

# Optimal Position for Beating TWAP with Prediction in T + 0 and T + 1Markets

# Xinyi Zhang<sup>1</sup>, Ming Ma<sup>2</sup>\*

<sup>1</sup>United World College Changshu, 215500, China

<sup>2</sup>\*Higgs Research Lab, Hangzhou Higgs Asset Management Co., Hangzhou, 310051, China

Keywords: optimal position, recursive mean-variance, dynamic programming, position lines

**ABSTRACT.** In this paper, we consider a trading algorithm for the investor with some prediction of the stock price in the market in order to gain excess returns, achieved by controlling the position of the stock held according to the prediction. We aim to find an optimal feedback control which maximizes the recursive mean-variance of the excess return. In this paper, we define the excess return to be the summation of every individual mean-variance for every operation. By mathematical induction, we generate an explicit expression of the different position lines, in terms of straight lines, corresponding to different signals in T + 0 market. Also, we prove that time-weighted average price (TWAP) strategy is the optimal strategy in T + 1 market when no information is available and present the numerical simulation for the kinked position lines in T + 1 market when prediction is possible. Particularly, we discover that the optimal feedback control, regardless of the type of market, is determined by all the straight or kinked position lines corresponding to all the different signals.

# **1. INTRODUCTION**

With the arrival of Digital age, the landscape of finance has been shaped greatly by the extensive use of computer programming. Traders have been gradually freed from checking the stocks constantly and repeatedly in order to carry out operations with the emergence of trading algorithms. Trading algorithms are ways to execute orders from automatic instructions determined by pre-set variables like time, volume, price and so on. Popular trading algorithms include Volume Weighted Average Price (VWAP), Time Weighted Average Price (TWAP), Implementation Shortfall (IS). VWAP trading strategy refers to the trading of stocks with the benchmark of the volume-weighted average price over a certain period of time; similarly, TWAP refers to the stock trading with the time-weighted average price over a certain time period as the benchmark. On the other hand, IS trading strategy focuses on minimizing the difference between the decision price, including both the close price and the arrival price, and the final execution price which includes terms of taxes, commission fees, and etc. Usually, when an investor employs TWAP as his or her trading algorithm, the position will be opened or closed uniformly over time; however, in our research paper, we investigate the trading strategies that could be employed for an investor with some ability to predict the market price for excess returns as an improvement on the basis of TWAP.

To determine an optimal position-opening strategy, from the mathematical view, we consider it as an optimization problem in order to maximize the utility of the stocks. Traditionally, the mean-variance analysis proposed by Markowitz (1952) is widely employed due to its prominent simplicity and intuitive appeal.<sup>[5]</sup> This analysis is typically useful under a single-period framework or for a myopic investor who aims to maximize the next period objective under a multi-period framework (e.g. Jagannathan and Ma, 2003; Bansal, Dahlquist and Harvey, 2004); however, the results generated by traditional mean-variance analysis may not provide the investor with the optimal dynamic strategy; that is to say, the optimal strategy over the period of total trading time may not be compatible with the optimal strategy over the one specific time segment of the total time<sup>[1][4]</sup>. Therefore, to eliminate this time-inconsistency, we consider the dynamically optimal mean-variance policy brought up by Basak and Chabakauri (2009) where dynamic programming could be applied for the convenience of numerical simulation and the dynamic time-consistency is preserved.<sup>[2]</sup> In our paper, we first build

the discrete-time model for the change of the stock price. Because we only consider the change of the price within one day, we could ignore the compound interest associated and use the simple drifted random walk model as a description of the stock price. Furthermore, we set the signal for prediction to conform to the Bernoulli distribution, and with the minimal time interval for operation  $\delta t$  being small enough, the Bernoulli distribution will approach to the Poisson distribution. The size of the signal could also be described by the Bernoulli distribution, with the probability of a price rise to be p and the probability of a price drop to be 1 - p. With this model, we aim to find a feedback control that maximizes the utility of the excess return, which is the difference between the realized average price of the investor and the time-weighted average price.

In a T + 0 market, where stocks could be traded freely regardless of how long the investor has hold the stocks, the optimization problem is relatively easy. Because there are no restrictions on the feedback control, we generate the explicit expression of the value function for each operation optimal feedback control. Moreover, we represent the optimal feedback control as different position lines on the graph and associate every signal to a different position line. In particular, when we have only three types of signals, three different positions line is drawn on the graph respectively. Even though the mathematical expression of the optimal feedback control is complicated, the philosophy is simple: every time a signal is observed by the investor, the next position should be adjusted to the line corresponding to the signal. However, in a T + 1 market, such as Chinese market, where investors are forced to hold the stocks for at least one day before any trading, there are many restrictions in the admissible feedback control. In fact, these restrictions make it challenging for us to obtain an explicit expression in T + 1 market. Still, we prove that the optimal feedback control is TWAP when no information is available and no prediction could be made in T + 1 market and have worked out the numerical solutions to the optimal feedback control through computer simulation. Particularly, we discover that the optimal feedback control also takes the form of different position lines. On the contrary to the T + 0 market, the position line of no signal coincides with the line of preferred signal, depending on whether the investor aims to open or close the position. Only the undesired signal would result in the change of the position.

As a remainder for the readers, this paper is organized as the following. In Section 2, we build the model of the stock price, strategy employed, and the prediction ability of the investor. In Section 3, we represent the return and risk of the excess return in terms of mean and variance and set up the recursive mean-variance strategy, to which dynamic programming could be applied. Furthermore, in section 4, we formulate the explicit expression for T + 1 market with no prediction and the one for T + 0 market by mathematical induction, while in sections 5, the graphs from our computer simulation are presented visually and explained to a detailed extent to illustrate the optimal feedback control in the T + 1 market. Finally, we conclude our paper in Section 6.

## 2. Model Setting

### 2.1.Stock price

In this paper, we first define the following variables in order to build a mathematical model to describe the stock price in the T + 1 stock market:

- 1. Total time of the trading hours in the stock market :  $\overline{T}$ ;
- 2. the daily volatility of the stock :  $\bar{\sigma}$ ;
- 3. the minimal time interval required between operations:  $\Delta t$ ;
- 4. the elapsed trading time:  $\bar{t}$ .

Hence, with the  $\overline{T}$  given, the total number of operations that could be taken within one day can be calculated as following:  $T \triangleq \overline{T}/\Delta t$ . In this case, we could simply suppose the T to be an integer because the number of operations could only be a natural number.

Consequently, we could define the number of operations that could be taken for the elapsed trading time as  $t = \bar{t}/\Delta t$ , and t takes value of any integers between 0 and T including 0 and T with the same reason explained above.



In addition, because the volatility of the stock is a statistical measure of the dispersion of returns during a certain period of time for stocks, security or market index, often measured by the standard deviation, with the given daily volatility, the volatility of one operation can is  $\sigma \triangleq \bar{\sigma}/\sqrt{T} = \bar{\sigma}\sqrt{\Delta t/\bar{T}}$ . Only with this  $\sigma$  above can the sum of the volatility of all operations in one day generate the mathematically correct daily volatility of the day.

Finally, we could now denote  $P_t$  as the stock price at time  $\bar{t}$ . Usually, in a time-continuous model of the stock price, the stock price is consist of three elements: the smooth time-varying volatility of the stock, which is also called as the drift of the stock, the micro-structure noise, and jump. However, we consider the case to be under the time-discrete condition because the operation of the stock market could not happen instantaneously. Rather, it takes time for the operations to be carried out. In this case, with a set  $\Delta t$ , there will be no difference between the drift and the jump. Hence, we would combine them in the following recurrence formula:

$$P_t = P_{t-1} + \mu_{t-1}\Delta t + \sigma Z_t, \tag{1}$$

where,  $\mu_{t-1}$  is the drift between t-1 and t which is a variable that can be considered to be predictable in our model, and  $Z_t$  is the noise between t-1 and t which takes the form of a standard normal distribution. Usually, we consider the noise at different time to be random points on the standard normal distribution and therefore, the terms of  $\{Z_t\}_{t=0}^{T-1}$  to be independent.

Accordingly, it is easy to give an explicit expression of  $P_t$ :

$$P_t = P_0 + (\sum_{s=1}^{t-1} \mu_s) \Delta t + \sigma \sum_{s=1}^t Z_s, \qquad (2)$$

which is a summation of all predictions and all noises.

#### 2.2.Strategy

In this paper, we consider generating a position with an average price lower than that of the timeweighted average price (TWAP). Here, we would like to define the purchasing strategy of one stock within one day by the following series:

$$\{V_t\}_{t=0}^T$$

where  $V_t$  indicates the volume of stocks purchased already at time t, and we assume that the investor has a specific investment target, that is, a certain volume of stocks to purchase within one day. Hence we have our boundary conditions where

$$V_0 = 0, V_T = 1.$$

In a T + 0 market, our purchasing strategy is already well defined because T + 0 market permits free trade of stocks without any restrictions on time. However, in a T + 1 market where investors are required to hold the stocks for at least one day before trading, our strategy series must increase monotonically since we are considering the operations within one day. Therefore, we have

$$0 \le V_{n-1} \le V_n \le 1, 0 \le n \le N$$

## 2.3.Prediction

Since we have already built the mathematical model of the stock price and the strategy, we now need to determine the prediction of the stock price from the view of an investor. We assume the investor to be able to predict the drift of the stock price,  $\mu_t$ , one step ahead based on some indications; for example, at t - 1, the investor knows the drift between t - 1 and t, including the one at t. Also, the plus-minus sign of  $\mu_t$  indicates the increasing or decreasing tendency of the stock price and the size of the  $\mu_t$  reveals the magnitude of the price change predicted.

Now, we define the  $\lambda$  to be the probability of a non-trivial indication within time  $\Delta t$ , that is,  $P(\mu_n \neq 0) = \lambda$ . Consequently, we can deduce  $P(\mu_n = 0) = 1 - \lambda$ . When  $\mu$  is not zero, we suppose it is a binary distribution:

$$P(\mu_n = \varepsilon | \mu_n \neq 0) = p, P(\mu_n = -\varepsilon | \mu_n \neq 0) = (1 - p),$$



where p is the probability of a positive prediction, and  $\varepsilon$  is the size of a prediction. Therefore, the marginal distribution  $\mu_{t+1}$  is a trinomial distribution:

$$P(\mu_n = \varepsilon) = p\lambda$$
,  $P(\mu_n = -\varepsilon) = (1 - p)\lambda$ ,  $P(\mu_n = 0) = 1 - \lambda$ .

### **3.Return and Risk**

#### **3.1.Excesses return**

We want to attain a position with lower average price than Time Weighted Average Price (TWAP) and determine a strategy to realize this position.

By definition, due to our definition of the strategy, the Weighted Realized Price at time t is

$$WRP_t = P_t(V_{t+1} - V_t).$$

Summing up, we get the Total Weighted Realized Price:

$$TWRP = \sum_{k=0}^{T-1} P_k (V_{k+1} - V_k).$$

In particular, when  $V_t = \frac{t}{T}$ , the Total Weighted Realized Price represents the TWAP:

$$TWAP = \sum_{k=0}^{T-1} \frac{P_k}{T}$$

Hence, the Excess Return can be calculated as following:

E

$$R \triangleq \sum_{k=0}^{T-1} P_k (V_{k+1} - V_k - \frac{1}{T}).(3)$$

By Abel transformation, summing it by parts, we can rewrite the Excess Return as the following:  $ER = \sum_{k=1}^{T-1} \left(\frac{k}{T} - V_k\right) (P_k - P_{k-1}). \quad (4)$ Noticing  $V_0 = 0$  and  $V_T = 1$ , we have the calculation as following:  $\Gamma^{T-1}$ 

$$ER = P_{T-1} \left[ \sum_{n=0}^{T-1} (V_{n+1} - V_n - \frac{1}{T}) \right] \\ + \sum_{k=0}^{T-2} (P_k - P_{k+1}) \left[ \sum_{n=0}^{k} (V_{n+1} - V_n - \frac{1}{T}) \right] \\ = P_{T-1} (V_T - V_0 - \frac{T}{T}) \\ + \sum_{k=0}^{T-2} (P_k - P_{k+1}) (V_{k+1} - V_0 - \frac{k+1}{T}) \\ = \sum_{k=0}^{T-2} (P_k - P_{k+1}) (V_{k+1} - \frac{k+1}{T}) \\ = \sum_{k=1}^{T-1} (\frac{k}{T} - V_k) (P_k - P_{k-1}) \\ = \sum_{k=1}^{T-1} (\frac{k}{T} - V_k) (\mu_{k-1} \Delta t + \sigma Z_k),$$

where we denote the  $(\frac{k}{T} - V_k)(\mu_{k-1}\Delta t + \sigma Z_k)$  as  $ER_k$ , the k th term of the summation, representing the excess return gained on the *k*th operation.

By Abel transformation, we have successfully changed not only the format of the Excess Return, but more importantly, our understanding to the Excess Return as well. Before, it could only be viewed



as a whole on the time scale; now, it could be understood as the summation of excess return gained from different operations at separate time intervals. This is crucial to the later employment of dynamic programming, where we only concern the mean and variance of the excess return for one operation.

### 3.2.Mean

Now we could calculate the expectation of the Excess Return separately in terms of one operation. The Excess Return gained on the kth operation is

$$\mathbb{E}[ER_k] = \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(P_k - P_{k-1})\right]$$
$$= \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(\mu_{k-1}\Delta t + \sigma Z_k)\right]$$
$$= \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(\mu_{k-1}\Delta t)\right] + \mathbb{E}\left[\sigma Z_k(\frac{k}{N} - V_k)\right].$$

Since the investor could only predict the drift in this case, the strategy of the investor must be completely independent from the noise. Accordingly, we obtain

$$\mathbb{E}[ER_k] = \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(\mu_{k-1}\Delta t)\right] + \mathbb{E}\left(\frac{k}{N} - V_k\right)\mathbb{E}[\sigma Z_k].$$

Because  $Z_k$  takes the form of a standard normal distribution, the expectation of it equals to zero. Hence,

$$\mathbb{E}[ER_k] = \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(\mu_{k-1}\Delta t)\right].$$

Due to the additivity of the expectation, the expectation of the Excess Return as a whole can simply be expressed as the summation of all the expectation of the single operation, that is

$$\mathbb{E}[ER] = \sum_{k=1}^{T-1} \mathbb{E}\left[\left(\frac{k}{N} - V_k\right)(\mu_{k-1}\Delta t)\right].$$

#### 3.3.Variance

Besides the expectation of the Excess Return, we also want to determine the variance of the Excess Return because it measures the significance of the risk associated with one specific strategy.

$$\operatorname{Var}[ER_{k}] = \operatorname{Var}\left[\left(\frac{k}{T} - V_{k}\right)(\mu_{k}\Delta k + \sigma Z_{k+1})|V_{k}, \mu_{k}\right]$$
$$= \operatorname{Var}\left[\left(\frac{k}{T} - V_{k}\right)(\mu_{k}\Delta k) + \sigma Z_{k+1}\left(\frac{k}{T} - V_{k}\right)\right].$$

Since all the variables in  $(\frac{k}{T} - V_k)(\mu_k \Delta k)$  are well known at k, this term could be seen as a constant when calculating the variance. Thus,

$$\operatorname{Var}[ER_k] = \operatorname{Var}\left[\sigma Z_{k+1}(\frac{k}{T} - V_t)\right] = \sigma^2 (\frac{k}{T} - V_t)^2.$$

### **3.4. Recursive Mean-Variance**

In this paper, the optimal strategy and the optimal result regarding both mean and variance would be determined by dynamic programming. Consequently, we first define the utility function  $U(V_{t+1}; V_t, \mu_t)$  and the value function  $u(t, V_t, \mu_t)$  as follows:

$$U(V_{t+1}; V_t, \mu_t) \\ \triangleq \mathbb{E}\left[ (\frac{t+1}{T} - V_{t+1})(P_{t+1} - P_t) | V_t, \mu_t \right] \\ -\gamma \operatorname{Var}\left[ (\frac{t}{T} - V_t)(P_{t+1} - P_t) | V_t, \mu_t \right],$$



$$u(t, V_t, \mu_t) \triangleq U(V_{t+1}; V_t, \mu_t) + \mathbb{E}_{\mu_{t+1}}[u(t+1, V_{t+1}, \mu_{t+1})],$$

for any t and a given  $\gamma$  as a positive constant that measures the degree of risk aversion.

If we consider  $\mu_{t+1}$  to be a discrete distribution, namely, a trinomial distribution  $(-\varepsilon, 0, +\varepsilon)$  with  $P(\mu_{t+1} = +\varepsilon) = p\lambda$  and  $P(\mu_{t+1} = -\varepsilon) = (1-p)\lambda$ , then under this circumstance, we can write our value function as  $u(t, V_t, \mu_t; v)$  where v is a function of the strategy employed at t + 1. Hence, for the signal with binary distribution,

$$= \{ \begin{array}{l} u(t, V_t, \mu_t; v) \\ (\frac{t+1}{T} - v)(\mu_t \Delta t) - \gamma(\frac{t}{T} - V_t)^2 \sigma^2 + (1 - \lambda \Delta t)u(t+1, v, 0) \\ + \lambda \Delta t p u(t+1, v, +\varepsilon) + (\lambda \Delta t (1-p))u(t+1, v, -\varepsilon) \}. \end{array}$$

Furthermore, we could define the optimal value function as  $u^*(t, V_t, \mu_t)$  in the following way:  $u^*(t, V_t, \mu_t)$ 

$$= \max_{V_{t+1} \ge V_t} \{ \begin{array}{l} (\frac{t+1}{T} - V_{t+1})(\mu_t \Delta t) - \gamma(\frac{t}{T} - V_t)^2 \sigma^2 \\ + (1 - \lambda \Delta t)u(t+1, V_{t+1}, 0) \\ + \lambda \Delta t pu(t+1, V_{t+1}, +\varepsilon) + (\lambda \Delta t(1-p))u(t+1, V_{t+1}, -\varepsilon) \}. \end{array}$$

Consequently, this optimal value function offers a solution that is a function of the strategy at t + 1, where we denote it as  $v^*(t, V_t, \mu_t)$ , that maximizes the value function:

$$v^{*}(t, V_{t}, \mu_{t}) = \arg\max_{V_{t+1} \ge V_{t}} \{ \mathbb{E}\left[ (\frac{t+1}{T} - V_{t+1})(P_{t+1} - P_{t}) | V_{t}, \mu_{t} \right] \\ -\gamma \operatorname{Var}\left[ (\frac{t}{T} - V_{t})(P_{t+1} - P_{t}) | V_{t}, \mu_{t} \right] \\ + \mathbb{E}_{\mu_{t+1}} [u(t+1, V_{t+1}, \mu_{t+1})]. \}$$

Therefore, we have

$$u^{*}(t, V_{t}, \mu_{t}) = \max_{v \in \mathcal{V}_{1}} u(t, V, \mu; v) = u(t, V, \mu; v^{*}).$$

where  $\mathcal{V}_1$  is the admissible set of all the strategy in T + 1 market, which satisfies  $v(t, V, \mu) \ge V, \forall v \in \mathcal{V}_1$ .

In addition, because the stock market closes at the last operation T, no more value could be added at T, the value function at T does not have an input strategy function. We have our boundary conditions defined :

$$u(T, V_T, \mu_T) = \begin{cases} 0, & V_T = 1, \\ -\infty, & V_T < 1. \end{cases}$$
(5)

Notice, for the boundary conditions, we have a punishment term of negative infinity when  $V_T < 1$  in order to stress the significance of achieving a specific target of the day.

By the definition above, because  $V_T$  can only be 1, we can easily calculate the value function and the optimal strategy at time T - 1 as follows.

$$v^*(T-1, V_{T-1}, \mu_{T-1}) = 1,$$

and

$$u^*(T-1, V_{T-1}, \mu_{T-1}) = -\gamma (\frac{T-1}{T} - V_{T-1})^2 \sigma^2.$$

They are both independent of the signal at T - 1, that is the  $\mu_{T-1}$ .

#### 4. Theoritical Analysis

### 4.1.No predication in T+1 market



Hence

In this case, we consider  $\mu_t = 0$  for every positive integer *t*. In other words, the investor makes completely no prediction about the drift part of the stock price. By mathematical induction, we could conclude that the optimal strategy for the investor is TWAP, where  $V_t = \frac{t}{T}$ .

Lemma 4.1 In the case of no predication, i.e.,  $\lambda = 0$ , the value function is always

$$u(t, V_t, 0; v) \le -\gamma (\frac{t}{T} - V_t)^2 \sigma^2, \forall 0 \le t \le T, t \in Z, 0 \le V_t \le \frac{t+1}{T}$$

Proof. We can tell easily that our Lemma 4.1. is equivalent as

$$u(T-n, V_{T-n}, 0; v) \leq -\gamma \left(\frac{T-n}{T} - V_{T-n}\right)^2 \sigma^2,$$
  
$$\forall 0 \leq n \leq T, n \in Z, 0 \leq V_{T-n} \leq 1.$$

By the boundary condition Eq. (3.5), for n = 0, we have

$$u(T, V_T, 0) \le 0, \forall 0 \le V_T \le 1$$

Next, we assume our lemma hold for all values of n up to some natural number k. Therefore, we have

$$u(T - n, V_{T-n}, 0; v) \le -\gamma (\frac{T - n}{T} - V_{T-n})^2 \sigma^2,$$
  
$$\forall n \le k < T, n \in Z, 0 \le V_{T-n} \le 1.$$

for 
$$n = k + 1$$
, we have  

$$u(T - k - 1, V_{T-k-1}, 0; v)$$

$$= \{(\frac{T-k}{T} - V_{T-k})(\mu_{T-k-1}\Delta t) -\gamma(\frac{T-k-1}{T} - V_{T-k-1})^2\sigma^2 + u(T-k, V_{T-k}, 0; v)\}$$

$$= u(T - k, V_{T-k}, 0; v) - \gamma(\frac{T-k-1}{T} - V_{T-k-1})^2\sigma^2.$$
Since  $u(T - k, V_{T-k}, 0; v) \leq -\gamma(\frac{T-k-1}{T} - V_{T-n})^2\sigma^2$ , we get  
 $u(T - k - 1, V_{T-k-1}, 0) \leq -\gamma(\frac{T-k-1}{T} - V_{T-k-1})^2\sigma^2.$ 

Thus, when our lemma is true for n = k, then it is true for n = k + 1. As it is true for n = 0, then it must be true for n = 0 + 1(n = 1). As it is true for n = 1 then it must hold true for n = 2 and so on for all positive integers n where  $0 \le n \le T$ .

From Lemma 4.1, we know that with no prediction, the value function could only generate a negative result or zero. In other words, investors should not expect positive results on the investment when no information could be known beforehand for any prediction to be made.

Lemma 4.2 In the case of no predication, i.e.,  $\lambda = 0$ , the optimal strategy is

$$v*(t,V,0)=t+1T, \forall 0 \leq t < T, t \in \mathbb{Z}, 0 \leq V \leq tT.$$

Proof. Because of Lemma 4.1., we only need to verify that

$$u(t, V_t, 0; \frac{t+1}{T}) = -\gamma (\frac{t}{T} - V_t)^2 \sigma^2.$$

First, we can rewrite our lemma as

$$v^*(T-m,V_{T-m},0)=\frac{T-m+1}{T},$$

 $\forall 1 \le m \le T, m \in Z, 0 \le V_{T-m} \le 1.$ Note in this case,  $V_T = 1$ . Hence, our boundary condition would be  $u(T, V_T, 0) = 0.$ 

Therefore, for 
$$m = 1, v^*(T, V_T, 0) = 1$$
, we have



$$u(T-1, \frac{T-1}{T}, 0; 1)$$

$$= (\frac{T}{T}-1)(\mu_{T-1}\Delta t) - \gamma(\frac{T-1}{T}-V_{T-1})^2\sigma^2 + u(T, V_T, 0)$$

$$= -\gamma(\frac{T-1}{T}-V_{T-1})^2\sigma^2.$$

Next, we assume our lemma holds for all values of m up to some natural number k. Therefore, we have

$$u(T - k, V_{T-k}, 0; \frac{T - k + 1}{T}) = -\gamma (\frac{T - k}{T} - V_{T-k})^2 \sigma^2,$$
  
$$\forall m \le k \le T, m \in Z, 0 \le V_{T-m} \le 1$$

Hence for m = k + 1, we get

$$u(T-k-1,\frac{T-k-1}{T},0;\frac{T-k}{T})$$

$$= (\frac{T-k}{T}-\frac{T-k}{T})(\mu_{T-k-1}\Delta t) - \gamma(\frac{T-k-1}{T}-V_{T-k-1})^2\sigma^2$$

$$+u(T-k-1,\frac{T-k}{T},0)$$

$$= -\gamma(\frac{T-k-1}{T}-V_{T-k-1})^2\sigma^2 - \gamma(\frac{T-k}{T}-\frac{T-k}{T})^2$$

$$= -\gamma(\frac{T-k-1}{T}-V_{T-k-1})^2\sigma^2.$$

As a result, if this lemma is true for m = k, then it is true for m = k + 1. As it is true for m = 1, then it must be true for m = 2. As it is true for m = 2 then it must hold true for m = 3 and so on for all positive integers m where  $1 \le m \le T$ .

To conclude, TWAP is the optimal strategy at any time regardless of the previous position for the investor if no information could be gained from the stock market and no prediction could be made. In such situation, the best or highest return that could be generated by every value function will be the outcome value when TWAP is employed. If the current position is lower or equal to the next position offered by TWAP, then follow TWAP. On the other hand, if the current position is higher, then no action is the best action as the next move.

#### 4.2.T + 0 market

In a T + 0 market, we do not have the restrictions on our strategy besides the boundary conditions. Therefore, we have

$$= \max_{V_{t+1}} \{ (\frac{t+1}{T} - V_{t+1})(\mu_t \Delta t) - \gamma (\frac{t}{T} - V_t)^2 \sigma^2 + (1 - \lambda)u(t+1, V_{t+1}, 0) + \lambda pu(t+1, V_{t+1}, +\varepsilon) + \lambda (1 - p)u(t+1, V_{t+1}, -\varepsilon) \}.$$

Now we could calculate some results from specific value functions:



$$u^{*}(T-2, V_{T-2}, \mu_{T-2})$$

$$= \arg \max_{V_{T-1}} \{ \mathbb{E} \left[ \left( \frac{T-1}{T} - V_{T-1} \right) (P_{T-1} - P_{T-2}) | V_{T-2}, \mu_{T-2} \right] -\gamma \operatorname{Var} \left[ \left( \frac{T-2}{T} - V_{T-2} \right) (P_{T-1} - P_{T-2}) | V_{T-2}, \mu_{T-2} \right] + \mathbb{E}_{\mu_{T-1}} [u(T-1, V_{T-1}, \mu_{T-1})] \}$$

$$= \arg \max_{V_{T-1}} \{ \left( \frac{T-1}{T} - V_{T-1} \right) (\mu_{T-2} \Delta t) - \gamma \left( \frac{T-2}{T} - V_{T-2} \right)^{2} \sigma^{2} -\gamma \left( \frac{T-1}{T} - V_{T-1} \right)^{2} \sigma^{2} \}.$$

Solving this quadratic function, we get

$$v^*(T-2, V_{T-2}, \mu_{T-2}) = \frac{T-1}{T} - \frac{\mu_{T-2}\Delta t}{2\gamma\sigma^2}$$

After substitution, we get

$$u^{*}(T-2, V_{T-2}, \mu_{T-2}) = \frac{\mu_{T-2}^{2}\Delta t^{2}}{4\gamma\sigma^{2}} - \gamma(\frac{T-2}{T} - V_{T-2})^{2}\sigma^{2}.$$

Furthermore, we have

$$v^*(T-3, V_{T-3}, \mu_{T-3}) = \frac{T-2}{T} - \frac{\mu_{T-3}\Delta t}{2\gamma\sigma^2},$$

and

$$u^{*}(T-3, V_{T-3}, \mu_{T-3}) = \frac{\mu_{T-3}^{2} \Delta t^{2}}{4\gamma \sigma^{2}} - \gamma (\frac{T-3}{T} - V_{T-3})^{2} \sigma^{2} + \frac{\lambda \Delta t \varepsilon^{2} \Delta t^{2}}{4\gamma \sigma^{2}}.$$

From careful observation of the numerical results, we found the pattern and proved it mathematically as following:

Lemma 4.3 The optimal value function in T + 0 market takes the form of

$$u^{*}(T-t, V_{T-t}, \mu_{T-t}) = \frac{\mu_{T-t}^{2} \Delta t^{2}}{4\gamma \sigma^{2}} - \gamma (\frac{T-t}{T} - V_{T-t})^{2} \sigma^{2} + C_{t},$$

where  $\{C_t\}_{\{0 \le t \le T\}}$  satisfies:

$$C_{k+1} = C_k + \lambda (\frac{\varepsilon^2 \Delta t^2}{4\gamma \sigma^2}),$$
$$C_1 = 0$$

and is free from the influence of  $V_{T-\eta}$  and  $\mu_{T-\eta}$  and  $2 \le t \le T, 0 \le V_{T-\eta} \le 1$ . For t = 1,  $\mu^*(T-1, V_{T-1}, \mu_{T-1}) = -\gamma (\frac{T-1}{2} - V_{T-1})^2 \sigma^2$ .

$$u^*(T-1, V_{T-1}, \mu_{T-1}) = -\gamma(\frac{1}{T} - V_{T-1})^2 \sigma^2.$$

Proof. Base case is already presented in the numerical examples above. We could assume this lemma holds true for all values of t up to some natural number k,

$$u^{*}(T-t, V_{T-t}, \mu_{T-t}) = \frac{\mu_{T-t}^{2} \Delta t^{2}}{4\gamma \sigma^{2}} - \gamma (\frac{T-t}{T} - V_{T-t})^{2} \sigma^{2} + C_{t},$$
  
$$\forall t \le k \le T, 0 \le V_{T-t} \le 1.$$

Hence, for t = k + 1,



$$\begin{aligned} u^{*}(T-k-1,V_{T-k-1},\mu_{T-k-1}) \\ &= \max_{V_{T-k}} \{ \begin{array}{l} (\frac{T-k}{T}-V_{T-k})(\mu_{T-k-1}\Delta t) - \gamma(\frac{T-k-1}{T}-V_{T-k-1})\sigma^{2} \\ &+ (1-\lambda)u(T-k,V_{T-k},0) \\ &+ \lambda pu(T-k,V_{T-k},+\varepsilon) + \lambda(1-p)u(T-k,V_{T-k},-\varepsilon) \} \\ &= \max_{V_{T-k}} \{ \begin{array}{l} (\frac{T-k}{T}-V_{T-k})(\mu_{T-k-1}\Delta t) - \gamma(\frac{T-k-1}{T}-V_{T-k-1})\sigma^{2} \\ &+ (1-\lambda)(-\gamma(\frac{T-k}{T}-V_{T-k})^{2}\sigma^{2}+C_{k}) \\ &+ \lambda p(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}} - \gamma(\frac{T-k}{T}-V_{T-k})^{2}\sigma^{2}+C_{k}) \\ &+ \lambda(1-p)(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}} - \gamma(\frac{T-k}{T}-V_{T-k})^{2}\sigma^{2}+C_{k}) \} \\ &= \max_{V_{T-k}} \{ \begin{array}{l} (\frac{T-k}{T}-V_{T-k})(\mu_{T-k-1}\Delta t) - \gamma(\frac{T-k-1}{T}-V_{T-k-1})\sigma^{2} \\ &- \gamma(\frac{T-k}{T}-V_{T-k})^{2}\sigma^{2}+C_{k} + \lambda(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}}). \} \end{array} \end{aligned}$$

Still, this is a quadratic function with solution

$$v^*(T-k-1,V_{T-k-1},\mu_{T-k-1}) = \frac{T-k}{T} - \frac{\mu_{T-k-1}\Delta t}{2\gamma\sigma^2}.$$

Substituting, we get

$$u^{*}(T-k-1, V_{T-k-1}, \mu_{T-k-1})$$

$$= \frac{\mu_{T-k-1}^{2}\Delta t^{2}}{4\gamma\sigma^{2}} - \gamma(\frac{T-k-1}{T} - V_{T-k-1})^{2}\sigma^{2} + C_{k} + \lambda\Delta t(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}})$$

$$= \frac{\mu_{T-k-1}^{2}\Delta t^{2}}{4\gamma\sigma^{2}} - \gamma(\frac{T-k-1}{T} - V_{T-k-1})^{2}\sigma^{2} + C_{k+1},$$

where

$$C_{k+1} = C_k + \lambda \Delta t \left(\frac{\varepsilon^2 \Delta t^2}{4\gamma \sigma^2}\right).$$

By mathematical induction, we know that this proposition holds true for every positive integer t where  $2 \le t \le T$ .

Based on this conclusion, we could simply deduce the general expression of the optimal strategy from the optimal value function followed by the investor throughout the day.

Theorem 4.4 In the T + 0 market, the optimal strategy is

$$v^*(t, V, \mu) = \frac{t+1}{T} - \frac{\mu \Delta t}{2\gamma \sigma^2}, \forall 0 \le t < T, t \in \mathbb{Z}, 0 \le V \le \frac{t}{T},$$

which is independent of the position V.

Proof. For any t,

$$\begin{aligned} u^{*}(T-t,V_{T-t},\mu_{T-t}) \\ &= \max_{V_{T-t+1}} \{ \begin{array}{l} (\frac{T-t+1}{T} - V_{T-t+1})(\mu_{T-t}\Delta t) - \gamma(\frac{T-t}{T} - V_{T-t})\sigma^{2} \\ &+ (1-\lambda)u^{*}(T-t+1,V_{T-t+1},0) \\ &+ \lambda p u^{*}(T-t+1,V_{T-t+1},+\varepsilon) \\ &+ \lambda (1-p)u^{*}(T-t+1,V_{T-t+1},-\varepsilon) \} \\ &= \max_{V_{T-t+1}} \{ \begin{array}{l} (\frac{T-t+1}{T} - V_{T-t+1})(\mu_{T-t-1}\Delta t) - \gamma(\frac{T-t}{T} - V_{T-t})\sigma^{2} \\ &- \gamma(\frac{T-t+1}{T} - V_{T-t+1})^{2}\sigma^{2} + C_{t-1} + \lambda(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}}) \}, \end{aligned} \end{aligned}$$

where the last equality is guaranteed by Lemma 4.3.

Hence, the optimal strategy at time T - t is



$$v^*(T-t, V_{T-t}, \mu_{T-t}) = \frac{T-t+1}{T} - \frac{\mu_{T-t+1}\Delta t}{2\gamma\sigma^2},$$

which is a maximizer of

$$(\frac{T-t+1}{T} - V_{T-t+1})(\mu_{T-t-1}\Delta t) - \gamma(\frac{T-t}{T} - V_{T-t})\sigma^{2} - \gamma(\frac{T-t+1}{T} - V_{T-t+1})^{2}\sigma^{2} + C_{t-1} + \lambda(\frac{\varepsilon^{2}\Delta t^{2}}{4\gamma\sigma^{2}}).$$

#### 5. Numerical Simulation

In this section, we set up our model for the stock price and implement the optimization through the maximization of the value function in Python in order to determine the position line of the optimal strategy employed.

In this simulation, we set up our parameters in the following way:  $\overline{T} = 4, \Delta t = \frac{1}{60}, \ \overline{\sigma} = 0.01, \ \lambda = \frac{1}{60}, \ \varepsilon = 0.002, \ p = 0.5. \ \overline{T}$  represents the total four trading hours of the stock market per day, and the value of  $\Delta t$  shows that the minimal time interval required between operations we consider in this case is one second. The daily volatility  $\overline{\sigma}$  is set to be 0.01, a realistic measure associated with relatively low risk, considering the median of the Standard & Poors 500 daily volatility index from 1962 to 2018 as 0.014 (Easterling, 2019)<sup>[3]</sup>. The value of  $\lambda$  results in  $P(\mu_n \neq 0) = \frac{1}{60}$  and  $P(\mu_n = 0) = \frac{59}{60}$ , meaning that the signals appear in a relatively rare manner. Moreover, the values of  $\varepsilon$  and p represent the fact that  $P(\mu_n = 0.002 | \mu_n \neq 0) = \frac{1}{2}$ ,  $P(\mu_n = -0.002 | \mu_n \neq 0) = \frac{1}{2}$ , showing the fact that the drop and rise of the price are not only equally likely to happen, but also occur with the same magnitude.

All of the parameters mentioned above are independent of each other, and the values of T and  $\sigma$  could be calculated from them by the formulas above: T = 240,  $\sigma \approx 6.5 \times 10^{-5}$ . For the convenience of the readers, all the parameters are shown in Table 1

T	4	Δt	$\frac{1}{60}$
ō	0.01	λ	$\frac{1}{60}$
ε	0.002	р	0.5
Т	240	σ	$6.5  imes 10^{-5}$

Table 1 Default values of parameters

With all the parameters set, we run the simulation and generate the following three graphs. Notice that every coloured position line in the first and the last graph indicates the situation that the investor receives its corresponding signals consistently over the total trading time.

In the first graph of Figure 1, the position lines in the T + 0 market are straight line parallel to the TWAP line, which corresponds to Theorem 4.4. The position lines in the T + 1 market are shown in the third graph of Figure 1, which



Figure 1 The signal lines and simulated optimal position lines in T+0 and T+1 markets

are kinked straight lines parallel to the TWAP line. Same as the position lines in the T + 0 market, the undesired signal line (green) above the no signal line (yellow) and the no signal line above the preferred signal line (red).

Additionally, with the same signals appeared over time indicated by the second graph of Figure 1, the optimal position line generated in T + 0 market and T + 1 market has shown great difference. For T + 0 market, with no restrictions, the next position could shift freely among the three lines depending on the next signal; while for T + 1 market, due to the restrictions, the next position could only go up or stay the same as the current position. In this case, the optimal position line will be mainly between the undesired signal line and the no signal line. Because it is rather impossible for all the signals to be preferred signals, when the optimal position react to an undesired signal or no signal, it could not go down to the preferred signal line in the future.

## 6. Conclusion

In this paper, we solve the optimal strategy through the recursive mean-variance analysis, when the investor is able to make prediction about the market price .

In T + 0 market, we obtain the explicit expression of the value function for each optimal feedback control. In our example, given three different signals, three position lines corresponding to different signals are parallel to each other and the optimal position is determined by the three position lines. The next optimal position lands on the position line corresponding to the signal seen by the investor at this moment.

In T + 1 market, we prove that the TWAP is the optimal strategy when the investor is unable to formulate any prediction. For the general case, the optimal strategy is given by the numerical simulation in Section 5, and we discover that the optimal feedback control is determined by the three lines. Further analyses about the explicit expression of the value function still remain an open question due to the constraints of the T + 1 market.

### References

[1] Bansal, R., Harvey, C. and Dahlquist, M. (2004). Dynamic Trading Strategies and Portfolio Choice. SSRN Electronic Journal.



- [2] Basak, S. and Chabakauri, G. (2009). Dynamic Mean-Variance Asset Allocation. SSRN Electronic Journal.
- [3] Easterling, E. (2019). [online] Crestmontresearch.com. Available at: https://www.crestmontresearch.com/docs/Stock-Volatility-Perspective.pdf [Accessed 14 Sep. 2019].
- [4] Jagannathan, R. and Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. The Journal of Finance, 58(4), pp.1651-1683.
- [5] Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), p.77.