High Frequency Trading

Wei Hong Ho
Finsam Samson
Joachim Sasson

Stanford University

April 27th 2021

Introduction

Long Term Goals: Design and implement a trading strategy based on high frequency stocks data.

Data:

```
[112]: df.head(10)
```

[112]:

	datetime	bid	bs	ask	as	mid	imb	wmid	spread	time
0	2021-01-05 09:00:00.000	128.40	988	128.42	1561	128.410	0.387603	128.407752	0.02	0
1	2021-01-05 09:00:00.100	128.40	1088	128.42	1561	128.410	0.410721	128.408214	0.02	1
2	2021-01-05 09:00:00.200	128.40	1088	128.42	1561	128.410	0.410721	128.408214	0.02	2
3	2021-01-05 09:00:00.300	128.40	1088	128.42	1561	128.410	0.410721	128.408214	0.02	3
4	2021-01-05 09:00:00.400	128.40	1088	128.42	1561	128.410	0.410721	128.408214	0.02	4
5	2021-01-05 09:00:00.500	128.39	12	128.42	1661	128.405	0.007173	128.390215	0.03	5
6	2021-01-05 09:00:00.600	128.39	12	128.42	1661	128.405	0.007173	128.390215	0.03	6
7	2021-01-05 09:00:00.700	128.39	12	128.42	1661	128.405	0.007173	128.390215	0.03	7
8	2021-01-05 09:00:00.800	128.39	12	128.42	1661	128.405	0.007173	128.390215	0.03	8
9	2021-01-05 09:00:00.900	128.39	12	128.42	1661	128.405	0.007173	128.390215	0.03	9

Figure: A glimpse into AAPL (2021-01-05)

Introduction (2)

Question: From the point of view of a market maker, what is the fair price given the state of the order book?

If we know that fair price $\hat{P}$, we are able to place an order (P^b_{t+1}, P^a_{t+1}) such that:

$$P^b_t \leq P^b_{t+1} \leq \hat{P} \leq P^a_{t+1} \leq P^a_t \tag{1}$$

Thus, we provide liquidity all the while being covered against price variations.

State of the Art

- Mid-Price

$$M_t = \frac{P_t^b + P_t^a}{2} \tag{2}$$

- Weighted Mid-Price

$$W_t = I_t P_t^a + (1 - I_t) P_t^b \tag{3}$$

$$I_t = \frac{Q_t^b}{Q_t^a + Q_t^b} \tag{4}$$

where P_t^a (resp. P_t^b) denotes the price at best ask (resp. best bid) and Q_t^a (resp. Q_t^b) denotes the volume at best ask (resp. at best bid)

State of the Art (2)

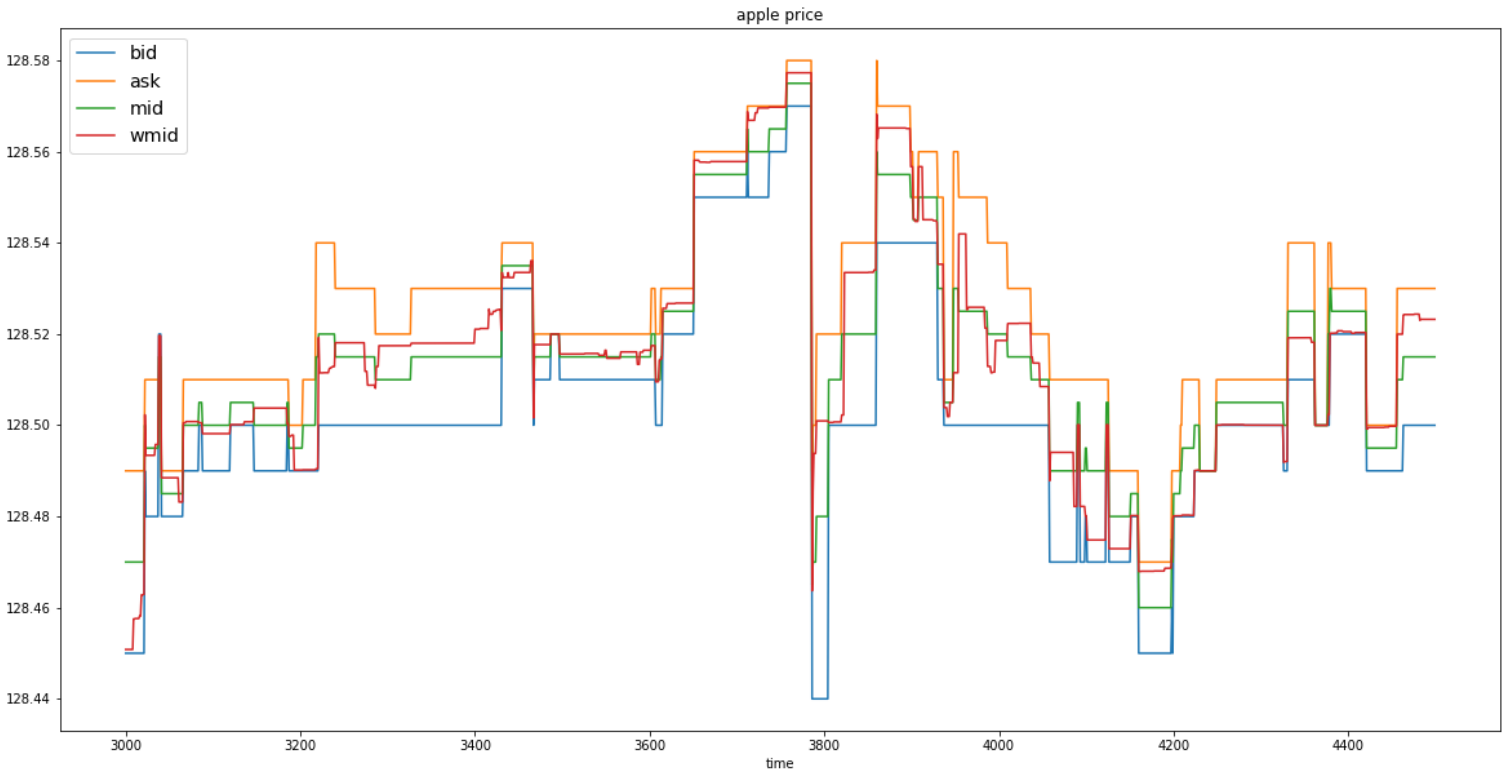


Figure: Bid, ask, mid and weighted mid prices AAPL stock (2021-01-05)

The Micro Price

Goal

Build a fair estimator of the price P_t of the stock at time t.

Definitions

- Times when the mid-price changes:

$$\tau_1 = \inf\{u > t, M_u - M_{u^-} \neq 0\} \tag{5}$$

$$\tau_{i+1} = \inf\{u > \tau_i, M_u - M_{u^-} \neq 0\} \tag{6}$$

- Micro-price:

$$P_t^{micro} := \lim_{i \to \infty} \mathrm{E}\left[M_{\tau_i} | \mathcal{F}_t\right] \tag{7}$$

Interpretation: If we are at time t, we consider that the fair price is the conditional expectation of future mid-prices based on the current state of the order book (analogy with Black-Scholes theory).

The Micro Price (2)

Assumptions

- The information in the order book is determined by the processes of the mid, the imbalance and the spread:

$$\mathcal{F}_t = \sigma(M_s, I_s, S_s; s \leq t) \tag{8}$$

- Mid price increments are independent from mid-price level:

$$\begin{aligned} \mathrm{E}[M_{\tau_{i+1}} - M_{\tau_i} | M_t = M, I_t = I, S_t = S] = \\ \mathrm{E}\left[M_{\tau_{i+1}} - M_{\tau_i} | I_t = I, S_t = S\right] \end{aligned}$$

The Micro Price (3)

Theorem:

Given these two assumptions, the prediction of the i^{th} mid-price can be written as:

$$\mathrm{E}\left[M_{\tau_i}|\mathcal{F}_t\right] = M_t + \sum_{k=1}^{i} g^k(I_t, S_t) \tag{9}$$

where

$$g^1(I, S) = \mathrm{E}\left[M_{\tau_1} - M_t | I_t = I, S_t = S\right] \tag{10}$$

and

$$g^{i+1}(I, S) = \mathrm{E}\left[g^i(I_{\tau_1}, S_{\tau_1}) | I_t = I, S_t = S\right] \tag{11}$$

The Micro Price (4)

The finite-space model

We denote by $X_t := (I_t, S_t)$ the state of the order book. For a finite number of possible values of the imbalance I_t and spread S_t, we can express the micro-price as:

$$P_t^{micro} = M_t + \sum_{k=1}^{\infty} B^k G^1 \tag{12}$$

where:

$$B := (I - Q)^{-1} T \tag{13}$$

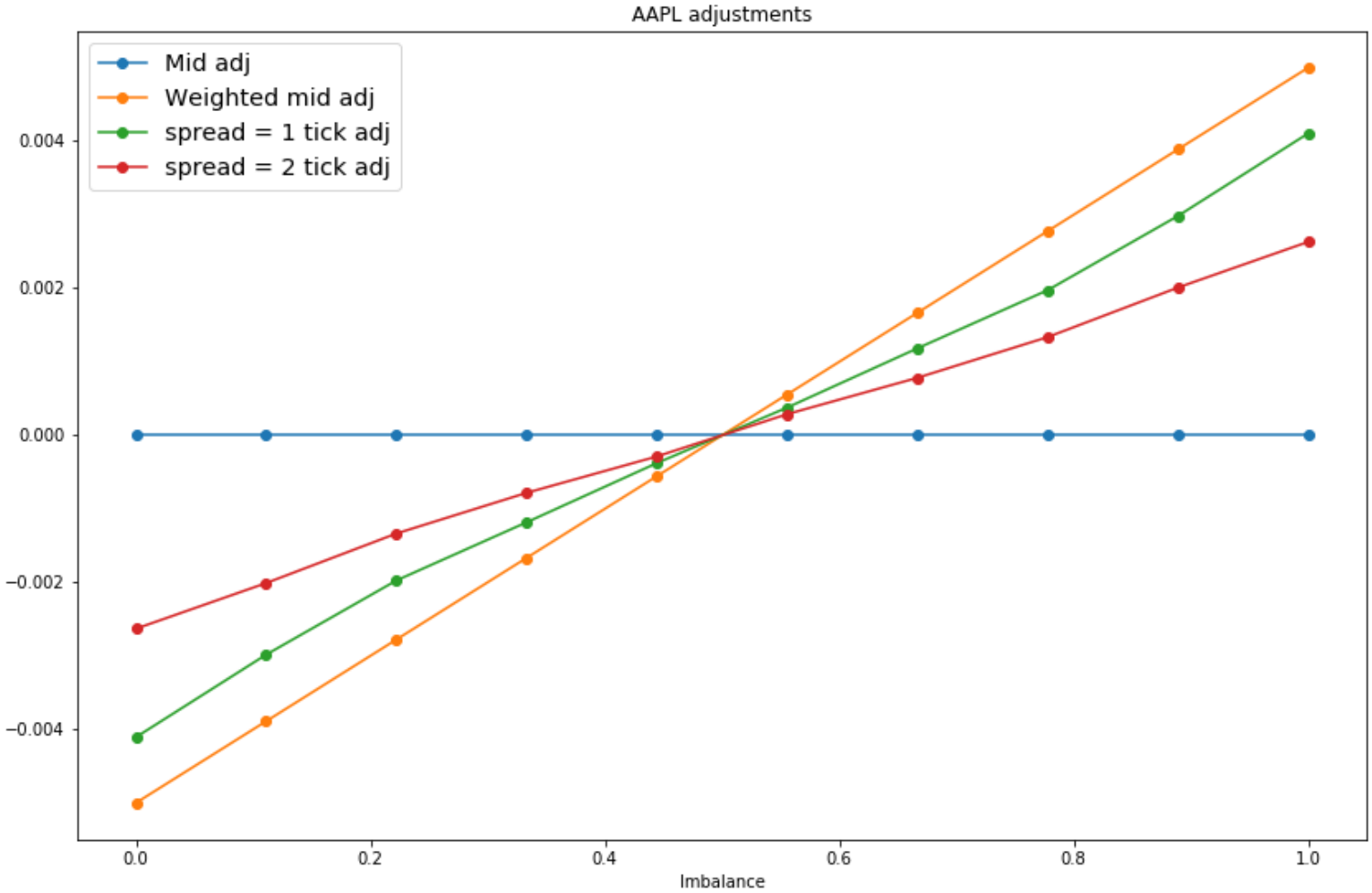
$$G^1 := (I - Q)^{-1} RK \tag{14}$$

$$Q_{xy} := \mathrm{P}(\mathrm{M_{t+1}} - \mathrm{M_t} = 0 \cap \mathrm{X_{t+1}} = \mathrm{y} | \mathrm{X_t} = \mathrm{x}) \tag{15}$$

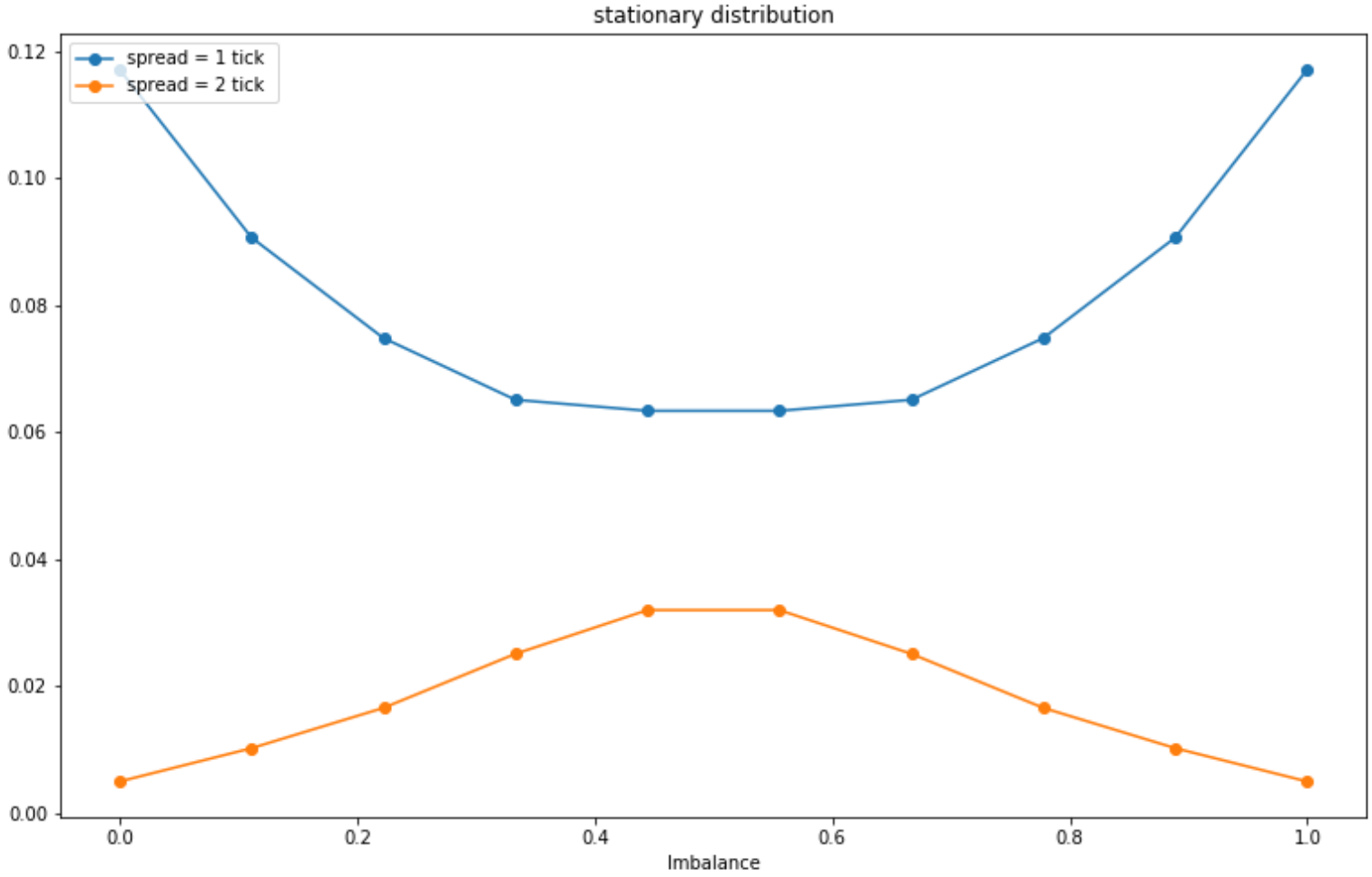
$$T_{xy} := \mathrm{P}(\mathrm{M_{t+1}} - \mathrm{M_t} \neq 0 \cap \mathrm{X_{t+1}} = \mathrm{y} | \mathrm{X_t} = \mathrm{x}) \tag{16}$$

$$R_{xk} := \mathrm{P}(\mathrm{M_{t+1}} - \mathrm{M_t} = \mathrm{k} | \mathrm{X_t} = \mathrm{x}) \tag{17}$$

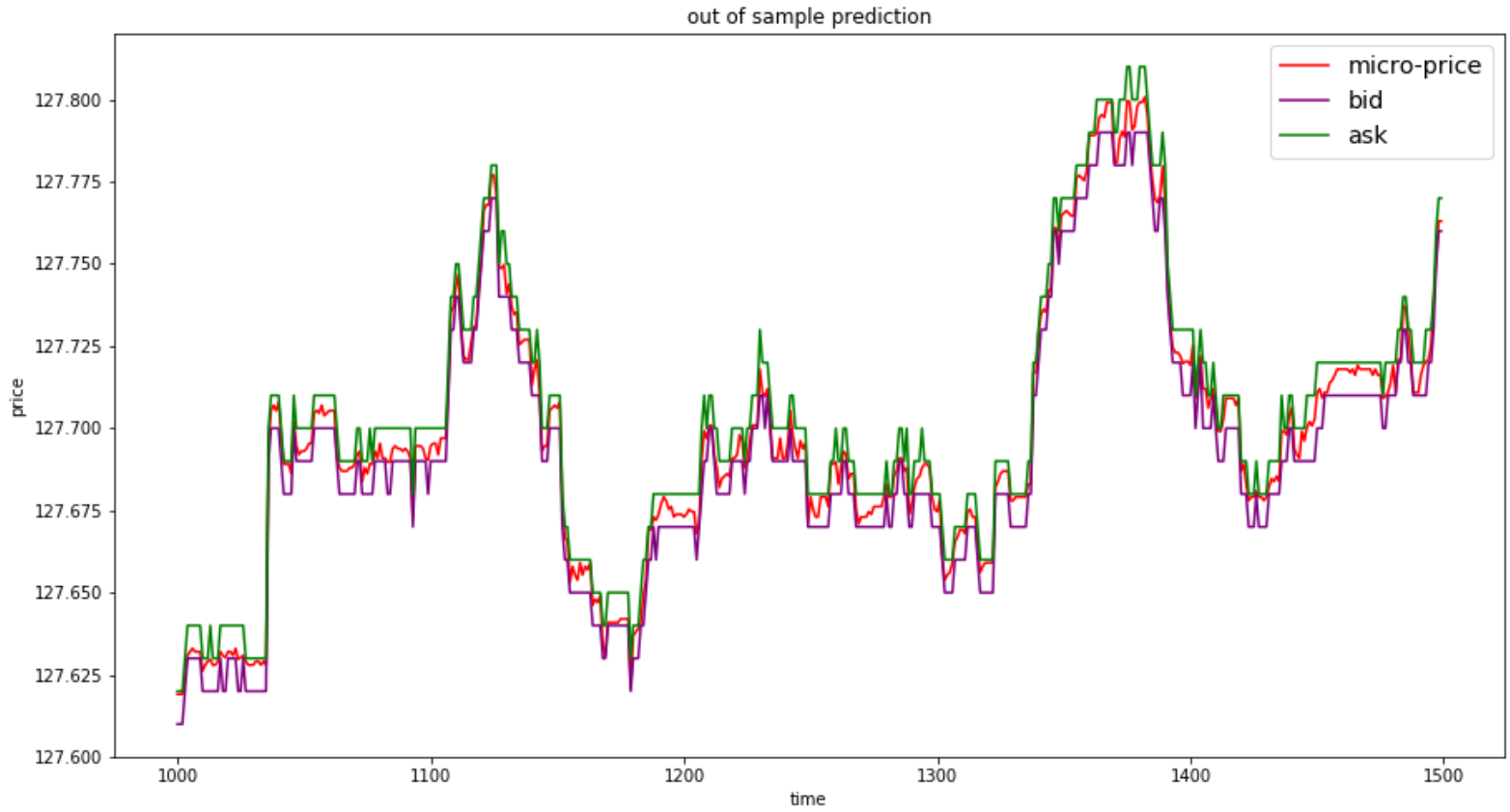
Results adjustment



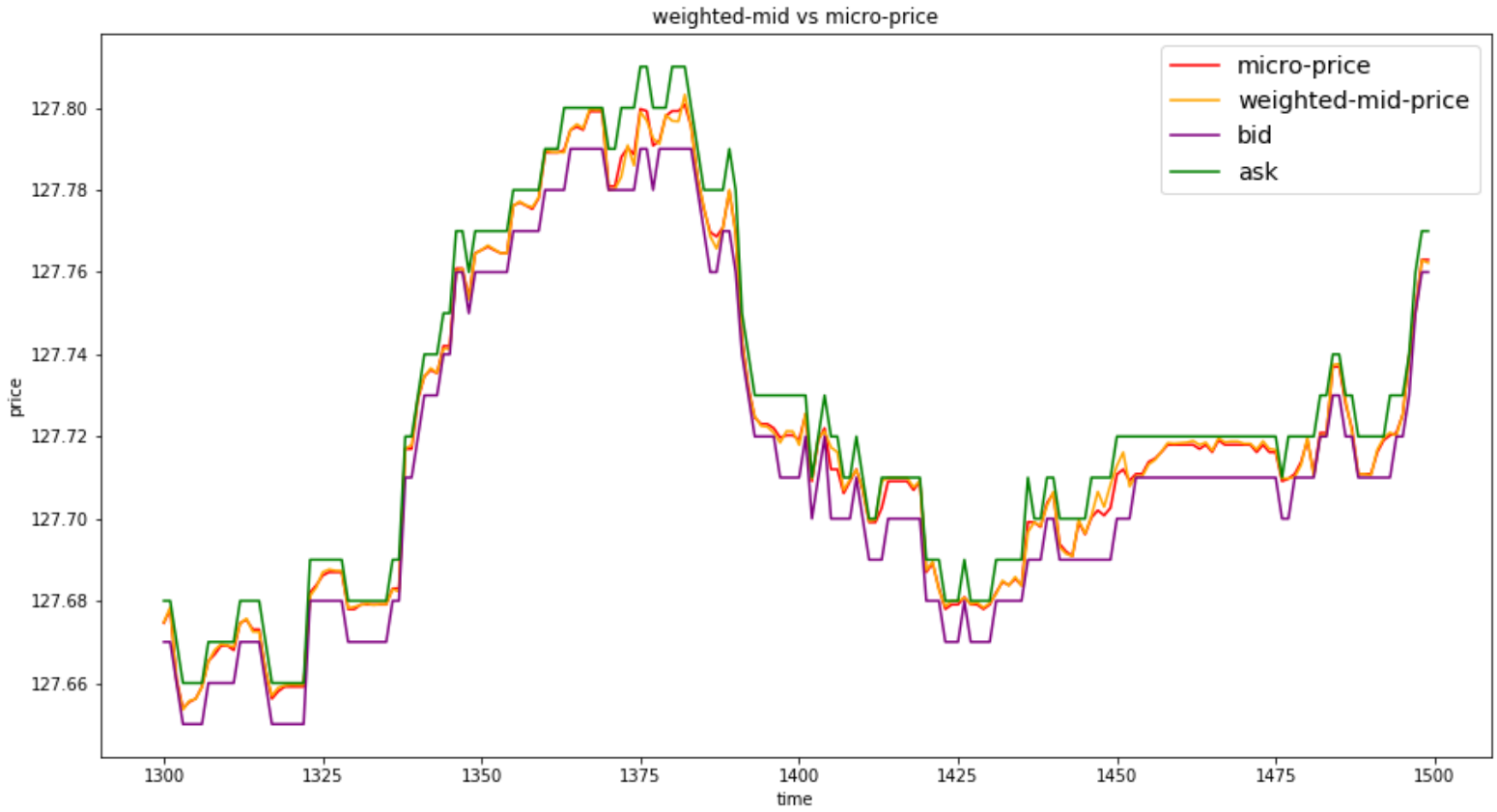
Results state distribution



Out of sample Predictions



Out of sample wMid vs MicroP



Next Steps

- Explore ways to limit the impact of an order using the microprice as future price.
- Use this micro-price as a feature for prediction problems in HFT.