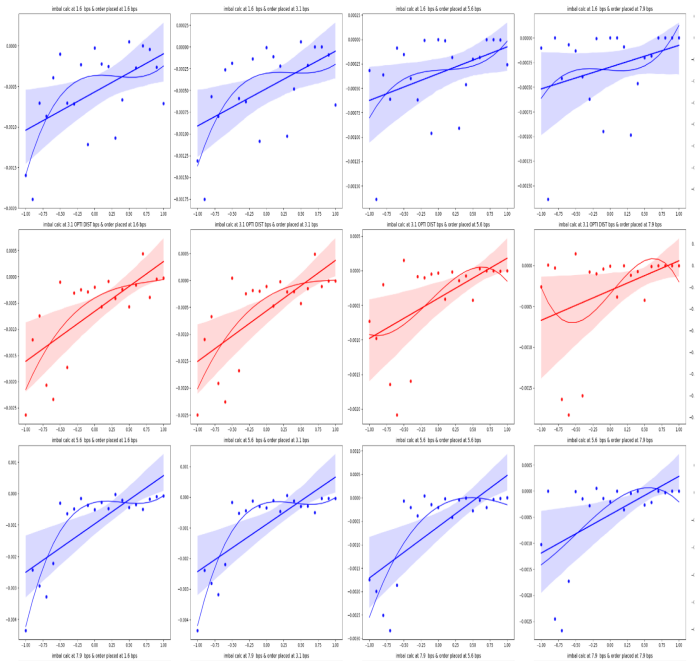
Methodology (2)

Calculating the helper functions

- We compute $\mathbb{E}_x[\phi_b^{\delta,Q}]$, etc for values of δ, Q (multiple of ATS), and values of x : The expectancy is a function of the order book imbalance and depth at which it is calculated.
- We compute and plot the functions with depth and imbalance values as arguments, and select the best depths with the following heuristic (given for fill-ratio as an example)
  1. For each depth, calculate the spearman correlation of imbalances and fill ratios. Map this correlation to the correct subset : Zone 1 : [0, 0.33], ..., Zone 3 : [0.66, 1]. If only one depth corresponds to the best zone, take it. If tie :
  2. For each depth that ties, calculate the pearson correlation, and do the same thing. If tie :
  3. Select the biggest depth.
  4. Perform a linear regression to parametrize the model.

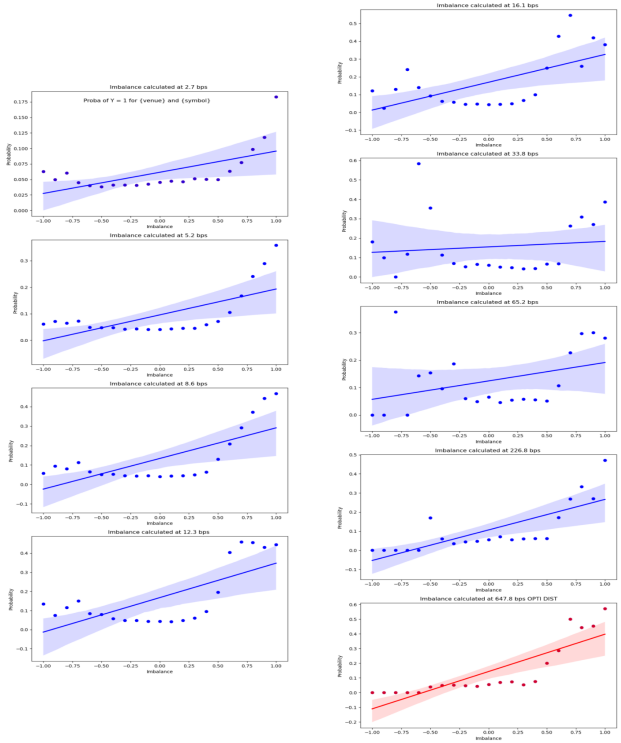
Methodology (3)

Impact of depths on precalculated data



$\mathbb{E}[\phi_b^{\delta,Q}]$

ETH-USD, gemini



$\mathbb{P}(\phi_a^{0,1} = 1)$

XRP-USD, kraken

Methodology (4)

Running the spoofing suspicion algorithm

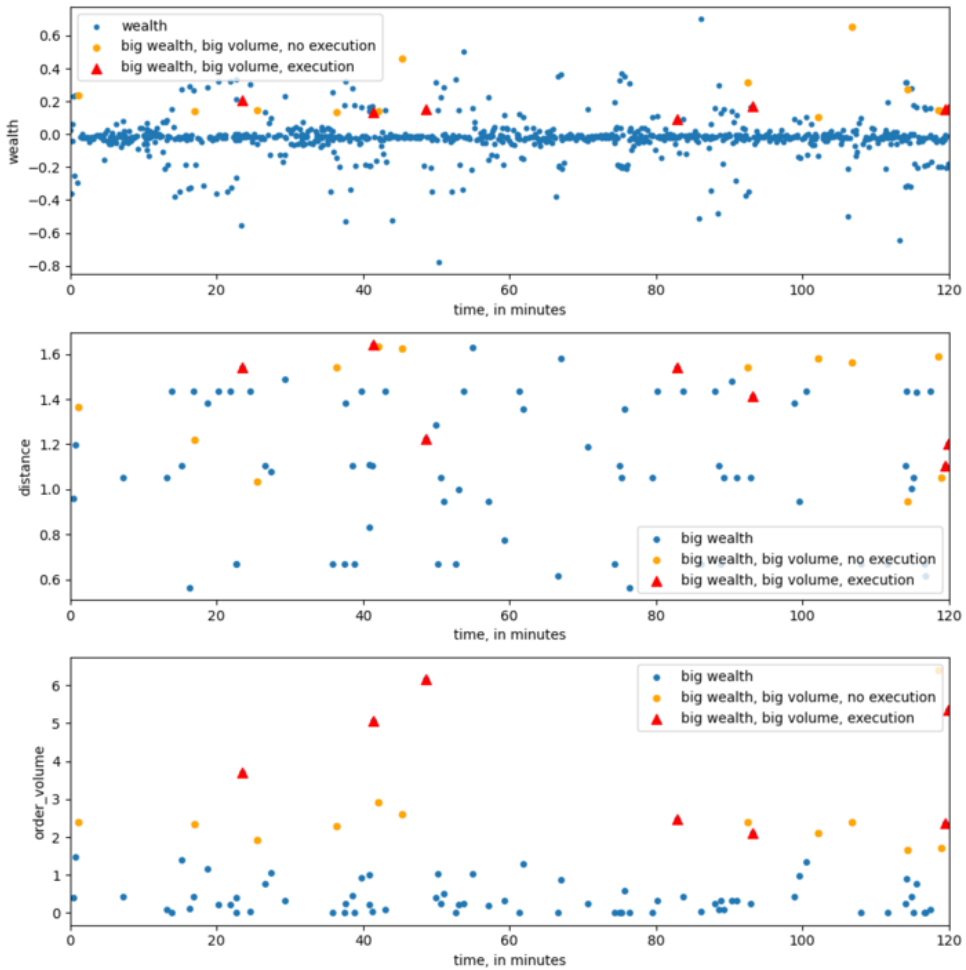


Figure: 2 hours of bitstamp, BTC-USD

Post Analysis

Gain of execution probability, when Wealth is in top 10%.

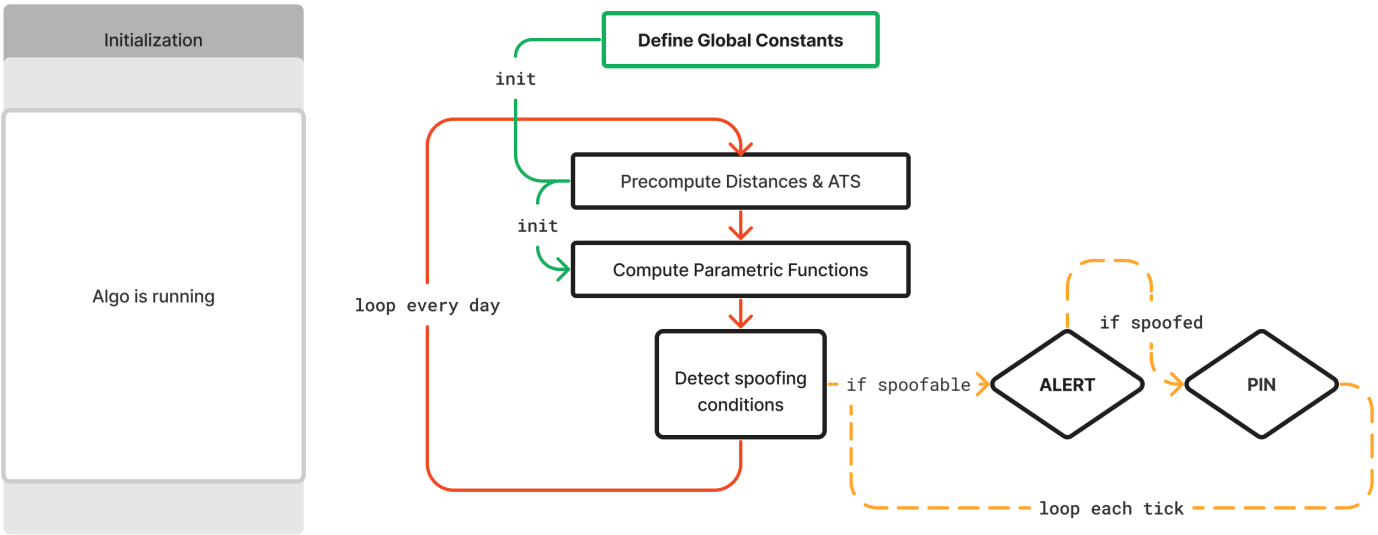
Symbol & Venue	Exec Proba Gain
Kraken BTC-USD	(2%)
Bitstamp BTC-USDT	82%
Bitstamp SHIB-USD	0.17 %
Bitstamp XRP-USD	(22%)
Gemini BTC-USD	6 %
Gemini SHIB-USD	15 %

What it tells us

The relevance of the wealth function is established, but varies in the time, from an exchange to another, from a symbol to another.

Live Detection

Algorithm Lifecycle



Live Spoofing Detection Algo Design Pattern

What's next ?

- Detect Taker spoofer.
- Launch production.
- Think about a visual interface.

Thank you for your attention