TAQ Write-up

Sihan Qi, Sihan Liu, Ruihan Zhuang

March 2022

1 Building an Impact Model from Public Data

Please refer to readme file for a structure of code and tests.

We used the 200 most liquid stocks in sp500, sorted using their number of trades.

After reading in the data, we dropped the Nan values. If one date have more than 5 missing value, then we drop that date; Otherwise, we drop the column. (We perform this process in both Find_Largest.py and Regression.py)

In Find_Largest.py, we dropped the stock "SUNW"; (the missing value is caused by absent of trade and quote files) In Regression.py, we dropped the stock "NE"; Because there are days in NE without trades or quotes.

1.1 Regression

1.1.1 i

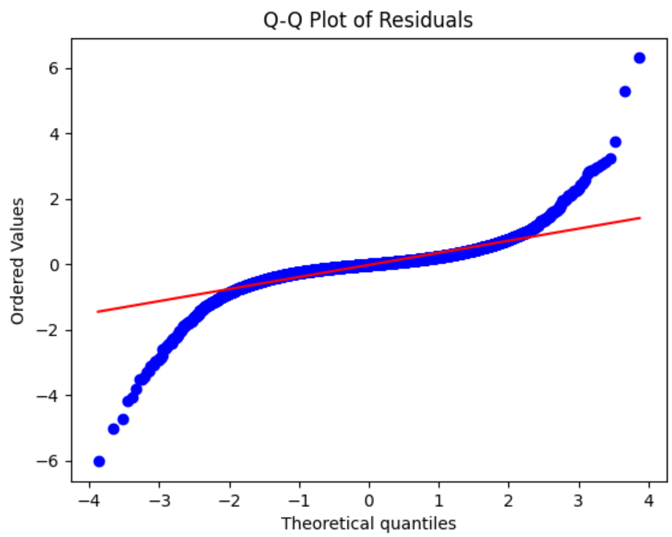
According to data in params_part1.txt:

The estimate of eta is statistically significant as in both bootstrap tests, the p value are smaller than 0.05; While that of beta is not because both greater than 0.05

1.1.2 ii

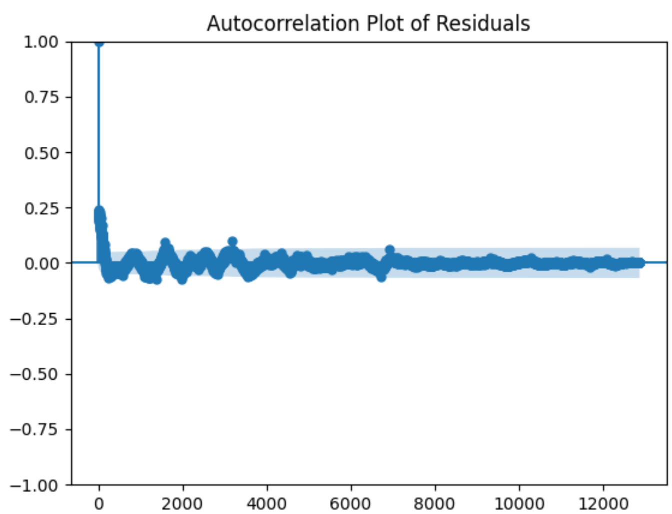
We examined the assumptions in residual_analysis.py:

The QQ_plot of the residual looks like:



The plot deviate significantly from a straight line, then the residuals may not be normally distributed.

We then conducted a independence test using the autocorrelation plot.



where points on the plot are basically randomly scattered around a horizontal line at zero, which means that the residual are independent.

1.1.3 iii

We performed the comparison in Regression$_{h}alf.py$

Here are the eta and beta values of the largest half stocks (_left)
eta_left: 0.05569069617052793
beta_left): 4.742905111034668e-12
eta_right: 0.3688050355026979
beta_right: 0.3772504093461929

They are significantly different.

1.1.4 iv

In residual_analysis.py, we performed a Goldfeld-Quandt test for homoscedasticity. The result is:

```
Goldfeld-Quandt Test:
F test statistic: 6.829
F test p-value: 0.009
```

Since the p-value is smaller than 0.05, we conclude that the residuals are not homoscedastic.

2 Optimal Execution

2.1 Part (a)

2.1.1 Part 1

By using HJB equation to the optimal liquidation problem, we have that

$$\partial_t H(t,Q,S,x) + sup_v(\mathcal{L}_t^v H(t,Q,S,x) - \phi Q_t^2) = 0$$

We omit the under-script t and super-script v for simplicity if there is no ambiguity in the following argument.

Note that if we write the state varibles evolution in multi-deminsional version, we have

$$d\begin{pmatrix} Q \\ S \\ x \end{pmatrix} = \begin{pmatrix} -v \\ -g(v) \\ v\hat{S}\partial_x \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} dW_t$$

Note that in our setting $g(v) = bv$, $\hat{S} = S - f(v)$ and $f(v) = kv$.

Applying multi-dimensional Ito's formula, we can simplify the HJB function in the following way:

$$\begin{aligned}
0 &= \partial_t H(t,Q,S,x) + sup_v(\mathcal{L}_t^v H(t,Q,S,x) - \phi Q_t^2) \\
&= \partial_t H - \phi q^2 + sup_v((-v,-g(v),v\hat{S})DH + \frac{1}{2}(0,\sigma,0)\begin{pmatrix} 0 \\ \sigma \\ 0 \end{pmatrix} D^2 H) \\
&= \partial_t H - \phi q^2 + sup_v((-v\partial_q - g(v)\partial_S + v\hat{S}\partial_x)H + \frac{1}{2}\sigma^2\partial_{SS}H) \\
&= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})H - \phi q^2 + sup_v((-v\partial_q - bv\partial_S + v(S-kv)\partial_x)H)
\end{aligned}$$

with terminal condition $H(T,q,S,x) = x + Sq - \alpha q^2$.

Note that

$$\begin{aligned}
&(-v\partial_q - bv\partial_S + v(S-kv)\partial_x)H \\
=& -k\partial_x H v^2 - (\partial_q + b\partial_S - S\partial_x)Hv \\
=& -k\partial_x H(v + \frac{(\partial_q + b\partial_S - S\partial_x)H}{2k\partial_x H})^2 + \frac{((\partial_q + b\partial_S - S\partial_x)H)^2}{4k\partial_x H}
\end{aligned}$$

Therefore, to maximize $(-v\partial_q - bv\partial_S + v(S-kv)\partial_x)H$ w.r.t v, we have that the optimal $v^* = -\frac{(\partial_q + b\partial_S - S\partial_x)H}{2k\partial_x H}$, as we desire.

Now we make the ansatz $H(t,q,S,x) = x + Sq + h(t,q,S)$ with $h(T,q,S) =$

$-\alpha q^2$. Inserting to the HJB equation and we obtain:

$$
\begin{aligned}
0 &= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})H - \phi q^2 + sup_v((-v\partial_q - bv\partial_S + v(S-kv)\partial_x)H) \\
&= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})(x + Sq + h) - \phi q^2 + \frac{((\partial_q + b\partial_S - S\partial_x)H)^2}{4k\partial_x H} \\
&= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})h - \phi q^2 + \frac{1}{4k}(b(q + \partial_S h) + (S + \partial_q h) - S)^2 \\
&= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})h - \phi q^2 + \frac{1}{4k}(b(q + \partial_S h) + \partial_q h)^2
\end{aligned}
$$

Since this PDE does not explicitly depend on S and the terminal condition is independent of S, we have $\partial_S h(t,q,S) = 0$ and $h(t,q,S) = h(t,q)$, using this fact, we can further simplify our optimal control and HJB function into the following 2 equations:

$$
\begin{aligned}
v^* &= -\frac{(\partial_q + b\partial_S - S\partial_x)H}{2k\partial_x H} \\
&= -\frac{\partial_q h + bq}{2k}
\end{aligned}
$$

and

$$
\begin{aligned}
0 &= (\partial_t + \frac{1}{2}\sigma^2\partial_{SS})h - \phi q^2 + \frac{1}{4k}(b(q + \partial_S h) + \partial_q h)^2 \\
&= \partial_t h - \phi q^2 + \frac{1}{4k}(bq + \partial_q h)^2
\end{aligned}
$$

Let's make the separation of variable ansatz $h(t,q) = h_2(t)q^2$, plugging into the $\partial_t h - \phi q^2 + \frac{1}{4k}(bq + \partial_q h)^2 = 0$, and we obtain,

$$
\begin{aligned}
0 &= \partial_t h - \phi q^2 + \frac{1}{4k}(bq + \partial_q h)^2 \\
&= q^2\partial_t h_2 - \phi q^2 + \frac{1}{4k}(bq + 2h_2(t)q)^2
\end{aligned}
$$

Devide q^2 on both side of the equation (since if $q = 0$, this case is trivial and is not meaningful in our practical case), we have $\partial_t h_2 - \phi + \frac{1}{4k}(b + 2h_2(t))^2 = 0$ with $h_2(T) = -\alpha$.

Now we set $a = \phi - \frac{b^2}{4k}$, $d = -\frac{b}{k}$ and $c = -\frac{1}{k}$, then the ODE becomes $\partial_t h_2 = a + dh_2 + ch_2^2$. Let $u = ch_2$, plugging in and we will get $u' = ca + du + u^2$.

Then let $u = \frac{-w'}{w}$, then

$$
\begin{aligned}
u' &= -\frac{w''}{w} + (\frac{w'}{w})^2 \\
&= -\frac{w''}{w} + u^2
\end{aligned}
$$

Therefore, $-\frac{w''}{w} = u' - u^2 = du + ca = d\frac{-w'}{w} + ca$, thus $w'' - dw' + caw = 0$, i.e. $w'' + \frac{b}{k}w' + (\frac{b^2}{4k^2} - \frac{\phi}{k})w = 0$. Now consider $\lambda^2 + \frac{b}{k}\lambda + (\frac{b^2}{4k^2} - \frac{\phi}{k})w = 0$, since solution to this quadratic equation is $\lambda_{1,2} = \frac{-\frac{b}{k} \pm \sqrt{\frac{4\phi}{k}}}{2}$, therefore, $w = Ae^{\lambda_1 t} + Be^{\lambda_2 t}$, insert in $h_2 = -ku = k\frac{w'}{w}$, with boundary condition $h_2(T) = -\alpha$

Then we have $h_2 = k\frac{\lambda_1 \frac{A}{B} e^{(\lambda_1-\lambda_2)t} + \lambda_2}{\frac{A}{B} e^{(\lambda_1-\lambda_2)t} + 1}$ and set $t = T$, we can deduce form the boundary condition that $\frac{B}{A} = -\frac{\lambda_1 + \frac{\alpha}{k}}{\lambda_2 + \frac{\alpha}{k}} e^{(\lambda_1-\lambda_2)T}$. Then let $\gamma = \sqrt{\frac{\phi}{k}}$ and let $\zeta = \frac{\lambda_1 + \frac{\alpha}{k}}{\lambda_2 + \frac{\alpha}{k}} = \frac{\alpha - \frac{1}{2}b + \sqrt{k\phi}}{\alpha - \frac{1}{2}b - \sqrt{k\phi}}$, note that$\lambda_1 - \lambda_2 = 2\gamma$, $\lambda_1 + \lambda_2 = -\frac{b}{k}$ we obtained:

$$\begin{aligned}
h_2(t) &= k\frac{\lambda_1 - \lambda_2 \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} \\
&= k\frac{(\frac{\lambda_1+\lambda_2}{2} + \frac{\lambda_1-\lambda_2}{2}) - (\frac{\lambda_1+\lambda_2}{2} - \frac{\lambda_1-\lambda_2}{2})\zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} \\
&= k(-\frac{b}{2k} + \gamma\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}}) \\
&= -\frac{b}{2} + \sqrt{k\phi}\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}}
\end{aligned}$$

Then

$$\begin{aligned}
v_t^* &= -\frac{\partial_q h + bq}{2k} \\
&= -\frac{2h_2 q + bq}{2k} \\
&= -\frac{1}{2k}(-b + 2\sqrt{k\phi}\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} - b)q \\
&= -\gamma\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} Q_t^{v^*}
\end{aligned}$$

as we desire.

What's more, since $dQ_t^v = -v_t dt$, with $Q_0^v = q_0$, we have

$$\begin{aligned}
Q_t^{v^*\prime} &= \gamma\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} Q_t^{v^*} \\
\frac{Q_t^{v^*\prime}}{Q_t^{v^*}} &= \gamma\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} \\
\log(Q_t^{v^*})' &= \gamma\frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} \\
\log(Q_t^{v^*}) &= \int_0^t \gamma\frac{1 + \zeta e^{2\gamma(T-s)}}{1 - \zeta e^{2\gamma(T-s)}} ds + c \\
&= \frac{\ln(\zeta e^{2\gamma(T-t)} - 1)}{\gamma} - \frac{\ln(\zeta e^{2\gamma T} - 1)}{\gamma} + c
\end{aligned}$$

Using $Q_0^v = q_0$ and we obtain the formula for $Q_t^{v^*}$:

$$Q_t^{v^*} = \frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} q_0$$

as we desire.

2.1.2 Part 2

As $\phi \to 0$, we apply L'Hôpital's rule to both expression:

$$
\begin{aligned}
\lim_{\phi \to 0} Q_t^{v^*} &= \lim_{\phi \to 0} \frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} q_0 \\
&= \lim_{\phi \to 0} \frac{\frac{-((T-t)k\phi - (T-t)(\alpha - 0.5b)^2 - 2(\alpha-0.5b)k)e^{(T-t)\sqrt{\frac{\phi}{k}}}}{2\sqrt{k\phi}(\sqrt{k\phi}+\alpha-0.5b)^2} + \frac{(T-t)e^{-(T-t)\sqrt{\frac{\phi}{k}}}}{2k\sqrt{\frac{\phi}{k}}}}{\frac{-(Tk\phi - T(\alpha-0.5b)^2 - 2(\alpha-0.5b)k)e^{T\sqrt{\frac{\phi}{k}}}}{2\sqrt{k\phi}(\sqrt{k\phi}+\alpha-0.5b)^2} + \frac{Te^{-T\sqrt{\frac{\phi}{k}}}}{2k\sqrt{\frac{\phi}{k}}}} q_0 \\
&= \lim_{\phi \to 0} \frac{\frac{(T-t)e^{-(T-t)\sqrt{\frac{\phi}{k}}}}{2k\sqrt{\frac{\phi}{k}}}}{\frac{Te^{-T\sqrt{\frac{\phi}{k}}}}{2k\sqrt{\frac{\phi}{k}}}} q_0 \\
&= \frac{T-t}{T} q_0
\end{aligned}
$$

and

$$
\begin{aligned}
\lim_{\phi \to 0} v_t^* &= -\gamma \frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} Q_t^{v^*} \\
&= \lim_{\phi \to 0} -\gamma \frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} \frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} q_0 \\
&= \lim_{\phi \to 0} -\gamma \frac{\zeta e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} q_0 \\
&= \frac{1}{T} q_0
\end{aligned}
$$

Intuitively, when$\phi \to 0$, we do not penalize the inventory during the execution, but only penalize the inventory at the end. Therefore, to maximize utility, we follow the suggestion by Almgren-Chriss model and use Small Delta Continuous Trading, which is our optimal liquidation rate.

2.2 Part (b)

2.2.1 Part 1

By using HJB equation to the optimal liquidation problem, we have that

$$\partial_t H(t, S, q) + sup_v(\mathcal{L}_t^v H(t, S, q) - (S - kv)v) = 0$$

Note that if we write the state varibles evolution in multi-deminsional version, we have

$$d \begin{pmatrix} S \\ q \end{pmatrix} = \begin{pmatrix} 0 \\ -v \end{pmatrix} dt + \begin{pmatrix} \sigma \\ 0 \end{pmatrix} dW_t$$

Applying multi-dimensional Ito's formula, we can simplify the HJB function in the following way:

$$
\begin{aligned}
0 &= \partial_t H(t,Q,S,x) + sup_v(\mathcal{L}_t^v H(t,Q,S,x) + (S-kv)v) \\
&= \partial_t H + sup_v((0,-v)DH + \frac{1}{2}(\sigma,0)\begin{pmatrix}\sigma\\0\end{pmatrix} D^2H + (S-kv)v) \\
&= \partial_t H + sup_v(-v\partial_q H + \frac{1}{2}\sigma^2\partial_{SS}H + (S-kv)v) \\
&= \partial_t H + sup_v(-kv^2 + (S-\partial_q H)v + \frac{1}{2}\sigma^2\partial_{SS}H) \\
&= \partial_t H + sup_v(-k(v+\frac{\partial_q H - S}{2k})^2 + \frac{(\partial_q H - S)^2}{4k} + \frac{1}{2}\sigma^2\partial_{SS}H) \\
&= \partial_t H + \frac{(\partial_q H - S)^2}{4k} + \frac{1}{2}\sigma^2\partial_{SS}H
\end{aligned}
$$

We time $4k$ on both side and we obtain $0 = 4k\partial_t H + (\partial_q H - S)^2 + 2k\sigma^2\partial_{SS}H$ and $v^* = -\frac{\partial_q H - S}{2k}$ is our optimal liquidation rate.

2.2.2 Part 2

$v^* = -\frac{\partial_q H - S}{2k}$ is our optimal liquidation rate, as it gives us the maximum to the quadratic form in the sup. Now, in order to get a more explicit expression to v^*, we need to solve H, i.e. the HJB function $0 = 4k\partial_t H + (\partial_q H - S)^2 + 2k\sigma^2\partial_{SS}H$.

We use the ansatz $H(t,S,q) = h_2(t)q^2 + h_1(t)q + h_0(t) + qS$, while $H(T,S,q) = qS - \alpha q^2$, since h_2, h_1, h_0 are not depend on q, therefore $h_1(t) = h_0(t) = 0$ and $h_2(T) = -\alpha$. Inserting this into the ODE and we obtain:

$$\partial_t h_2 + \frac{1}{k}h_2^2 = 0$$

Let $u = -\frac{1}{k}h_2$ and we have $u' = u^2$. Then let $u = \frac{-w'}{w}$ and we obtain that $\frac{-w''}{w} = u' - u^2 = 0$, thus $w'' = 0$ and we have $w(t) = c_1 t + c_2$, then $h_2 = -ku = k\frac{w'}{w} = \frac{kc_1}{c_1 t + c_2} = \frac{k}{t + \frac{c_2}{c_1}}$. Note that $h_2(T) = -\alpha$, therefore, we have $\frac{c_2}{c_1} = -\frac{k}{\alpha} - T$. Therefore, $h_2(t) = \frac{k}{(t-T)-\frac{k}{\alpha}}$. Back to the ansatz and we have $H(t,S,q) = \frac{k}{(t-T)-\frac{k}{\alpha}}q^2 + qS$.

Plugging into $v^* = -\frac{\partial_q H - S}{2k}$ and we obtained $v_t^* = \frac{\alpha}{\alpha(T-t)+k}Q_t^{v^*}$

2.2.3 Part 3

Since $dQ_t^v = -v_t dt$, with $Q_0^v = q_0$, we have

$$
\begin{aligned}
Q_t^{v^*\prime} &= -\frac{\alpha}{\alpha(T-t)+k}Q_t^{v^*} \\
\log(Q_t^{v^*}) &= \int_0^t -\frac{\alpha}{\alpha(T-s)+k}ds + c \\
&= \ln(a(T-t)+k) - \ln(aT+k)
\end{aligned}
$$

Thus $Q_t^{v^*} = \frac{\alpha(T-t)+k}{\alpha T+k} q_0$ as the boundary condition let us know that $Q_0^{v^*} = q_0$. Therefore

$$\begin{aligned} v_t^* &= \frac{\alpha}{\alpha(T-t)+k} Q_t^{v^*} \\ &= \frac{\alpha}{\alpha(T-t)+k} \frac{\alpha(T-t)+k}{\alpha T+k} q_0 \\ &= \frac{\alpha}{\alpha T+k} q_0 \end{aligned}$$

2.2.4 Part 4

As $\alpha \to \infty$, $v_t^* = \frac{\alpha}{\alpha T+k} q_0 \to \frac{q_0}{T}$. Intuitively, we penalize the remaining inventory (as long as the inventory is not 0 at time T) to $-\infty$ as $\alpha \to \infty$, therefore the only thing we care is to make inventory to be 0 at time T, and then the optimal liquidation rate under the only restriction that inventory is 0 at time T is to make Small Delta Continuous Trading, which is exactly what Almgren-Chriss model told us.

3 Standard Concepts of Statistical Trading

(a) Describe four of the most commonly used high-frequency trading strategies. Support you claim of those being "most common" by providing some references. Distinguish between "alpha" strategies and other strategies.

Based on the article "Concept Release on Equity Market Structure" as mentioned in the homework instruction, I got the following answer.

- **Passive Market Making** "Passive market making primarily involves the submission of non-marketable resting orders (bids and offers) that provide liquidity to the marketplace at specified prices. While the proprietary firm engaging in passive market making may sometimes take liquidity if necessary to liquidate a position rapidly, the primary sources of profits are from earning the spread by buying at the bid and selling at the offer and capturing any liquidity rebates offered by trading centers to liquidity-supplying orders".
- **Arbitrage** "An arbitrage strategy seeks to capture pricing inefficiencies between related products or markets. For example, the strategy may seek to identify discrepancies between the price of an ETF and the underlying basket of stocks and buy (sell) the ETF and simultaneously sell (buy) the underlying basket to capture the price difference."
- **Structural** "Some proprietary firm strategies may exploit structural vulnerabilities in the market or in certain market participants. For example, by obtaining the fastest delivery of market data through co-location arrangements and individual trading center data feeds (discussed below in section IV.B.2.), proprietary firms theoretically could profit by identifying market participants who are offering executions at stale prices."
- **Directional** "There may, however, be a wide variety of short-term strategies that anticipate such a movement in prices. Some "directional" strategies may be as straightforward as concluding that a stock price temporarily has moved away from its "fundamental value" and establishing a position in anticipation that the price will return to such value."

"Alpha" strategies is the contrary of "beta" strategies. It means that you are trying to bit the market, and to earn profit exceeding overall market return.

(b) Provide a "back of the envelope" estimate of the profitability of high frequency traders in today's equity market. How do they use leverage? Motivate your assumptions, and how you come to your conclusion.

Here I discuss profits of market making style high frequency traders. All the data are obtained from the article "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market".

At the table 2 of the paper, the authors mentioned that on average, there are 446,340 trades taking place on a day and the high frequency traders are responsible for 32.56% trades, which is 145328. Besides, there are 15 traders on average for a day and that trade size is 5.69. Then, we can estimate that on average, a high frequency trader trades 55127.75 units per day. As mentioned in reference 5, some of trades are 'scratch' trades, which does not give high frequency traders profit. In table 8, the author mentioned that 2.84% trades are 'scratch' trades for high frequency traders. Then, the trades a high frequency traders make profits from should be around 53562 unites.

Assume that the bid-ask spread is 0.01 dollar, which is the 'tick' size of equity market. This gives 535.62 dollars per day.

Assume there are 252 trading days every day. Then yearly profit of a high frequency trader is 134,976.24.

For the high frequency traders who are doing market making, I don't think they are using leverage.

(c) Does high frequency trading impose risks of systemic nature? Find a journal paper or white paper no older than two years, that addresses this question. Do you agree or disagree with their findings?

The journal paper I found is *What (If Anything) is Wrong with High-Frequency Trading?*. It is published on 7 May 2022 on Journal of Business Ethics. https://doi.org/10.1007/s10551-022-05145-7

In this paper, the authors argue that "I find that it is by no means easy to make a convincing analytical case that LLT is related to systemic risk (let alone that individual traders are to blame)—or that, empirically speaking, the threat of harm is substantial" (page 2).

There are two main arguments made by the authors. First, they argued that "Note that the transactions effected by low-latency traders do not resemble the transactions which have contributed to the 2008 financial crisis. Arguably, risky transactions featuring not information asymmetries but expectation asymmetries played a significant role in causing that crisis". It means that, they think systemic risk are caused by expectation asymmetries, which is not what high frequency trader impose to the market.

Second, they said that when looking back on the flash crash event, "it is easy to see that what characterizes flash crashes is not only that stock prices fall at unprecedented speeds. They also bounce back rather quickly" (page 12). Further on, they said that events like flash crash do not produce significant effect like the crisis on 2008. People do not panic when facing flash crash, because they know it is because of some trading algorithm rather than some genuine bad news.

I agree with the authors. I don't think high frequency traders really produce any risk of significant impact. Furthermore, I don't think high frequency traders are producing high volatility environment. As shown in the paper

"The Flash Crash: The Impact of High Frequency Trading on an Electronic Market", even though high frequency traders creates large volume of trading during the flash crash, they are just trading with each other.

(d) Propose your own intraday equity "alpha" trading strategy. Describe what methodology and data you would use to research your idea. Provide a "guestimate" of its performance (Sharpe ratio).

I would just adopt the market maker high frequency strategy. For example, when there is selling pressure. I will deplete the best bid level of orders and set new bid and ask spread. If the selling pressure continues, my bid order will be executed with a lower price than my previous trade and then I make a profit. If the selling pressure stops, then I will try to buy with a same price as the previous trade, and then I do not lose.

I would use historical high frequency data such as TAQ to backtest my strategy.

I don't think its performance would be good, because there are already a lot of people using this strategy in the market. I guess my Sharpe ratio would be less than 1.