Homework 2 Technical Write-up

Contributors: Ziheng Chen (zc2068), Cong Wang (cw3140), Chengyang Gu (cg3765)

Problem 1

We focus on the estimation of related statistics in this write-up. For an illustration of the code, see readme.

Problem 1 (b)

1. Parameters and t-statistics

We used the S&P stocks to run the nonlinear regression. Out of the 508 stocks, we kept 493 stocks in the end after removing those that have missing value for a few days. Our parameter estimated are as followed and also attached in the text file called “params_part1.txt”. Only eta and beta are estimated as we only regress the temporary impact on the inputs: sigma, X and V.

We conclude our results in the table below.

	Estimation	T-statistics
η	4.092	3.153
β	0.574	4.585

2. Formula and Statistical Tests

We want to estimate a nonlinear regression function of the form

$$y = m(x, \beta) + \epsilon, \beta \in \Theta \subseteq \mathbb{R}^p$$

where Θ is the "permissible" parameter space.

The nonlinear least squares estimator beta is found by solving

$$min_{\beta\in\theta} SSR(\beta) := min_{\beta\in\theta} \frac{1}{2}\sum_{i=1}^{n}(y_i - m(x_i, \beta))^2$$

In our case,

$$m = \sigma\eta sgn(X)|\frac{X}{V(6/6.5)}|^b$$

where $x = [\sigma, X, V]$ and $\beta = \{\eta, b\}$.

We use the scipy.optimize.curve_fit to find the optimal parameters $\beta = \{\eta, b\}$ which minimizes the sum of the squared residuals of $m(x, \beta) - y$.

The asymptotic standard error of each element of β is the square root of the appropriate diagonal element of the $var[\hat{\beta}_i]$ matrix. The heteroskedasticity-robust standard errors is the square root of

$$v_{HCE}[\hat{\beta}] = (X'X)^{-1}(X'diag(u_1^2, \ldots, u_n^2)X)(X'X)^{-1}$$

where u_i are residuals.

The output from the `scipy.optimize.curve` will output popt (optimal parameter) and pcov (estimated covariance of popt). The diagonals provide the variance of the parameter estimate. To compute one standard deviation errors on the parameters use we use `perr = np.sqrt(np.diag(pcov))` as shown in our code.

$$t = \frac{\hat{\beta} - 0}{\sqrt{(X'X)^{-1}(X'diag(u_1^2, \ldots, u_n^2)X)(X'X)^{-1}}}$$

Our estimated coefficients and t-statistics are as below. Based on NLS assumption, for hypothesis testing we can still use same t as for OLS. From t table, we find that our estimated parameters are statistically significant at 95% confidence level.

Alternatively, we tried to transform the nonlinear equation into linear equation as below such that we can take advantage of using OLS regression to estimate the parameters. Since we cannot take log for negative value, we drop those nan in the code.

$$\log h - \log \sigma = \log \eta + sgn(X)\beta \log |\frac{X}{(6/6.5)V}| + \epsilon$$

The linear regression enables us to show both regular standard errors and heteroskedasticity-robust standard error. First, we show the regular standard error t-statistics. Please note that this is the result from the linear regression, so we need to transform the intercept and coefficient back to eta and beta.

	Estimation	**T-statistics**
η	4.854	N/A
β	-0.002	-0.654 / -0.662 (robust)

However, the coefficient for beta is not statistically significant at 95% confidence level as shown in the following result. It may be due to the model specification that we dropped those temporary impact with negative value.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.000
```

```
Model:                            OLS   Adj. R-squared:                 -0.000
Method:                 Least Squares   F-statistic:                    0.4275
Date:                Mon, 22 Mar 2021   Prob (F-statistic):              0.513
Time:                        21:23:00   Log-Likelihood:                -22253.
No. Observations:               13124   AIC:                         4.451e+04
Df Residuals:                   13122   BIC:                         4.452e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.5798      0.012    137.192      0.000       1.557       1.602
x1            -0.0021      0.003     -0.654      0.513      -0.008       0.004
==============================================================================
Omnibus:                     2181.890   Durbin-Watson:                   1.549
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4645.860
Skew:                          -0.986   Prob(JB):                         0.00
Kurtosis:                       5.146   Cond. No.                         3.65
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
```

Next, we show the result with heteroskedasticity-robust standard error. The coefficients stays the same but still not statistically significant at 95% confidence level .

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.000
Method:                 Least Squares   F-statistic:                    0.4387
Date:                Mon, 22 Mar 2021   Prob (F-statistic):              0.508
Time:                        21:44:44   Log-Likelihood:                -22253.
No. Observations:               13124   AIC:                         4.451e+04
Df Residuals:                   13122   BIC:                         4.452e+04
Df Model:                           1
Covariance Type:                  HC0
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.5798      0.012    137.177      0.000       1.557       1.602
x1            -0.0021      0.003     -0.662      0.508      -0.008       0.004
==============================================================================
Omnibus:                     2181.890   Durbin-Watson:                   1.554
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4645.860
Skew:                          -0.986   Prob(JB):                         0.00
Kurtosis:                       5.146   Cond. No.                         3.65
```

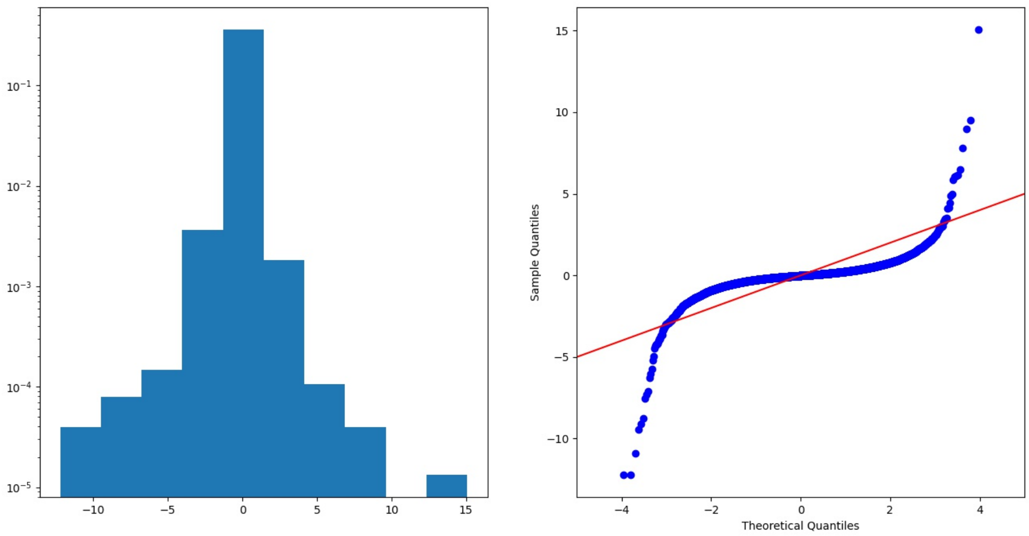
```
================================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC0)
```

3. Residual Tests

We computed the following statistics of the residuals. As we can see the residual has a mean close to 0 and standard deviation about 0.5. But it's highly skewed with a fat tail (leptokurtic).

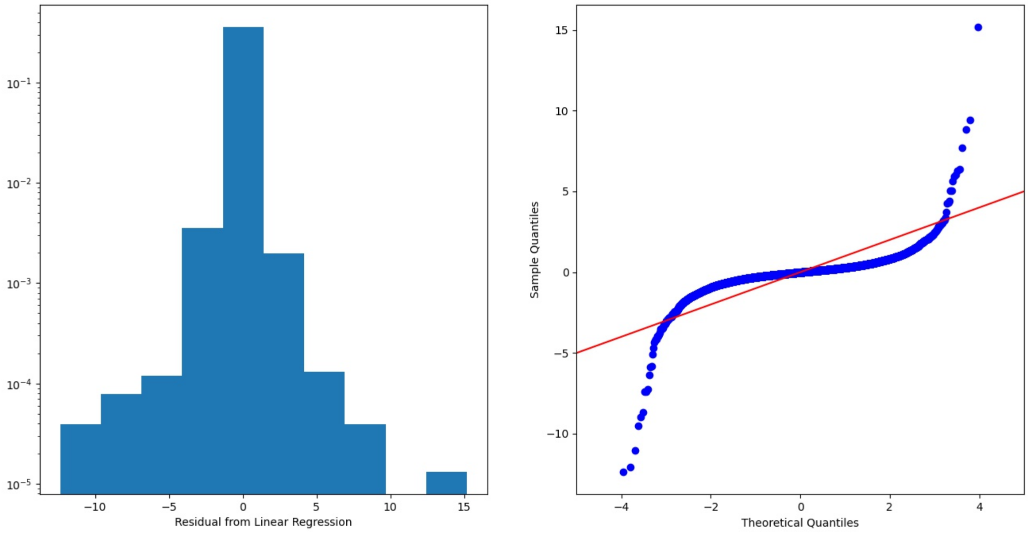
mean	std	skew	kurtosis
-0.024156	0.477624	-1.177237	125.860216

We also plotted the histogram and QQ-plot and confirm that our residual is not normally distributed. Still, given the number of samples and under moderate assumptions, we still believe the estimations are significantly away from zero.



Next, we ran a residual analysis for the result from the linear regression model. Below is the statistics of the linear regression model residual. It's actually very close to what we have from the nonlinear regression model.

Mean	Standard deviation	Skew	Kurtosis
-0.022389	0.494974	-1.064469	109.373691



4. Cross-validation of Estimation

We split our dataset into two parts based on the median daily volume. The top half has daily volume greater than the median while the bottom half has daily volume smaller than median. Then we run NLS and have the following result:

Liquid Stocks	Estimation	T-statistics
η	5.766	2.875
β	0.654	4.818

Illiquid Stocks	Estimation	T-statistics
η	3.370	1.846
β	0.550	2.458

As we can see from the result, both eta and beta for liquid stocks are larger than those from illiquid stocks.

5. Heteroskedasticity Tests

In statistics, the White test is a statistical test that establishes whether the variance of the errors in a regression model is constant: that is for homoskedasticity.

The White test is based on the estimation of the following:

$$\hat{\epsilon}^2 = \alpha_0 X_1 + \ldots + \alpha_p X_p + \gamma_1 X_1^2 + \ldots + \gamma_p X_p^2 + \ldots$$

Alternatively, in White test, we can use the following formula:

$$\hat{u}^2 = b_0 + b_1\hat{y} + b_2\hat{y}^2$$

The null hypothesis is that it doesn't have heteroskedasticity

$$H_o : b_0 = b_1 = 0$$

Then we perform F-test or LM test. If we reject the null hypothesis, then it indicates that it does have heteroskedasticity.

We performed White test to our nonlinear regression using the regression

$$\hat{u}^2 = b_0 + b_1\hat{y} + b_2\hat{y}^2$$

In our case, we plugged the estimated eta and beta to the following equation to compute $\hat{y}$

$$m = \sigma\eta sgn(X)|\frac{X}{V(6/6.5)}|^b$$

Then by using statsmodels module's white test, we have the following result. We found that both LM and F-test is statistically significant at 95% level and thus reject the null hypothesis. So it indicates that our residuals are heteroskedastic. This also testifies that we need to use robust standard errors for statistical tests.

```
The lagrange multiplier statistic=16.335700936737837,
p_value(LM)=0.0002836270269319864
Fvalue=8.171798182793141, p_value(F)=0.0002831935215390168
```

Problem 2 (a)

i

The corresponding HJB equation to this problem is

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{SS}H - \phi\nu^2 + \sup_{\nu}\{\nu(S - k\nu)\partial_x H - b\nu\partial_S H - v\partial_q H\} = 0$$

We only look at the sup part, we can find that it is in fact in the form of

$$-ax^2 + bx = 0$$

and we can easily see that this can be rewritten as

$$-a(x - \frac{b}{2a})^2 + \frac{b^2}{4a}$$

Hence, we can know that this part takes supreme when ν equals $\frac{\partial_x HS - b\partial_S H - \partial_q H}{2\partial_x Hk}$,

and this part has supereme value equals $\frac{(\partial_x HS - b\partial_x H - \partial_q H)^2}{4\partial_x Hk}$, we take this back to the HJB equation and get a PDE without sup:

$$\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H - \phi\nu^2 + \frac{(\partial_x HS - b\partial_x H - \partial_q H)^2}{4\partial_x Hk} = 0$$

The next step is to solve this PDE, we have the santz

$$H(t, x, S, q) = x + Sq + h(t, S, q)$$

with terminal $h(T, X, q) = -\alpha q^2$

Hence, we can know that $\partial_x H = 1$, $\partial_t H = \partial_t h(t, s, q)$, $\partial_S H = q + \partial_S h(t, s, q)$, $\partial_q H = S + \partial_q h(t, s, q)$.

We take these partial derivatives back and we can see that this PDE turns into

$$\partial_t h - q^2\phi + \frac{1}{4k}(b(q + \partial_S h(t, s, q)) + \partial_q h)^2 = 0$$

We note that this does not explicitly depend on S, so $\partial_S h = 0$, and this PDE becomes

$$\partial_t h - q^2\phi + \frac{1}{4k}(b(q + \partial_q h))^2 = 0$$

To solve this PDE, we seperate variable by letting

$$h = q^2 h_2$$

Taking this back to the PDE, we have ODE

$$0 = h_2' - \phi + \frac{1}{k}(h_2 + \frac{1}{2}b)^2$$

with terminal $h_2(T) = \alpha$

To solve this ODE, we let $h_2(t) = -\frac{1}{2}b + h_3(t)$, and we get

$$\frac{h_3'}{k\phi - h_3^2} = \frac{1}{k}$$

we separate variables and get

$$\frac{1}{k\phi - h_3^2} dh_3 = \frac{1}{k} dt$$

and integral both sides, we can get

$$\frac{1}{2(k\phi)^{1/2}} ln \frac{(k\phi)^{\frac{1}{2}} + h_3}{(k\phi)^{\frac{1}{2}} - h_3} = t + C$$

This time, we only integrate from t to T and rearrange terms, we can get

$$h_3(t) = \gamma \frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}}$$

where $\gamma = (k\phi)^{1/2}$, $\zeta = \frac{a - \frac{1}{2}b + \gamma}{a - \frac{1}{2}b - \gamma}$, $H = X + Sq + (-\frac{1}{2}b + h3)q^2$, and we can see that now

$$\nu* = \frac{S - bq - 2q(h_3 - \frac{1}{2}b) - S}{2k} = \frac{-qh_3}{k} = -(\frac{\phi}{k})^{1/2}\frac{1+\zeta e^{2\gamma(T-t)}}{1-\zeta e^{2\gamma(T-t)}}q$$

which is the same as the conlusion in 2(a)

We still need to find out the value of q*, when ν is optimal. We have

$$dQ_t^{v*} = -v_t dt$$

So,

$$Q_t^{v*} = q_0 exp(\int_0^t \frac{h_3(s)}{k} ds) = q_0 exp(log\frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}}) = q_0\frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}}$$

which is the same as our conclusion in 2(a)

ii

when $\phi \to 0$, $\gamma \to 0$, $\zeta \to 1$ and Q_t^{v*} has a $\frac{0}{0}$ form, so we can use L' hopital rule and we get

$$\lim_{\gamma \to 0} Q_t^{v*} = \frac{(T-t)e^{\gamma(T-t)} + (T-t)e^{-\gamma(T-t)}}{Te^{\gamma T} + e^{-\gamma T}} = \frac{T-t}{T}$$

and we can see from Q_t^{v*}, $v_t = \frac{1}{T}$.

When there is no punishment on holding inventory , the optimal trading strategy is just a TWAP strategy. Intuitively, when there is no urgent to liquid the position, we will just liquid on averge of time and this happens to be the optimal trading strategy when $\phi \to 0$.

Problem 2 (b)

i

The corresponding HJB equation to the problem is

$$\partial_t H + \mu\partial_S H + \frac{1}{2}\sigma^2\partial_{ss}H + \sup_\nu\{-\nu\partial_q H + (S - k\nu)\nu\} = 0$$

with the terminal condition

$$H(T, S, q) = q(S - \alpha q)$$

To find the optimal control, we maximize the $\sup$ part (by taking the derivative w.r.t. ν, for example, since it's a quadratic function of ν) and get

$$-\partial_q H + S - 2k\nu = 0 \implies \nu^* = \frac{S - \partial_q H}{2k}$$

ii

Substituting the optimal control back to the HJB equation gives

$$\partial_t H + \mu\partial_S H + \frac{1}{2}\sigma^2\partial_{ss}H + \frac{(S-\partial_q H)^2}{4k} = 0$$

To solve this PDE, we first consider the *ansatz* $H(t,S,q) = qS + h(t,S,q)$ and obtain

$$\partial_t h + \mu q + \mu\partial_S h + \frac{1}{2}\sigma^2\partial_{SS}h + \frac{(\partial_q h)^2}{4k} = 0$$

where the terminal condition is

$$h(T,S,q) = -\alpha q^2$$

Since neither the PDE nor the terminal condition explicitly rely on S, we further take the form $h(t,S,q) = h(t,q)$ (with a little abuse of notation) and simplify the PDE as

$$\partial_t h + \mu q + \frac{(\partial_q h)^2}{4k} = 0$$

We now guess the solution has the form

$$h(t,q) = \phi_0(t) + \phi_1(t)q + \phi_2(t)q^2$$

and thus the PDE becomes

$$(\phi_2' + \frac{1}{k}\phi_2^2)q^2 + (\phi_1' + \mu + \frac{1}{k}\phi_1\phi_2)q + (\phi_0' + \frac{1}{4k}\phi_1^2) = 0$$

with

$$\phi_2(T) = -\alpha, \quad \phi_0(T) = \phi_1(T) = 0$$

The PDE above holds for any t and q. Thus, we have

$$\begin{aligned} \phi_2' + \frac{1}{k}\phi_2^2 &= 0 \\ \phi_1' + \mu + \frac{1}{k}\phi_1\phi_2 &= 0 \\ \phi_0' + \frac{1}{4k}\phi_1^2 &= 0 \end{aligned}$$

We start with the first ODE. Using separation of variables, we have

$$\begin{aligned} \frac{d\phi_2(t)}{\phi_2(t)^2} &= -\frac{1}{k}dt \\ \implies \frac{1}{\phi_2(t)} &= \frac{t}{k} + C_2 \end{aligned}$$

With $\phi_2(T) = -\alpha$, we conclude that

$$\phi_2 = -\frac{k}{T - t + \frac{k}{\alpha}}$$

Now we solve the second equation. The equation now becomes

$$\phi_1' - \frac{k}{T-t+\frac{k}{\alpha}}\phi_1 + \mu = 0$$

Multiplying both sides with the integrating factor $e^{\int -\frac{k}{T-t+\frac{k}{\alpha}}} \sim T-t+\frac{k}{\alpha}$, we have

$$\begin{aligned}
\phi_1(t) \times (T-t+\frac{k}{\alpha}) &= \int -\mu(T-t+\frac{k}{\alpha})dt \\
&= \int \mu(T-t+\frac{k}{\alpha})d(T-t) \\
&= \frac{1}{2}\mu(T-t)^2 + \mu\frac{k}{\alpha}(T-t) + C_1 \\
&= \frac{1}{2}\mu(T-t)(T-t+\frac{2k}{\alpha}) + C_1
\end{aligned}$$

Again, with $\phi_1(T) = 0$, we have $C_1 = 0$, and

$$\phi_1(t) = \frac{\frac{1}{2}\mu(T-t)(T-t+\frac{2k}{\alpha})}{T-t+\frac{k}{\alpha}}$$

Finally, we have

$$\begin{aligned}
\nu^* &= \frac{S-\partial_q H}{2k} \\
&= -\frac{\partial_q h}{2k} \\
&= -\frac{\phi_1(t)+2\phi_2(t)q}{2k} \\
&= \frac{q}{T-t+\frac{k}{\alpha}} - \frac{1}{4k}\mu(T-t)\frac{T-t+\frac{2k}{\alpha}}{T-t+\frac{k}{\alpha}}
\end{aligned}$$

where $q = Q_t^{\nu^*}$.

iii

Denote $q = q_\alpha(t)$ parametrized with α. From the optimal control above, we have

$$-q'(t) = \frac{q}{T-t+\frac{k}{\alpha}} - \frac{1}{4k}\mu(T-t)\frac{T-t+\frac{2k}{\alpha}}{T-t+\frac{k}{\alpha}}$$

We want to find $q_\infty(t)$ with $\alpha \to \infty$. In other words, We solve

$$q_\infty(t) = \lim_{\alpha\to\infty} -\int_0^t \left(\frac{q_\infty(u)}{T-u+\frac{k}{\alpha}} - \frac{1}{4k}\mu(T-u)\frac{T-u+\frac{2k}{\alpha}}{T-u+\frac{k}{\alpha}}\right)du$$

We interchange the limit and integral without proof, take the derivative w.r.t. t on both sides, and obtain an ODE

$$q_\infty'(t) = \frac{q_\infty(t)}{T-t} - \frac{1}{4k}\mu(T-t)$$

To solve this ODE, we multiply the integrating factor $e^{\int -\frac{1}{T-t}} \sim -\frac{1}{T-t}$ on both sides and obtain

$$\left(\frac{q_\infty(t)}{T-t}\right)' = \frac{1}{4k}\mu$$

Integrating and using $q_\alpha(0) = q_0$ gives

$$\frac{q_\infty(t)}{T-t} = -\frac{1}{4k}\mu(T-t) + \frac{\mu T}{4k} + \frac{q_0}{T} \implies q_\infty(t) = (T-t)\left(\frac{\mu t}{4k} + \frac{q_0}{T}\right)$$

Remember that this $q_\infty(t)$ is exactly $Q_t^{\nu^*}$. This concludes the proof.

Problem 3 (a)

> Describe four of the most commonly used high-frequency trading strategies.

1. Latency Arbitrage

Summary: To trade the same asset across markets with the existence of differences between prices.

How does it work? The same asset may have different prices across different markets, leaving space for arbitrage. The HFT traders can simply buy in assets at a lower price in one exchange and sell it at a higher price in another, earning the spread, if they observe the existence of such opportunities.

Is it effective? Like rebate arbitrage, traders earn little profit per share with latency arbitrage, but this can easily accumulate to a significant number. It also incurs very little risk as the trader is able to complete the transaction in a very short time.

Does it have alpha? No, the HFT traders do not bet on the movement of the price.

What does SEC think? Latency arbitrage is good to market efficiency as in theory each asset (especially stocks which do not have storage costs, etc.) should have a unanimous price across different markets. Via latency arbitrage, HFT traders eliminate such differences and lubricate the price determination.

Comments: With regard to latency arbitrage, this is purely a game of technology and computer science. Very little finance is involved as traders do not need to understand how prices are formed. Rather, they design fast programs and reduce the latency to make sure they get notified with such opportunities and take action faster than anybody else.

References:

Be Thankful for High-Frequency Trading

Getting Up to Speed on High-Frequency Trading

High Frequency Trading

2. Cross Asset/Market/ETF Arbitrage

Summary: To earn the profit between the spread between two matching portfolios.

How does it work? Unlike as predicted by Arrow-Debreu, there may exist two different portfolios on the market that should have almost identical prices. For example, the value of an ETF index fund should equal that of the portfolio constituting the index. If there is a difference between the two portfolios, traders then can buy in the undervalued portfolio and at the same time sell the overpriced.

In reality, this exact relationship seldom occurs. However, the HFT traders may discover two portfolios that are nearly perfectly matched. For example, two stocks may be assumed highly correlated. If this is true, the appreciation of one stock will result in another stock appreciating with a high probability. The HFT traders can then monitor the two stocks and take action when one of them starts to move away from its current price.

Is it effective? As this strategy sounds unreliable, the traders may have thousands or even more pairs of portfolios to track. Statistically, the HFT traders are still able to make a profit even if some of the executions result in losses. Also, just like many other strategies, HFT traders only hold the assets in a very short time (at most minutes), thus bearing limited risk.

Does it have alpha? Yes, the traders do need the ability to predict the asset prices to some extent.

What does SEC think? Such arbitrage strategies contribute to market efficiency as traders seek for wrongly priced assets. As a result, such misprices are gone and true prices are discovered.

Comments: Some people tend to think latency arbitrage and cross market arbitrage are the same. We suggest that these two strategies are fundamentally different in that the former one involves purely technological investment and does not rely on quantitative results. The latter one, however, allows higher latency and bets on the probability that the existence of difference, once detected, will converge eventually.

References:

Statistical Arbitrage: Defined & Strategies

Getting Up to Speed on High-Frequency Trading

High Frequency Trading

3. Order Anticipation Strategies

Summary: To formulate an anticipation of price movement by predicting the potential orders and make a profit with this foreknowledge.

How does it work? First, HFT traders need the ability to know the potential orders before anyone else on the market. Such orders are "potential" rather than listed on the market because the brokers have not posted these orders or have posted them in the dark pool. Since posting a large order can leak information that stimulates the market to act upon it, making the price move against the interest of the broker, the broker usually wants to split a large order into many smaller ones, which is done with a program. The logic of the program is unknown to others, but can possibly be guessed by playing with it. This allows HFT traders the ability to tell the existence of large orders and design a strategy to make a profit.

One possible approach to detecting such orders is by "pinging", a terminology comes from computer networks. The traders send small orders across many stocks and observe whether any of them are filled. They then send more small orders to these stocks until they have enough confidence that there is a demand.

Knowing the existence of such a demand, the HFT traders can then prepare to exploit this opportunity. With the best conditions (e.g. co-location, fast computers, etc.), these traders can act faster than the broker in placing orders. They then prepare their positions and function as the counterparty of the broker by inducing a price against the broker. In this way, the traders can earn a profit.

Is it effective? It might be surprising that about two thirds of orders placed are filled in the scenario (see this post). With careful design, such strategies can be extremely successful and produce a significant profit.

Does it have alpha? Yes, it does. The trader effectively predicts the movement of the price by guessing a large demand that will be fulfilled even if the price moves a little against its issuer.

What does SEC think? The primary concern of SEC is whether this action provides liquidity to the market and whether this is fair. With order anticipation strategies, the HFT traders do not offer extra liquidity. Instead, they make brave guesses and only act as the bridge, preying on any spread that exists between the market and the broker who demonstrates the demand. It is also fair because the traders do take the risk. Thus, unlike front-running and spoofing that are significantly detrimental to fairness, the SEC does not explicitly and totally prohibit order anticipation.

Comments: There are also many strategies that have been banned by SEC and FINRA. Such strategies include spoofing and front-running. They seem to be a special case of order anticipation but are fundamentally different because they induce market anticipation rather than detect it, nowadays regarded as an act of market manipulation.

References:

Concept Release on Equity Market Structure

Market Data Patterns, Order Anticipation, and An Example Trading Strategy

4. Rebate Arbitrage

Summary: To earn the rebate as a reward by the exchange to place limit orders as providing liquidity to the market.

How does it work? Limit orders prevail over market orders in that the former one is usually regarded as a source of liquidity, while the latter simply consumes liquidity. To encourage limit orders, exchanges usually offers a fee rebate to them. Traders then can operate on market orders and reverse operate to earn the rebate.

Is it effective? Even if the rebate on each share is small, they can easily accumulate to significant amounts for HFT traders. This strategy also incurs very little risk and is very ideal to HFT traders.

Does it have alpha? No, the HFT traders do not bet on the movement of the price.

What does SEC think? The SEC does not complain because this does not harm the liquidity of the market. This also facilitates price discovery by reducing the effect of commission fees.

Comments: Theoretically, all fees are reflected on the price. Thus, those who take a lower fee are able to accept a less favorable price yet still make a profit. This is analogous to taxing which gives rise to a difference between buyer's payment and seller's gain. In fact, HFT practitioners can indeed make money simply because they are favored, or using jargon, have the "rent".

References:

Limit Order

Complying with Arbitrage Requirements:
A Guide for Issuers of Tax-Exempt Bonds

Arbitrage and Rebate

Problem 3 (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.

First of all, we are going to answer this question in general. Do these HFT traders make a lot of money? The answer is YES! We can take a look at the salary of employers of these HFT trading companies. Using data from Glassdoor, we can see that a quantitative researcher focus on HFT at Citadel, a famous market-maker can make at least 230k base salary and an additional pay of 70k on average every year. We can also find a open HFT poisition at Citadel with base salary 250k, sign-in bonus 100k, and a possible annul bonus of 150k. This is not the only company giving this high salary, other HFT companies are also giving extremely high salary to their employees. Hence, of course, we know HFT companies earn a lot, otherwise they cannot afford to pay for these expensive and clever brains.

Next, we are going to have a rough estimate of how much do these HFT companies every year. Accroding to Bloomberg, Citadel has a revenue of around 6.7 billion and a EBIT of 4.1 billion last year. We assume that HFT trading companies earn nearly the same with a same trading volume. There is a common belief in the market that Citadel accoounts for 10% of US equity market trading volume and all HFT companies accounts for 90% of the trading volume. So, the profit of HFT industry should be around 9*4.1=36.9 billion dollars every year.

Then, we are going to have an estimate of HFT leverage. Luckily, Virtu financial, another famous HFT company is a listed company and we will simple use their public data. Do they use leverage? Yes! Their equity worths about 10 billion and they have a debt of about 80 billion. So, their debt on equity ratio is around 8. Two necessary assumoptions here are HFT only trade equities, not derivatives, because trading derivatives can highly increase their leverage, and all HFT companies have similar debt to equity ratios.

References:

[Citadel Salaries]

[An anatomy of a Citadel (High Frequency Trading) job offer]

[Morning Coffee: Citadel Securities' enormous profits per employee. Bad luck to bankers caught up in Brexit]

Problem 3 (c)

> Does high frequency trading impose risks of systemic nature?

The HFT is usually believed to have the following key characteristics: (1) Automated processes for trading, (2) high speed in the submission of orders and in the process of incoming information, and (3) submission of high numbers of orders and/or quotes. Although HFT with these characteristics plays a key role in potentially increasing the market liquidity and contributing to the reduction of bid-ask spread since 1980s when technology started to get into stock markets, we agree that it also imposes systematic risk to the market as discussed in the referenced paper.

1. Adverse selection in orders and market-making

As learned from Market Microstructure class, the adverse selection in orders refers to a situation where market participation is affected by asymmetric information with HFT acting as "informed" traders. In other words, adverse selection in HFT can be interpreted as a faster reaction to new pieces of information. Thus, this imposes an issue of particular relevance when a market operates with both HFT market-makers and non-HFT market makers (MM) because the latter may be trading with a counterparty that is much better informed and faster than they are. These non-HFT market makers would be forced to charge a higher premium for the information risk they are facing. As a result, the information embedded in prices decreases substantially and slower non-HFT market makes may be crowded out from the market. The ultimate consequence of this could be a dysfunctional and self-induced process of price formation. A good example may be the Flash Crash May 6 2010.

2. Market Power and Barriers to entry

I recall from the Market Microstructure class, the HFT can potentially game the market in the way that HFT market marker may potentially create a large number of sell/buy orders to manipulate the price of security and eventually cancel those orders after executing with an ideal price themselves. The HFT market maker would enjoy several advantages from their privileged position while the non-HFT or slower group does not. Due to the complicated algorithm and technology HFT adopts, it makes non-HFT market makers very hard to catch up and a barrier to entry. In principle, financial markets regulation should ensure all the market participant to have equal access to information and price.

3. Correlation

Intuitively, the optimization of algorithms in HFT would remove the human component in trading leading toward more homogenous strategies and reactions to new events. However, if a shock simultaneously affecting several HFT may have negative consequences for them in terms of solvency and ultimately leading to widespread failures of HFT imposing an undesired defect of systemic risk in terms of interconnectedness and contagion. Also, Jiang et al. (2014) find that HFT increase volatility in the US Treasury market during, shortly before and after the announcement of macroeconomic news.

In sum, we do agree with the findings from the paper and think HFT does potentially impose a systemic risk to the market.

Reference:

[“High-Frequency Trading and Systemic Risk: A Structured Review of Findings and Policies”]

Problem 3 (d)

> Propose an intraday “alpha” trading strategy. Describe what methodology and data you would use to research your idea. Provide a “guestimate” of its performance.

We would like to propose an intraday Dynamic Pairs Trading using correlation and cointegration. It’s based on market-neutral statistical arbitrage strategy using a two-stage correlation and cointegration approach.

1. Idea

The pairs trading strategy uses trading signals based on the regression residual e and were modeled as a mean-reverting process.

2. Data

In order to have more pairs with very high correlation, we would select stocks in a specific industry. Intuitively, we would choose stocks from the same sector since they may be more likely to be close substitutes with each other. If we choose N stocks from a sector, then we would have

$$\frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}$$

number of pairs. We would choose to use millisecond or second data for a certain period of time that market is not very volatile. And may apply the strategy to a period of time when market is indeed volatile to compare the performance.

3. Methodology

(1) Correlation Approach

We would like to compute the correlation between each of the two stocks as below. The higher the ρ, the more positive the association of stock A and B is.

$$\rho = \frac{\sum_{i=1}^{N}(A_i - \bar{A})(B_i - \bar{B})}{[\sum_{i=1}^{N}(A_i - \bar{A})^2 \sum_{i=1}(B_i - \bar{B})^2]^{1/2}}$$

However, the correlation coefficients don't necessarily imply mean-reversion between the prices of two stock pairs. Thus a cointegration approach would be used as a second step of the selection process for the pairs.

(2) Cointegration Approach

Assume that x_t and y_t are two time series that are non-stationary. If there is exists a parameter γ such that

$$z_t = y_t - \gamma x_t$$

then it's a stationary process, and thus x_t and y_t would be cointegrated. In our case, the following would be a stationary process if there exists a parameter γ

$$P_t^A - \gamma P_t^B = \mu + \epsilon_t$$

where μ is a mean of the cointegration model and ϵ_t is a stationary, mean-reverting process and is referred to as a cointegration residual. The γ is known as a cointegration coefficient.

If ϵ is positive within a given confidence interval, it would be a signal that stock A is relatively overpriced and B is relatively underpriced. In this case, we would short A and long B at the same time. If ϵ is negative, we would long A and short B instead.

4. Implementation of strategy

This pairs trading strategy uses trading signals based on the regression residual ϵ and were modeled as a mean-reverting process. The first step is to identify stock pairs from the sector we choose and then pick the pair stock that have a correlation greater than a threshold, say 0.90. The second step is to check if the pairs can pass the correlation test. For example, the null hypothesis is

$$H_0 : \gamma = 0$$

If it's rejected, then it implies that the residual ϵ is stationary and the pair passes the cointegration test. Next, we may ran the pairs based on cointegration test value with higher rank for those pairs having smaller cointegration test value.

Last step would be to determine trading rules. In our case, if ϵ_t cross over and down the positive σ standard deviation above the mean or cross down and over the negative σ standard deviation below the mean, then we consider to initiate a trade. Again, if the residual is positive, we short Stock B and long stock A at the same time, vice versa.

5. Performance measurement

We may compute the Shape ratio for the period of our training dataset as our return benchmark. We run back test with our strategy above and compare to the benchmark to determine how it's performed. Also, we can test how different performance corresponding to the threshold parameters we choose (e.g., correlation coefficient, number of standard deviation used to construct the confidence interval, etc). In addition, we would like to use the sector from different time periods to stress test the performance of the strategy in different scenario in order to comprehensively measure the performance.

Reference:

[Intraday Dynamic Pairs Trading using Correlation and Cointegration Approach]