

QF302 Homework Assignment #1

Due Date: March 13 (Wednesday). Instructions of the assignment: Your answers need to be typed using any document editing software (e.g. Word), but your final submission has to be saved in one single PDF file. Any supporting programs (Python or R) need to be submitted along with the PDF file with proper references.

Problem 1. [50 points] You are required to answer the following questions using the SHIFT system with Python programming. The bid price B_t represents the highest priced buy order that's currently available in the market. The ask price A_t is the lowest priced sell order that's currently available or the lowest price that someone is willing to go short or sell at. The bid/ask spread S_t is the difference in price between the bid and ask prices. However, in practice people use different ways to estimate bid/ask spread and use it as a proxy for market liquidity. In this problem, you will need to use the following two measures to measure the bid/ask spread of Netflix (tiker symbol: NFLX) stock price using the SHIFT system with time interval or period $\Delta t = 5$ seconds:

The **quoted spread** between ask A_t and bid B_t that is averaged over T periods equals:

$$S^Q = \frac{1}{T} \sum_{t=1}^T (A_t - B_t) \quad (1)$$

The **average spread** in terms of the asset fundamental price P_t^* is defined as:

$$S^E = \frac{1}{T} \sum_{t=1}^T 2q_t(P_t - M_t) \quad \text{where} \quad M_t = \frac{1}{2}(A_t + B_t) \quad (2)$$

and q_t is 1 for buy orders and -1 for sell orders.

The **illiquidity** (Amihud 2002) is based on the price impact caused by

trading volume, is used:

$$ILLIQ = \frac{1}{T} \sum_{k=1}^T |r_k|/V_k \quad (3)$$

where r_k and V_k are the return and trading volume for k th period. (Note: V_k is the total stock volume traded during the k th period.)

- (a) Please present a table of summary statistics of your B_t and A_t including min, median, mean, max, std. [10 points]
- (b) Please use NFLX data from SHIFT and calculate S^Q , S^E , and $ILLIQ$ for the same period of time. In this calculation, please use $\Delta t = 5$ seconds for each period and $T = 120$ for the number of periods (or run the program for 10 minutes). Please also plot all three measures in one single graph. (Note: You may pick any starting point, and use two y axes - one for bid-ask spread and another for liquidity for the plot.) [15 points]
- (c) Please calculate Pearson correlation of between S^Q and S^E . Repeat (b) step 10 times and calculate the average Pearson correlation between S^Q and S^E . (Note: Do not repeat plotting.) [5 points]
- (d) Please calculate Pearson correlation of between S^Q and $ILLIQ$. Repeat (b) step 10 times and calculate the average Pearson correlation between S^Q and $ILLIQ$. (Note: Do not repeat plotting.) [5 points]
- (e) Please provide your observations on the relationship between S^Q and S^E , and the relationship between S^Q and $ILLIQ$ from the results from (c) and (d). [15 points]

Problem 2. [20 points] Realized volatility is normally calculated based on n data points using $\sigma_t = \left(\frac{1}{n} \sum_{i=t-n+1}^t (r_i - \bar{r})^2\right)^{1/2}$. Consider the daily stock price for Netflix (NFLX) from October 22, 2013 to October 22, 2018 in the file NFLX_2013–2018.csv.

Use the following volatility forecasting method to calculate “expected volatility” or sample mean volatility using the given Netflix (NFLX) stock prices. Please use $n = 22$ days to calculate σ_t and $\beta = 2/(n + 1)$ for the following volatility forecasting methods. Note: please use logarithmic return.

(a) The *Random Walk* forecast: $\hat{\sigma}_t = \sigma_{t-1}$ [5 points]

(b) *Exponential smoothing average* (often called exponential moving average or EMA):

$$\hat{\sigma}_t = (1 - \beta)\sigma_{t-1} + \beta\hat{\sigma}_{t-1}, 0 < \beta < 1 \quad (4)$$

[5 points]

(c) *Exponentially weighted moving average* (EWMA), which is a truncated version of EMA:

$$\hat{\sigma}_t = \sum_{i=1}^n \beta^i \sigma_{t-i} / \sum_{i=1}^n \beta^i, 0 < \beta < 1 \quad (5)$$

[5 points]

(d) Compare and contrast the differences of these three estimation methods.

Write your analysis and recommendation. [5 points]

Problem 3. [30 points] In Roll’s model, we define efficient price by m_t , and we assume $m_t = m_{t-1} + u_t$, where u_t are i.i.d. zero-mean random variables ($u_t \sim N(0, \sigma_u^2)$). The bid-ask spread is $a_t - b_t = 2c$, a constant. At time t , there is a trade at transaction price p_t , which may be expressed as: $p_t = m_t + q_t c$ where q_t is a trade direction indicator set to $+1$ if the customer is buying and -1 if the customer is selling. We also assume that buys and sells are equally likely, serially independent (a buy this period does not change the probability of a buy next period), and that agents buy or sell independently of u_t (a customer buy or sell is unrelated to the evolution of

m_t). By definition we have $\Delta p_t = (q_t - q_{t-1})c + u_t$, and we can write the variance and autocovariance as:

$$\begin{aligned}\gamma_0 &\equiv \text{Var}(\Delta p_t) = 2c^2 + \sigma_u^2. \\ \gamma_1 &\equiv \text{Cov}(\Delta p_{t-1}, \Delta p_t) = -c^2.\end{aligned}$$

For this assignment, you need to do the following (please use logarithmic return):

- (a) Please derive the value $\gamma_\ell, \ell \geq 2$? [10 points]
- (b) Use the Roll model to estimate the bid-ask spread and fundamental volatility of Netflix stock during the period from October 22, 2013 to October 22, 2018. Please use the same dataset for the last problem (NFLX_2013–2018.csv). [20 points]

Homework Honor Policy: You are allowed to discuss the problems between yourselves, but once you begin writing up your solution, you must do so independently, and cannot show one another any parts of your written solutions. The HW is to be pledged (that it adheres to this).