

Accumulation Opportunity

January 25, 2023

1 Introduction

Here we assess the feasibility of accumulating large positions while attempting to maintain low trading costs in an electronic market.

2 Data

Load the high-frequency data for all 3 pairs from the class website. The trade and book data has the following structure

2.1 Format

The data has the following structure¹

2.1.1 Trades

received_utc_nanoseconds	timestamp_utc_nanoseconds	PriceMillionths	SizeBillionths	Side
1674521267814309000	1674521267874527000	22970120000	87069600	-1
1674521267814046000	1674521267874527000	22970150000	25797600	-1
1674611962312088000	1674611962347434000	22499070000	4801640	-1
1674611962339264000	1674611962375191000	22498910000	1120200	-1

The *Side* is actually a sum of trade sides at the same price and time.

2.1.2 Book

Ask1PriceMillionths	22972550000	22972550000	22502670000	22502670000
Bid1PriceMillionths	22970150000	22970150000	22498910000	22498910000
Ask1SizeBillionths	210000000	410000000	101856140	101856140
Bid1SizeBillionths	25797600	25797600	280050	280050
Ask2PriceMillionths	22972560000	22972560000	22502680000	22502680000
Bid2PriceMillionths	22970120000	22970120000	22498690000	22498690000
Ask2SizeBillionths	210000000	210000000	50000000	50000000
Bid2SizeBillionths	87069610	87069610	12560150	12560150
received_utc_nanoseconds	1674521267750919800	1674521267751154000	1674611962359972000	1674611962365237000
timestamp_utc_nanoseconds	1674521267806932000	1674521267807073000	1674611962398574000	1674611962400579000
Mid	22971350000	22971350000	22500790000	22500790000

(transposed)

Note: only the sign of the *Side* is important to us.

3 Exercise

Write a simple VWAP participation algorithm that takes a target quantity Q , start time τ_s (also known as *arrival time*) and target participation rate p as parameters, and simulates the accumulation possible for a VWAP algorithm starting at time τ_s , both for buying (positive Q) and for selling (negative Q) in one of the cryptocurrency market data sets provided.

¹Note that inaccuracies in clock settings, i.e. “clock skew”, can cause timestamps to appear later than the time at which they are recorded as having been received.

Your simulation function should use actual signed trading volumes to judge simulated accumulation. It should make the conservative assumption that, in a flurry of trades trading through several levels², only the worst of those was “qualifying”, i.e. available for participation. For example, if you see 7 sells in a row within about 0.1 seconds³ at successively lower prices, you should assume your own sale could have participated only in the seventh (and worst) of those trades.

Assume transaction fees are 50 basis points (0.5% of traded notional) for transactions between cryptotokens and traditional currencies, and are 10 basis points for trades between cryptotokens.

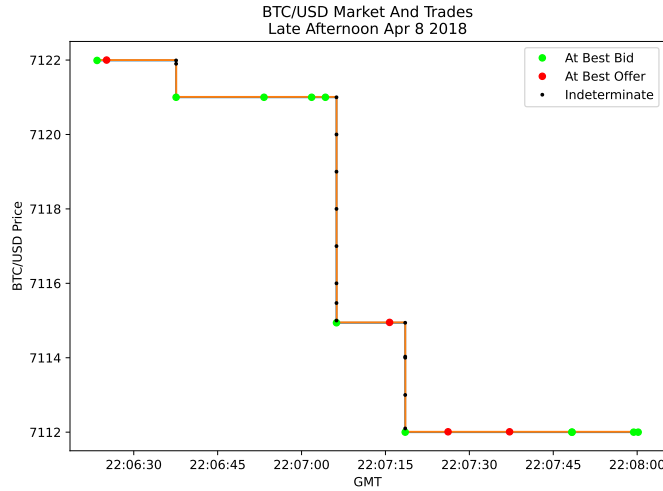


Figure 1: Cases of large sell orders crossing the spread, taking out several levels of resting buy orders.

4 Analysis

Assess and contrast the accumulation opportunities available in cryptotoken markets in 2021, 2022 and 2023. Concentrate on transactions that finish in 1-15 minutes or so, in order to give yourself a nice set of independent instances.

You can consider metrics such as $\frac{\text{Notional}}{\text{Time}}$, trading costs as a proportion of notional, statistical metrics comparing average price of accumulated positions to arrival prices, relationship between Q and the likelihood of completing the entire quantity, and/or deviation from expected accumulation rate.

²As, for example, when a large trade “takes out” several levels of the book.

³Use *timestamp_utc_nanoseconds* for your clocks.