Anomaly Correction by Optimal Trading Frequency

Yiqiao Yin
Columbia University

September 9, 2016

Abstract

Under the assumption that security prices follow random walk, we look at price versus different moving averages. Different periods of moving averages give investor different signals and we assume that a rational investor would want to buy more when the price goes down. This paper provides a theoretical model for an investor to systematically buy heavy when the security prices go down.

Contents

1 Introduction

Security prices are generally understood as random walk. Although some scholars still doubt the concept of efficient market hypothesis, there are scholars like Malkiel shown us empirical evidence that we do observe data in favor of efficient market hypothesis. He has shown us that professional investment managers do not outperform their index benchmarks and has provided us evidence that prices do digest all available information [4].

In this paper, we would use Malkiel's argument as an major assumption. That is, we assume market is efficient over the long run and security prices do reflect all available information out there. This is also to say that the first order derivative of price (price changes) follows binomial distribution (big probabilities that absolute values of return are small and small probabilities that absolute values of return are big). Intuitively speaking, we do not argue whether individual should spend his time or energy to pick stocks on his own because it is too costly to do pick small probabilistic events. Instead, we are providing a model to correct the anomaly in the market.

Jegadeesh and Titman (2007) has found evidence that strategies of buying stocks performed well in the past and selling stocks that performed poorly in the past generate significant positive returns in 3- and 12- month holding period [3]. Similar ideas are presented by De Bondt and Thaler (1985, 1987) [1][2]. This up-minus-down momentum factor in the security market provides another important ground work for us. If the strategies introduced by these scholars are worth running (which proved that they are), there would be rational investors out there who want to try these strategies. This provide liquidities for our model. If no one is willing to buy high sell low, it does not make sense for us to argue to buy low sell high. Hence, we would assume that there is some liquidities out there provided, that is, the market exists.

Another thing that we need to pay attention to is the concept of anomaly. A value strategy can be an anomaly if significant alpha can be found. A momentum strategy can be an anomaly if its alpha is significant. We define anomalies among a pool of expected returns to be the ones cannot be tolerated by an investor subjectively. This is a definition that has not been discussed before. As a matter of fact, this definition is difficult to start any quantitative approach because of the adjective "subjective". We set up this definition on purpose because we will introduce a trading frequency δ to be associated with the investor's tolerance of the security prices and hence we can start to have some discussions about anomalies.

The questions we want to answer in this paper originate from the assumptions above. If market is efficient and incorporates all publicly available information, it is rational to infer that a market grows by a various amount of factors out there, say GDP, NGDP, interest rates, fundamentals, technicals, algorithms, and etc. since all of the information are public. Then we would ideally see a relatively straightened line of prices throughout the history if we compare to other indicators. However, we observed two very different variations between security market and economic market. Scholars can find factors and use pricing models to find out why the security market is so volatile, hence, they can explain the reason why we have such observations, i.e., there are a lot more spikes in the price chart of stock market than there are in GDP chart (see Figure 1 and 2). Is it all possible to simply buy the market when market is low and just let the market grow with the economy? If so, how do we do it in a consistent and rational sense that the lower the market dips the more we buy?

Figure 1: The return after interest rate with $1 investment in the market from July of 1926 to June of 2016. *Source: From Ken French Data Library.*

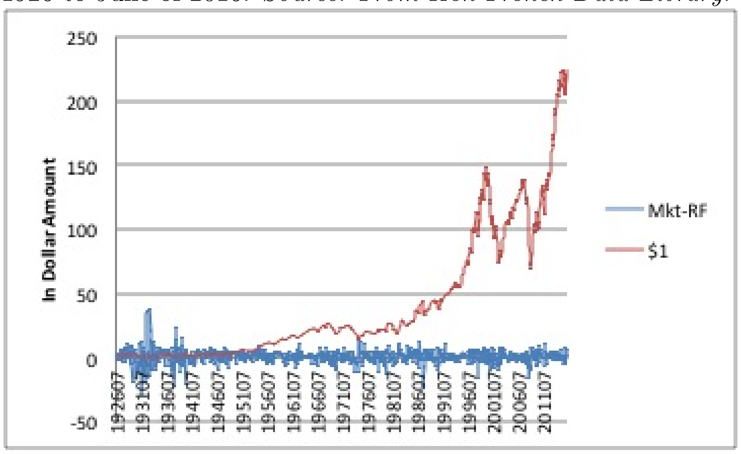
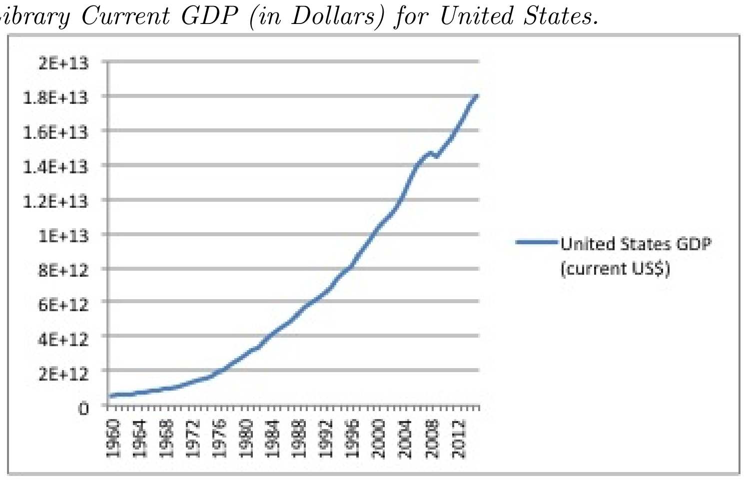


Figure 2: The GDP in US Dollars from 1960 to 2015. *Source: From World Bank Library Current GDP (in Dollars) for United States.*



2 Mathematical Model

2.1 Moving Averages

We look at price as a time-series function. That is, for each time node t, we have an executed price p. This can be seen as a discrete function with a map from time to price. We can denote price as p_t.

Anything can happen any time in the market and an individual price level is usually very volatile. We incorporate the concept of moving average so that we can look at the average price of some period in the past, which allows us to overlook the trading noise in the market.

Definition 2.1. Simple moving average (SMA) takes the average of prices traded n time period(s) in the past,

$$SMA_n = \frac{1}{n}\sum_{t=0}^{n} P_t$$

.

We can also implement the concept of exponential moving average, which is to add a certain amount of weight to the current prices when calculating the averages.

2.2 Price to Moving Averages

Based on the moving averages of the prices, we can establish the concept of price-to-moving averages.

Definition 2.2. Price-to-Moving Average is the ratio of price over a selected period of moving averages. For n periods, we have

$$P_t - SMA_n = \frac{P_t}{\frac{1}{n}\sum_{t=0}^{n} P_t} - 1$$

.

We can interpret this definition by our assumptions. We assumed that price follows random walk. The averages of a random walk should also be random walk. That is to say, we have two series of numbers following binomial distributions, which implies that the ratio should also be a binomial distribution. The data had an average expected return of 0.65%. If we imagine this to be the initial x-axis, we would observe a binomial distribution for price-to-moving averages for all three data series.

Figure 3: The graph shows returns after interest rate with $1 investment in the market along with three other moving averages with time periods to be 50-day, 100-day, and 150-day from July of 1926 to June of 2016. *Source: From Ken French Data Library.*

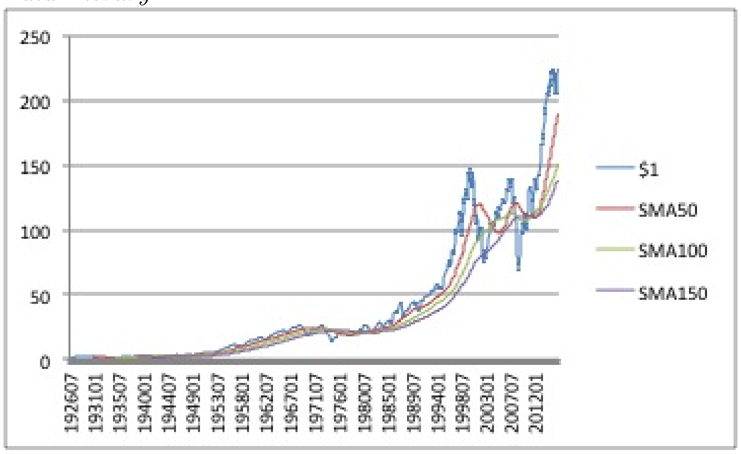
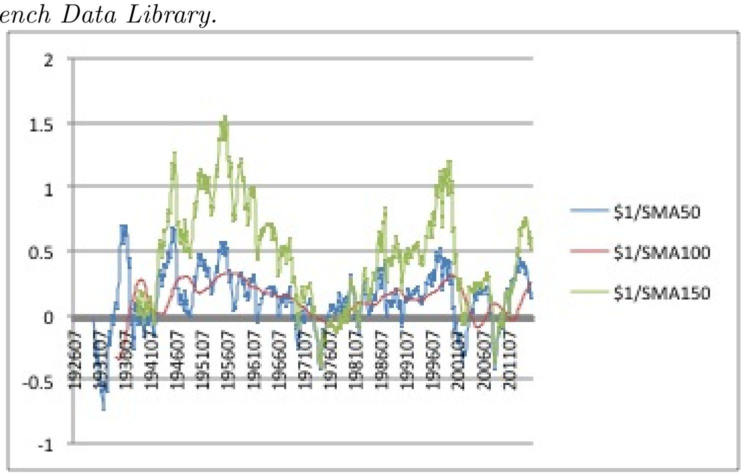


Figure 4: The graph shows price-to-moving averages with time periods to be 50-day, 100-day, and 150-day from July of 1926 to June of 2016. *Source: From Ken French Data Library.*



2.3 Trading Frequency

We defined in the introduction section of this paper that we call anomalies to be the ones among a pool of observations that cannot be tolerated by investors subjectively. This concept is associated with a trading frequency δ. An investor could be greedy or moody when he made an investment decision. We do not care. We summarize all of those emotions into δ.

Emotionally impatient investors tend to be less educated and would buy very often, hence a high trading frequency (δ is big). Emotionally patient investors tend to be more educated and would buy once in a while, hence a low trading frequency (δ is small).

Based on each price-to-moving average, we can calculate standard deviation, σ. For every $p_t - SMA_n$, we have a calculated $\sigma_{t,n}$. From binomial distribution, we know that the probability that $p_t - SMA_n$ is high is small and the probability that $p_t - SMA_n$ is low is big. Hence, for the observation space $X = (t, n)$ and a particular event $k = (i, j)$, we have

$$\forall i, j \in \mathbb{R}, Pr(p_t - SMA_n = p_i - SMA - j) = Pr(X = k)$$

$$= \binom{n}{k} p^k (1-p)^{n-k} \forall k \in \mathbb{Z}, \ where \ \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The interpretation is the following. There exists some i and j such that the probability that the price-to-moving average equals to this particular i and j is defined by the binomial equation.

The next question we need to think about is how do we understand or even quantify how many particular pairs of (i, j) we need to choose. The answer to this question is dependent on individual investor, which is associated with trading frequency δ. Moreover, the trading frequency δ is dependent on standard deviation $\sigma_{t,n}$. In other words, we need the following definition.

Definition 2.3. The particular event k is the product of δ and total observations with

$$\exists h \in \mathbb{R}, \delta = Pr(\sigma_{t,n} \geq h\sigma_{t,n}) = \frac{\sum_{(t,n)\to(t',n')}^{(t'',n'')} (h\sigma_{t,n})}{\sum_{(t,n)} \sigma_{t,n}}$$

while

$$\sigma_{t,n} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((p_t - SMA_i) - \frac{1}{n}\sum_{i=1}^{n}(p_t - SMA_i))^2}$$

This equation looks complicated, yet it is easy to understand. The amount of event k happens in a way that is dependent on trading frequency δ. Trading frequency δ is defined as a probability. It is a probability calculated to be the amount of times standard deviation of price-to-moving averages to be bigger than an ideal amount over the amount of times of all observations. In other words, we are looking at a certain area on the two tails of the binomial distribution. We are more interested in taking care of the left tail, i.e., when to buy instead of when to sell.

2.4 Optimal Trading Frequency

How often do we have to trade? Each of us has a unique risk tolerance as well as expected return for our investment. Then each of us has a particular k we are using when making investment decisions. This is to say, we need to find the right amount of k around the price when the price-to-moving average is the lowest. The bigger k an investor has the earlier he needs to start buying.

The first goal is to find the optimal price-to-moving averages level. We need to take the first order derivative of price-to-moving averages and set the result to zero. We have the following theorem.

Theorem 2.4. *$\forall n$, let $\triangle$ be a small unit of time period, when $\sum_{t=0}^{n} P_t = \sum_{t=0}^{n} P_{t+\triangle}$, we have a critical price to act on it.*

The interpretation is fairly intuitive. We look at a series of price and the price goes up or down. We calculate the average of prices in a certain period. Then we have a series of average prices. We look at the ratio of price over moving averages and we treat the ratio to be a new number series with some arbitrary mean. We calculate the sum of prices in the past certain of time nodes. The optimal price level occurs at a time when the sum of the averages at two time nodes does not change at all (or change tiny little). Intuitively speaking, this is a time when the price is going down too much and now it is starting to bounce back, causing the sum of the past prices to be relatively stable. For example, we calculate the sum of hundreds of prices in the past. If the price keeps going down, we are supposed to get a sum smaller and smaller, yet one day the sum stays the same all of the sudden. Then this price can be critical to buy and vice versa. The complete proof is shown in the Appendix 4.1.

In practice, we will probably find some price areas as critical price level. That is to say, we would not find a perfect equivalent sums of prices in the data. We would probably see a range of prices that satisfy the condition. Then we can apply trading frequency δ. The bigger the δ is the bigger the time node is. Therefore, a consistent investor who is true to his own preference would start to buy earlier and vice versa (sell earlier). The smaller the δ is the smaller the time nodes is. A consistent investor would buy later and vice versa.

3 Conclusion

In this paper we present a method to pick a critical price by using moving averages and investor trading frequency.

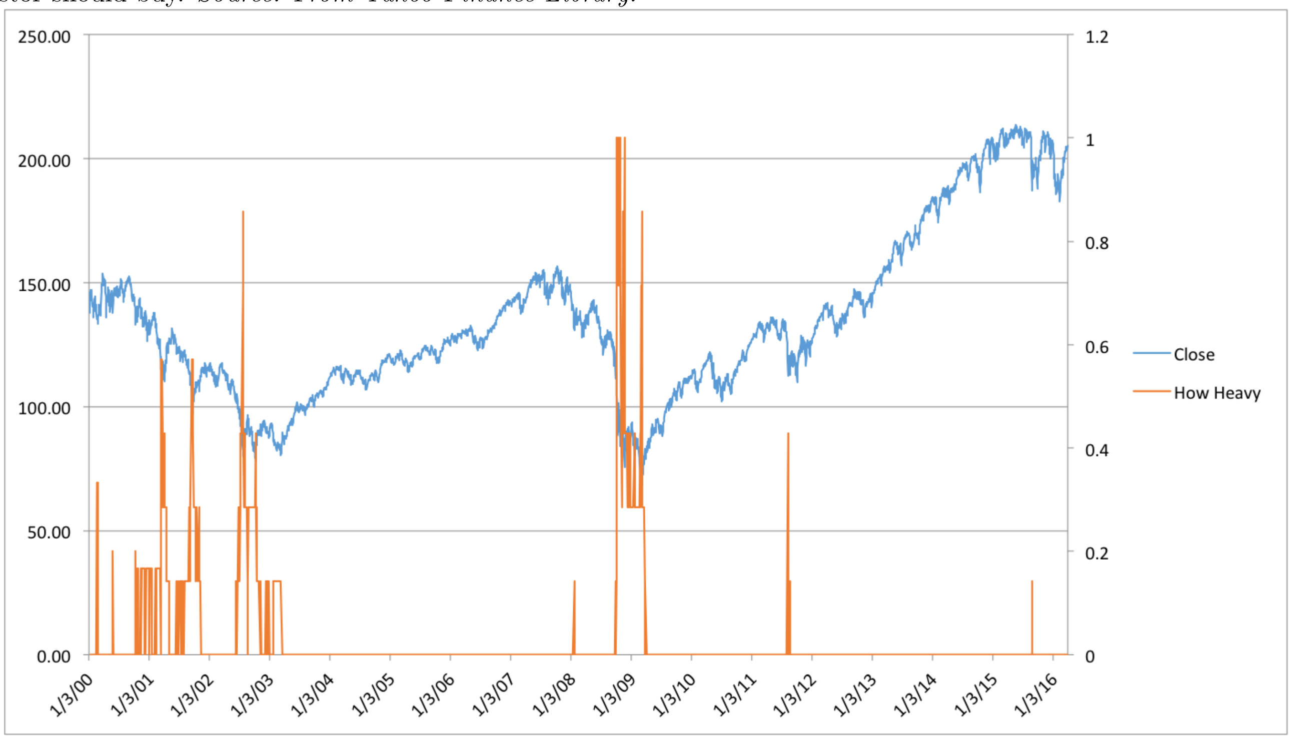
We assume price follows a random walk and we assume sufficient liquidities for to make a market. We provide a model such that an investor can find a critical price to be local minimum or local maximum to act on.

The essence of this paper rests in the goal of getting rid of the ups and downs of the equity market. We present this model hoping to help equity market to be less volatile.

3.1 Discussion of Applications

In practice, this model definitely can pick up some good price ranges. In Figure 5, we have shown the daily price plot for S&P 500 from January 2000 to January 2016 from Yahoo Data Library. We calculated the moving averages on 10-, 20-, 30-, 50-, 100-, 200-, and 300-day periods. For less than 10% trading frequency, we calculate the buy signal for each price-to-moving averages, which are presented as spikes on x-axis. The more price-to-moving averages based on different time periods satisfy the condition the higher the spikes are. The higher the spikes are the heavier an investor should buy. We observe that the spikes pretty much covered both financial crisis (2001 and 2008) and the model successfully tells us, by using only past data at each time node, that when to start buying and when to buy heavy.

Figure 5: The graph shows S&P 500 from January 2000 to January 2016. The model takes 10-, 20-, 30-, 50-, 100-, 200-, 300-day to be moving average periods and calculate the price-to-moving average ratio by investor trading frequency to be less than 10%. The spikes from x-axis are buy signal calculated based on critical price level. The higher the spikes are the heavier an investor should buy. *Source: From Yahoo Finance Library.*



4 Appendix

4.1 Proof of Theorem 2.4

Proof: We have price-to-moving average to be the following. For n periods, we have

$$P_t - SMA_n = \frac{P_t}{\frac{1}{n}\sum_{t=0}^{n} P_t} - 1$$

We want to know what time is the price the lowest. Thus, we want to take derivative in respect with time t. Then set the first order derivative to be zero and we have

$$\frac{\partial(P_t - SMA_n)}{\partial t} = lim_{\triangle \to 0} \frac{(\frac{P_{t+\triangle}}{\frac{1}{n}\sum_{t=0}^{n} P_{t+\triangle}} - 1) - (\frac{P_t}{\frac{1}{n}\sum_{t=0}^{n} P_t} - 1)}{t} = 0$$

$$\Rightarrow lim_{\triangle \to 0} \frac{\frac{P_{t+\triangle}}{\frac{1}{n}\sum_{t=0}^{n} P_{t+\triangle}} - \frac{P_t}{\frac{1}{n}\sum_{t=0}^{n} P_t}}{t} = 0$$

$$\Rightarrow lim_{\triangle \to 0} \left(\frac{1}{t}\right)\left(\frac{P_{t+\triangle}}{\frac{1}{n}\sum_{t=0}^{n} P_{t+\triangle}} - \frac{P_t}{\frac{1}{n}\sum_{t=0}^{n} P_t}\right) = 0$$

$$\Rightarrow lim_{\triangle \to 0} \left(\frac{n}{t}\right)\left(\frac{P_{t+\triangle}}{\sum_{t=0}^{n} P_{t+\triangle}} - \frac{P_t}{\sum_{t=0}^{n} P_t}\right) = 0$$

$$\Rightarrow lim_{\triangle \to 0} \left(\frac{n}{t}\right)$$

$$\left((P_{t+\triangle}) \times (\sum_{t=0}^{n} P_t) - (P_t) \times (\sum_{t=0}^{n} P_{t+\triangle})\right) \times (\sum_{t=0}^{n} P_t \times \sum_{t=0}^{n} P_{t+\triangle})^{-1} = 0$$

Then we need

$$\Rightarrow \left((P_{t+\triangle}) \times (\sum_{t=0}^{n} P_t) - (P_t) \times (\sum_{t=0}^{n} P_{t+\triangle})\right) = 0$$

Since $P_{t+\triangle} = P_t$ when $\triangle \to 0$ (the price this second is very close to the price next second), we can factor out one of them

$$\Rightarrow (P_{t+\triangle}) \times \left((\sum_{t=0}^{n} P_t) - (\sum_{t=0}^{n} P_{t+\triangle})\right) = 0$$

$$\Rightarrow (\sum_{t=0}^{n} P_t) - (\sum_{t=0}^{n} P_{t+\triangle}) = 0$$

Hence, when $(\sum_{t=0}^{n} P_t) = (\sum_{t=0}^{n} P_{t+\triangle})$, we have the optimal price in the market that has the first order derivative to be zero (or approximately zero), which shows that this price is a local minimum/maximum (a critical price or an actionable price).

Q.E.D.

References

[1] De Bondt, Werner F.M., and Thalor, Richard (1985), "Does the Stock Market Overreact?" *Journal of Finance*, 40, 793-805.

[2] De Bondt, Werner F.M., and Thalor, Richard (1987), "Further Evidence of Investor Overreaction and Stock Market Seasonality", *Journal of Finance*, 42, 557-581.

[3] Jegadeesh, Narasimha, and Titman, Sheridan (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *The Journal of Finance*, 48(1), 65-91.

[4] Malkiel, Burton G. (2007), "Reflections on the Efficient Market Hypothesis: 30 Years Later", *The Financial Review*, 40, 1-9.