INSTITUTIONAL ALGORITHMIC TRADING, STATISTICAL ARBITRAGE AND TECHNICAL ANALYSIS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Ning Shen

August 2009

ABSTRACT

Technical analysis tools are widely used by short term investors in the financial market to identify trading opportunities and generate abnormal profit. Two of the most popular ones, Moving Average Convergence – Divergence and Bollinger Bands, are adopted in this study for algorithmic traders and statistical arbitragers (intraday trading) to reveal their effectiveness in terms of realizing sizeable profit before and after transaction cost. The simple oscillator signals derived from MACD and BB fail to efficiently recognize optimal trading timing and negative profit before and after transaction cost are realized under both strategies. Numerical analysis describes the sensitivity of profit with and without transaction fee to the strategies parameters. The results disclose that the selection of relevant parameters is not able to improve the performance of the strategies. A Long Only Filter Strategy (LOFS) is created to further investigate the possible strategies employed by institutional investors. Successfully generating considerable profit after transaction cost with a significant lower level risk, LOFS outperforms the buy-and-hold benchmark strategy as well as MACD and Bollinger Bands. LOFS is a promising strategy for statistical arbitragers who aim to profit from trading after accounting for transaction costs.

BIOGRAPHICAL SKETCH

Ning Shen was born in 1983 in Hefei, the capital city of Anhui province in China. In 2001, she graduated from Hefei No. 1 high school, the top first high school in Anhui province. Due to her excellent ranking on the national entrance exams and first honor award in the high school Olympic chemistry contest, she was admitted to Peking University. At Peking University, she studied both atmospheric sciences and economics, graduating with dual bachelor degrees in 2005. She was awarded a full scholarship by Georgia Institute of Technology to continue her study of atmospheric science. In 2007, she graduated with a master of science and then moved to Cornell University to study applied economics. She finished her masters study at the department of Applied Economic and Management and received scholarship as a teaching-assistant.

To My Beloved Parents

ACKNOWLEDGMENTS

Finally it is time to write up the acknowledgements. It has been a long journey here and there were so many helping hands on the way. Firstly, I would like to express my deepest gratitude to my Chair, Professor Hazem Daouk, for his priceless advice, guidance, patience, and never-default support throughout the entire research. Also I would like to bring my appreciation to my committee member, Professor Vicki Bogan, for her precious suggestion and support for my work.

I would like to thank all my colleagues and my friends at Cornell who have provided me both intellectual stimulants as well as enjoyable social time. Also my appreciation should go to AEM and Cornell faculty from whom I learned necessary knowledge to complete this study.

Finally I would like thank my beloved family, from whom I received unconditional and unfailing support, trust and love. Without them, I would never be where I am. I love you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMEX: American Exchange

AP: Adjusted Profit

BB: Bollinger Bands

EMA: Exponential weighted Moving Average

EP: Excess Profit

ETF: Exchange Traded Funds

LOFS: Long Only Filter Strategy

M&A: Merger and Acquisition

MA: Moving Average

MACD: Moving Average Convergence Divergence

NASDAQ: National Association of Securities Dealers Automated Quotations

NYSE: New York Stock Exchange

PP: Portfolio Profit

SMA: Simple weighted Moving Average

SPY: S&P Dep Recipts S&P 500 ETF

STD: Standard Deviation

TA: Technical Analysis

1 INTRODUCTION

1.1 Background and Motivation

When many hedge funds, private equity firms and investment banks faded into history during the severe financial crisis from 2007, some less-known financial firms, such as Infinium Capital Management L.L.C., DRW Holdings L.L.C. and Traditum Group L.L.C., thrived and racked up record profits. (Crain's Detroit Business, Jan 2009)

Different from traditional financial institutions, these proprietary trading firms concentrate themselves on light-speed trading on short-term trading movements and heavily depend on software and high-speed connections to trade on a much larger scale. In stead of relying on fundamental analysis, these traders spend more time on technical analysis (TA) of the market movement to run after abnormal returns (Crain's Detroit Business, Jan 2009). Given the bearish condition of current financial markets, it is interesting to ask questions such as why those proprietary trading firms outperform the rest of the financial industry? How they utilize TA into trading strategies to realize such large profits?

There is literature that finds theoretical support for the usefulness of technical analysis that past prices contain information for predicting future returns. For example, Brown and Jennings (1986) derive a two-period dynamic model of equilibrium to demonstrate that rational investors use historical prices in forming their demands. Technical analysis is found to have value in a model in which prices are not fully revealing and traders have rational conjectures about the relation between prices and signals. The general goal of TA is to identify regularities in the time series of prices by extracting nonlinear patterns from noisy data. Especially when market prices do not follow a random walk (Lo et al. 2000), TA is an effective means for extracting useful information from market performance. While technicians use various methods and tools, some extensively use indicators, which are typically mathematical transformations of price or volume. These indicators are used to help determine whether an asset is trending, and its price direction. In another theoretical model, Hong and Stein (1999) model a market populated by two groups of boundedly rational agents: "news watchers" and "momentum traders". Each news watcher observes some private information, but fails to extract other news watchers' information from prices. If information diffuses gradually across the population, prices under react in the short run. The under reaction means that the momentum traders can profit by trend-chasing. However, if they can only implement simple (i.e., univariate) strategies, their attempts at arbitrage must inevitably lead to overreaction at long horizons. Hong and Stein (1999) provide a unified account of under- and overreactions. TA could be viewed as a tool to capitalize on the under- and overreaction in prices.

Can high frequency traders make money out of technical analysis? Brad Barber (2005) states that 'heavy day traders earn gross profits, but their profits are not sufficient to cover transaction cost.' Despite this finding, he provided strong evidence of persistent ability for a relatively small group of day traders. It is possible that the winners of the current recessional market are also the small bunch. However, very few studies have been done to discuss technical analysis strategies applied by institutional investors. Also, not much research has focused on the strategies adopted by these not so well known winning proprietary trading firms.

Taking a close look at these institutional investors reveals two major categories: algorithmic traders and statistical arbitragers. As a trading strategy, statistical arbitrage is a heavily quantitative and computational approach to equity trading. It involves data mining and statistical methods; as well as automated trading systems. The first job advertisement in the Appendix exhibits the characteristics of statistical trading: ultra-high frequency quantitative trading, short term strategies, and top percentage payouts. Broadly speaking, statistical arbitrage is usually a strategy that is bottom-up, beta-neutral and uses statistical or econometric techniques in order to provide signals for execution. Signals are often generated through a contrarian mean-reversion principle, but can also be formed by lead / lag effects, extreme psychological barriers, corporate activity, as well as short-term momentum. Because of the large number of stocks involved and the high portfolio turnover, the strategy is implemented in automated fashion and great attention is placed on reducing trading costs.

Different from the common risk free arbitrage, statistical arbitrage does not come without risk. In the general sense, statistical arbitrage only is demonstrably correct as the amount of trading time approaches infinity as well as the liquidity. Over any finite period of time, a series of low probability events may occur leading to a shortage in liquidity available to the trader, default may even occur.

Statistical arbitrage is also subject to model weakness as well as stock specific risk (Barber et al. 2005; Lo et al. 2000). The statistical relationship on which the model is based may be spurious, or may break down due to changes in the distribution of returns on the underlying assets. Factors which the model may not be aware of having exposure to could become the significant drivers of price action in the markets (Lo et al. 2000). The existence of the investment based upon model itself may change the underlying relationship, particularly if enough entrants invest with similar principles. The exploitation of arbitrage opportunities themselves increases the efficiency of the market, thereby reducing the scope for arbitrage, so continual updating of model is necessary. On a stock-specific level, there is risk of M&A activity which would immediately end any historical relationship assumed from empirical statistical analysis.

The second job advertisement in the appendix provides information about the other group of institutional investors: algorithmic traders. They are characterized as the users of computer programs for entering trading orders with the computer algorithm deciding on certain aspects of the order such as the time, price, or even the final quantity of the order. Algorithmic trading is widely used by hedge funds, pension funds, mutual funds and other institutional traders to generate and execute orders automatically, in this context algorithmic trading can be classified between buy side and sell side institutions. In sell side algorithmic trading large trades are divided into several smaller trades in order to manage market impact, opportunity cost and risk (Economist.com, Feb 2006). Computer technology is employed to make decisions to initiate orders based on information that is received electronically, before human traders are even aware of the information. Different from statistical arbitragers, algorithmic traders focus on the execution of the trades (minimizing transaction costs) rather than on profiting from a trading strategy. For that reason, the benefit from a good algorithmic strategy does not need to out-weight the transaction fee. Algorithmic traders have to make the trade anyway, thus incurring the transaction cost.

Algorithmic trading may be used in any investment strategy, including market making, inter-market spreading, arbitrage, or pure speculation (including trend following). The investment decision and implementation may be augmented at any stage with algorithmic support or may operate completely automatically.

A third of all EU and US stock trades in 2006 were driven by automatic programs, or algorithms, and this figure should reach 50% by 2010. In 2006 at the London Stock Exchange, over 40% of all orders were entered by algorithmic traders, with 60% predicted for 2007. American markets and equity markets generally have a higher proportion of algorithmic trades than other markets, and estimates for 2008 range as high as an 80% proportion in some markets. Foreign exchange markets also have active algorithmic trading (about 25% of orders in 2006) (Timmons 2006). Futures and options markets are considered to be fairly easily integrated into algorithmic trading (Economist.com, Apr 2007), with about 20% of options volume expected to be computer generated by 2010 (Economist 383 June 2007). Bond markets are moving toward more access to algorithmic traders (The Wall Street Journal Europe, Apr 2007).

This paper is interested in how technical analysis tools such as moving average convergence – divergence and Bollinger bands, the two most popular instruments among practitioners would fit into algorithmic trading strategies and statistical arbitrage. This will be stated in more details in section two. One customized technical analysis strategy is created in the last part of section two to further explore possible trading rules for institutional investors. The profitability and risk and return analysis are also included in the same section. As a final point, summary of this study together with the future research perspective is presented in the last section.

1.2 Data

The Trades And Quotes (TAQ) database is employed in this paper for all analysis. This database contains second-by-second intraday transaction data for all securities listed on the US Exchanges, such as AMEX, NASDAQ and NYSE. In the paper, we used Jan 1st 2002 to Dec 31st 2006 price time series of SPY (listed on AMEX) traded between 9:30 am to 16:00 pm every valid trading day. The prices are filtered to remove mistakes. The filter rule adopted is from Tanompongphandh (2008). A price outlier which has 0.1% price deviation from the center moving average of the nearest 10 prices is removed from the dataset. In order to reduce computational intensity, we transform the second-by-second data into minute-by-minute format by taking the simple weighted average of every 60 seconds prices.

$$newprice_i = \sum_{t=61i-60}^{61i} price_t$$

This study focuses on exploring the algorithmic trading strategies and statistical arbitrage strategies that institutional investors would apply to meet their financial needs. For algorithmic trading strategies sections, a five-year continuous trading rule is adopted. The position will not be forced to be closed by the end of very trading day. However, we restrict the number of shares to exactly one at every position, which means once one long position is held, if the sell signal shows up, we will sell the share and simultaneously short one share to maintain the position of one. Consequently, when the next long signal shows up, the short position will be recovered and simultaneously enter into one long position. Successive long or short/sell action is not allowed in the algorithmic trading section. In the statistical trading strategy section, the five-year continuous trading rule is still valid while the restriction of exact one share position and non-successive long/short/sell are removed. The trading rule for statistical arbitrage will be explained in more details in its own section. The transaction cost is fixed as 3 cents for all strategies. It is comprised of a bid/ask spread of 1 cent and a commission of 2 cents for one round-trip. The bid/ask spread of 1 cent is adopted following Tanompongphandh (2008) who documents the distribution of

Figure 1 Time Series of SPY price from 2002 to 2006. Both a bearish and a bullish trend are displayed in this time series plot. Before 2003, the market displayed the downward trend (bear market) and after 2003 the market showed an upward trend (bull market).

bid/ask spreads for our period for the S&P 500 ETF. The 2 cents commission is the fee for a round-trip trade with Interactive Brokers.

Figure 1 illustrates the entire dataset employed in this study. SAS is trusted in this study to perform algorithmic trading strategies and generate relevant figures. For the statistical arbitrage section, C++, the most widely used language in the industry, is employed to carry out the strategy, and associated figures are produced in Matlab.

All codes will be enclosed in the appendix.

2 TECHNICAL ANALYSIS

2.1 Moving Average Convergence – Divergence (MACD)

2.1.1 *Background*

Created by Gerald Appel, the MACD, one of the most popular technical analysis tools used by the financial world, is a trend-following momentum indicator which captures the change in momentum, crossover signals of new trends thresholds and measures the rates of ascent or descent and works best in wide-swing trading markets.

Three steps involved to form the MACD and its signal (Kaufman 2003). Firstly, one faster and one slower smoothed trendline from the original price series are derived, the MACD is generated by subtracting the slower one from the faster one, and finally another moving average of MACD is superimposed over the MACD as the signal line. MACD is what's known as a centered oscillator. In other words, the MACD fluctuates above and below a centering line. These types of oscillators are good for identifying strength or weakness or direction of momentum behind a security's move.

When MACD is below the signal line, a bullish mode is established, the market is considered overbought, and when MACD is above the signal line, a bearish mode is formed then the market is believed oversold. As MACD is rising, it is telling us that the gap between the fast and slow moving average is widening, therefore indicating the bullish momentum is increasing. As MACD crosses down below the signal line, it is known that the fast moving average has crossed down below the slow moving average and as it continues in its downward path, the distance between the two moving average is widening therefore bearish sentiment is increasing. The momentum is strengthening in the downward trend as long as MACD line is in descent. Figure 2 provides an example of the construction of MACD.

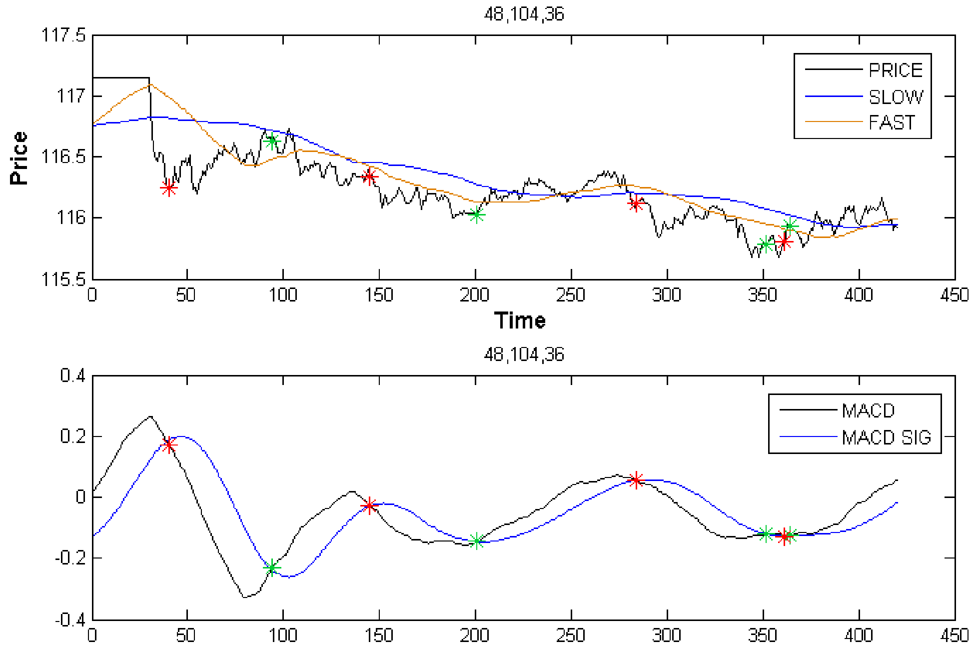

Figure 2 Demonstration of MACD with SPY minute-by-minute data. In the upper panel, price line is marked in black and the brownish curve represent the fast mood moving average, while the blue curve is the slow mood moving average. The lower panel presents the MACD and its signal line. The red and green stars symbolize the sell and long signals respectively.

Some investors also use zero line with MACD to identify trading signals. The standard interpretation is to buy when the MACD crosses up through the zero line, or sell when it crosses down through the signal line. Similar to the interpretation to the signal line, the crossing of the MACD up through zero is consider bullish while down through as bearish.

The third type of MACD signal occurs when the shorter MACD average rises or falls dramatically compared to the longer moving average, causing a sharply widening gap between the MACD line and the signal line. When the difference between these lines becomes more extreme, it suggests a price move has become overextended and is subject to pulling back or correcting itself.

Another powerful tool in the MACD is the creation of divergences between indicator and the price trend. If the prices have made a low but the difference between MACD and signal has not, it suggests that not as much pressure to push prices lower as there was earlier. This provides early evidence that the market cold move into an uptrend. Figure 3 and figure 4 demonstrate the convergence and divergence examples using SPY data.

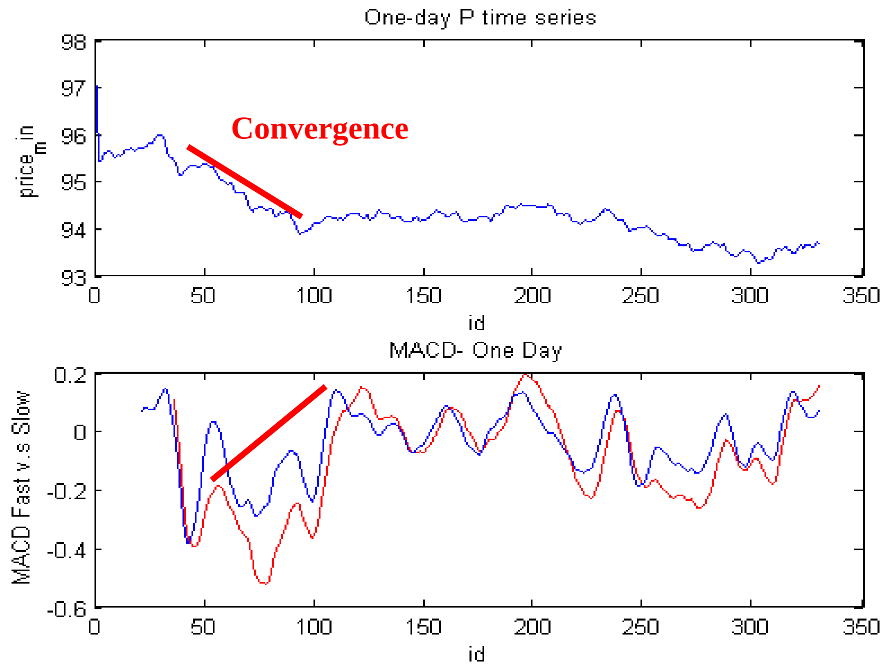

Figure 3 Demonstration of Convergence in MACD with SPY minute-by-minute data.

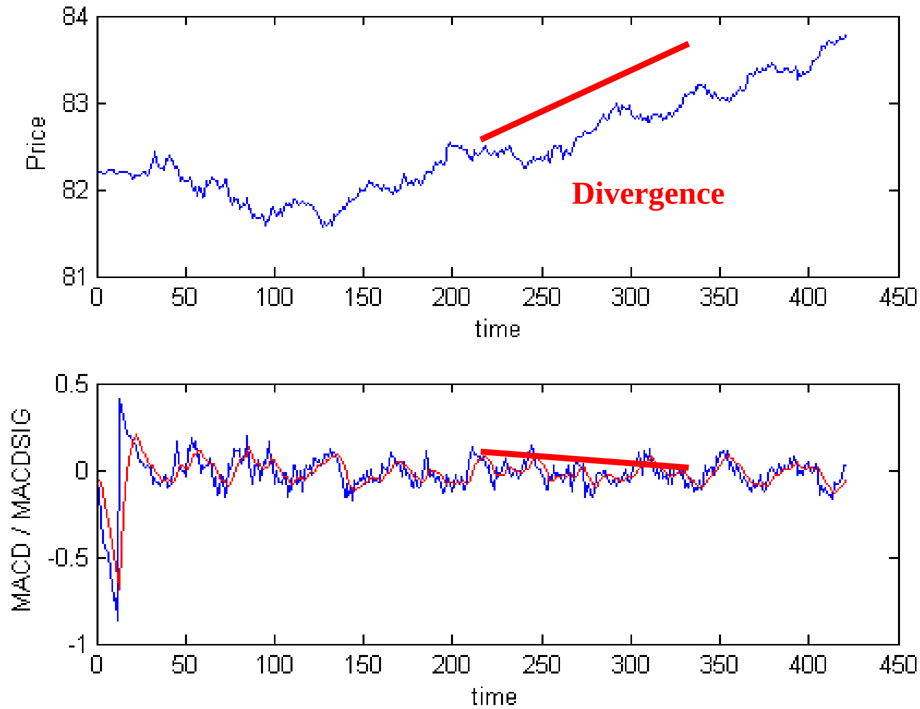

Figure 4 Demonstration of Divergence in MACD with SPY minute-by-minute data.

In terms of the calculation of moving average, simple moving average (SMA) and exponential moving average (EMA) are most widely applied. Simple moving average gives equal weights to prices at different times no matter it is the latest or the earliest. Exponential weight moving average weight the latest price most heavily and the weight decrease in an exponential manner as the price becomes less close to the current one.

Murphy (1999) suggested looking at a MACD on weekly scale before looking at a daily scale to avoid making short term trades against the direction of the intermediate trend. However, in our case, min-by-min price data is used, so some modifications should be adopted. In the next section, the parameter specifications and moving average calculation used in the MACD study will be presented.

2.1.2 *Methodology*

2.1.2.1 Parameter Specification and Calculation Approach

The MACD used in the analysis is created as follows:

Step 1. Choose the two calculation periods. Slow period with parameter S = 26 minutes increasing to 130 minutes with step length equal to 26 minutes; fast period with F = 12 minutes increasing to 60 minutes with step length equal to 12 minutes.

Step 2. Choose from simple weighted moving average and exponential weighted moving average approach to generate trendlines.

Step 3. Calculate MACD as the fast trendline minus the slow trendline. When the market is moving up quickly, the fast smoothing line will always be above the slow one, and the difference between the two will be positive. When prices go up, MACD line will go up by definition.

Step 4. The signal line is an M-day moving average of the MACD line. In the study, M = 9 minutes increasing to 45 minutes with step length equal to 9 minutes.

Figure 2 demonstrates the simple weighted moving average approach (SMA) which is simply the average of prices of the index over a specific time span. SMA is calculated for each trading minute for the previous period, and at the end of a trading minute, the last minute is added while the earliest minute is dropped.

$$\text{SMA(mins)}=\sum_{k=1}^{\text{mins}}\frac{\text{Price}_k}{\text{mins}}$$

The problem with SMA is that the earliest day of the time period has the same weighted in the average as the most rent day. It is also sensitive to the volatility of the market which may induce many false signals. If the earliest day was volatile, but the market has recently calmed, then the volatile day will have a large influence on the average which would not best represent the current market. Thus SMA based on shorter time spans with little volatility reflects the underlying current trend more appropriately, and it loses power when the time spans expanses.

In order to correct the anomaly generated by SMA, exponential weighted moving average (EMA) is also explored, where greater weight is given more recent prices. This greater weight causes the EMA to follow the underlying prices more closely most of the time than the SMA of the same duration.

EMA could be calculated in many different ways. In this study, we trust the traditional method of calculating EMA by adding an additional day to the simple moving average, but give greater weight to the last day. For a M-minute moving average, the formula to calculate the weight of the last day is:

$$\text{Weight}_{\text{current}} = 2/(M+1)$$

Since the sum of all of the weights must equal 100%, the weights of the preceding M minutes must equal:

$$\text{Weight}_{\text{MA}} = 100\% - \text{Weight}_{\text{current}}$$

Hence, the formula for calculating the exponential moving average is:

$$\text{EMA} = \text{last day weight} \times \text{last day price} + \text{weight of previous EMA} \times \text{previous EMA}$$

As stated earlier, MA can be calculated in many different ways, and, likewise, can be used in many different ways. The best use of moving average is in determining trends. The greater the slope of the moving average, the greater the strength of the trend. Generally, traders will choose a time period that is suitable to their investment time frame. Since we are dealing high frequency index price data and one of the research purposes is to figure out the optimal parameter set that would generate the optimal profit and return for given risk level, the comparison of parameter sets and the difference between SMA and EMA approach will be presented in the results section.

2.1.2.2 Trading Rules

As described in the background section, trading signals are identified with the MACD indicator in a number of ways:

Trigger line signal: Some argue that a buy signal is generated when the MACD crosses and goes above zero, and that a sell signal is generated when the MACD crosses and goes below zero.

Crossover signal: Others interpret a 'crossover' as a signal: when the MACD crosses and falls below its moving average, a sell signal is generated, and when the MACD crosses and rises above its moving average, a buy signal is generated.

Divergence signal: When the MACD does not follow the current trend and moves counter to the direction of the corresponding index price, then this is interpreted as a warning that the price trend may change. Hence, when the MACD is moving down while the price is still rising, then this may be interpreted as a sell signal. A strong sell signal occurs when the MACD is reaching new lows while the corresponding price is still moving up. Conversely, when the MACD is moving up while the price is still falling, then this could be interpreted as a buy signal. A strong buy signal occurs when the MACD is reaching new highs while the corresponding price is still going down.

Given these different signal choices, we focus on *Crossover signal* to trigger buy and sell transactions. A buy signal is generated when the MACD line crosses over and rises above the MACD moving average line; a sell signal is generated when the MACD line crosses over and falls below the MACD moving average line. If successive buy or sell signals appear, only the first buy or sell signal will trigger the action, the later same signal will be ignored. According to the one share position only restriction, two simultaneous transactions will occur at one buy or sell signal. If a long position is at hand when a sell signal comes, the long position will be closed and at the same time a short position is generated. Similarly for a short position, when a long signal comes, the short position will be closed and meanwhile a long position is set up.

2.1.3 *Results*

2.1.3.1 Parameter Test

Algorithmic traders who are interested in the optimal execution of trades will seek any edge in terms of the timing of their trades. Since they have to trade by design, they will have to incur the transaction cost whether they use MACD or not. Therefore, there is a difference between patterns that can be used for algorithmic trading, versus those that can be used for trading strategies. A pattern might not present a profit opportunity in the absence of other motives to trade, since strategies that attempt to take advantage of it lose money, after paying the transaction costs. However, algorithmic traders who have other exogenous motives for trading might wish to exploit these patterns. Thus, we first tested the sensitivity of raw profit before transaction cost to slow, fast and signal moving average parameters to reveal a potential optimal parameter set that would suit algorithmic traders. Figure 5 and 6 reveal the relationship between raw profits and fast, slow parameters. When both parameters are small, the shape of the distribution of raw profit looks like a circle. And the raw profits increase with the increase of both parameters. Raw profits maintain unchanged when both parameters are large.

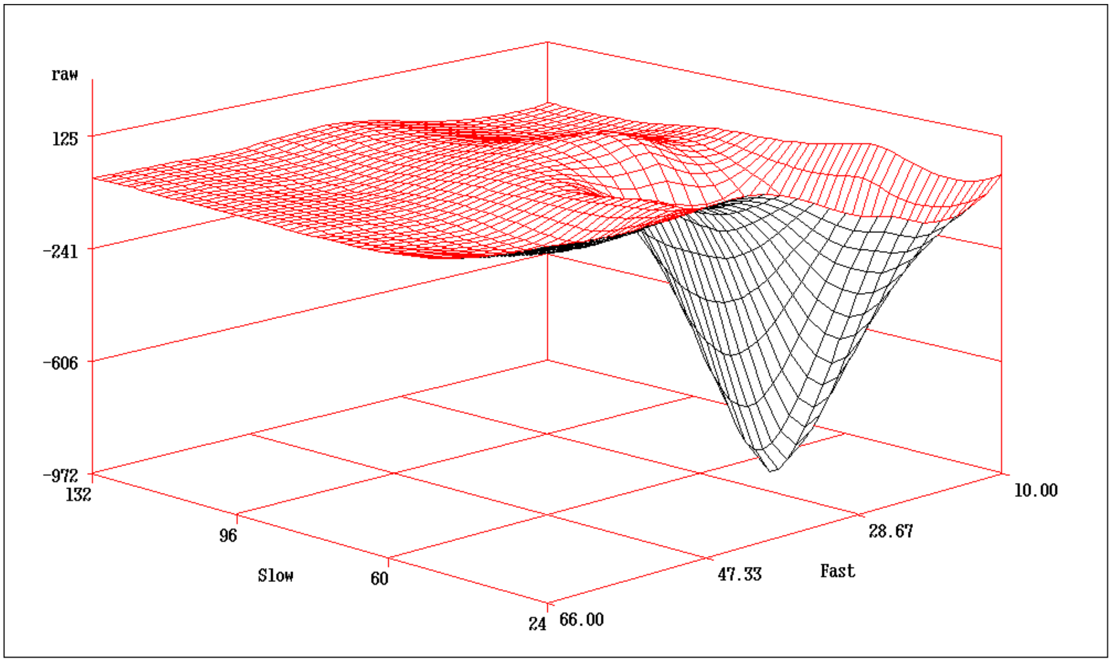

Figure 5 3D sensitivity of gross profit to slow and fast parameters plain.

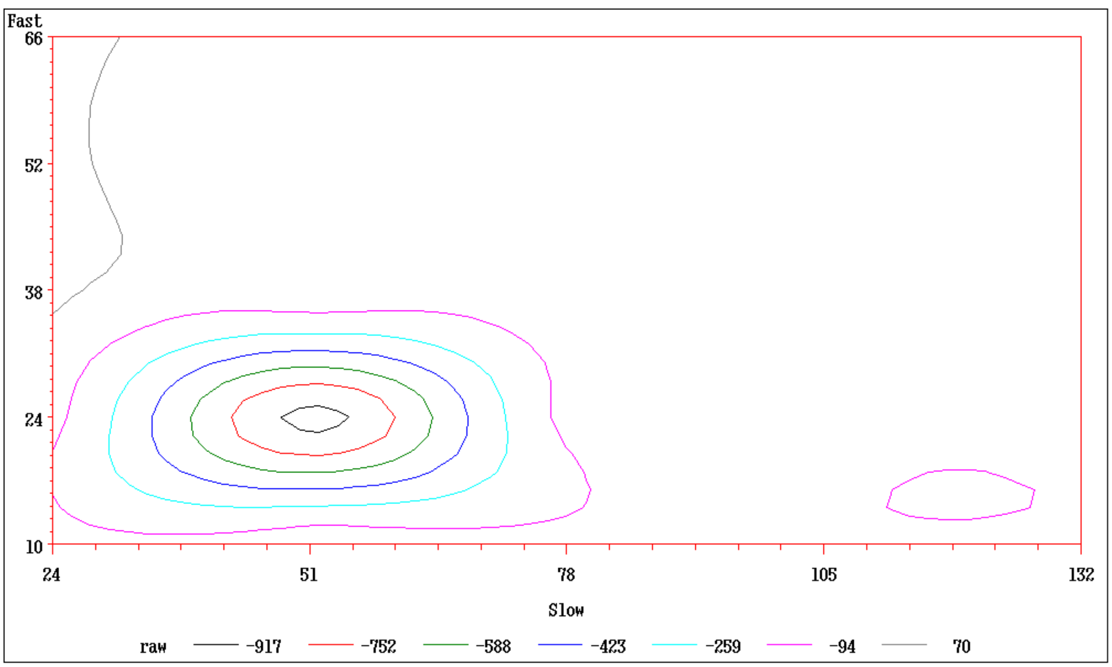

Figure 6 Contour of the sensitivity of gross profit to slow and fast parameters plain.

In figure 7 and 8, the relationship between raw profit and signal, fast parameters are very complex. No monotonic relationship holds between signal and raw profit. But raw profit increases with the increase of fast signal for a given signal parameter.

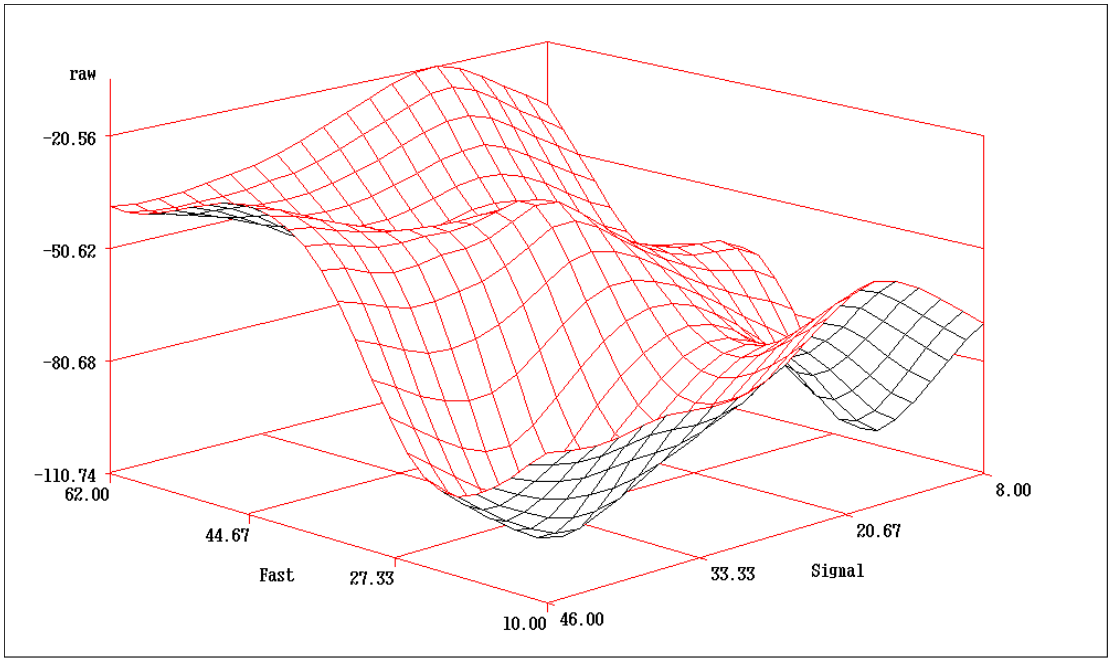

Figure 7 3D sensitivity of gross profit to fast and signal parameters plain.

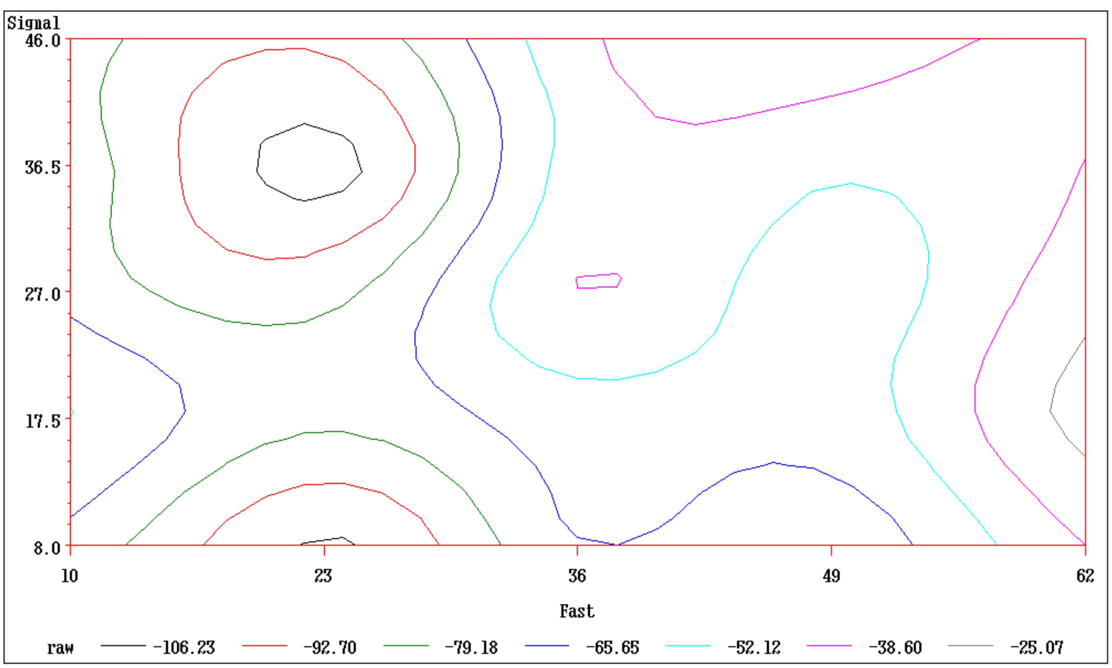

Figure 8 Contour of the sensitivity of gross profit to fast and signal parameters plain.

Based on figure 9 and 10 which state the relationship between raw profit and slow, signal parameters, the raw profit becomes inelastic to the change of slow for a given signal parameter, and for a known slow parameter, the raw profit increases as the signal parameter increases.

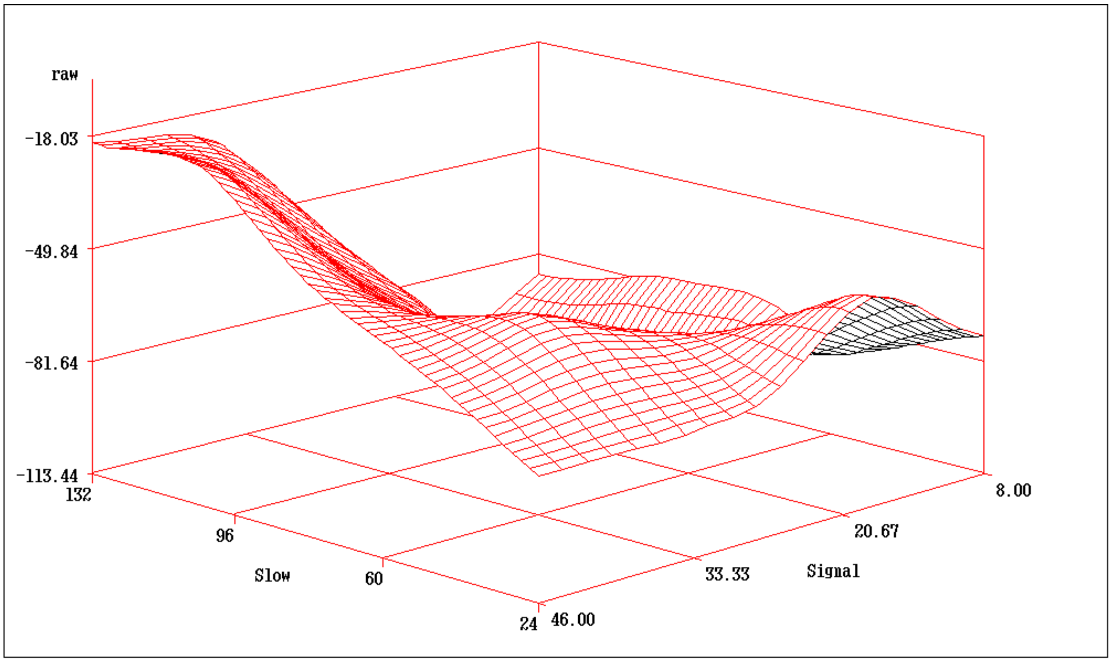

Figure 9 sensitivity of gross profit to slow and signal parameters plain.

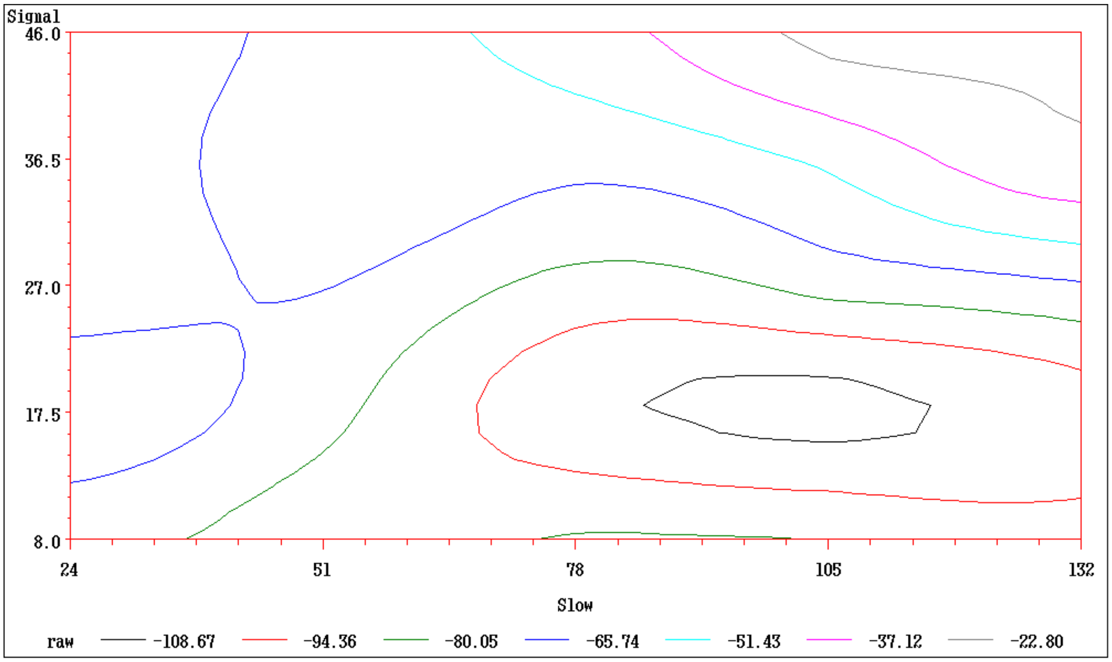

Figure 10 Contour of the sensitivity of gross profit to slow and signal parameters plain.

Taking transaction cost into account, the stories is quite different but to some extent, simpler. Figure 11 and 12 demonstrate the relationship between net profit with fast and slow parameter. When slow is small (between 26 and 52), a clear bimodal distribution of net profit displays with respect to the fast parameter. As slow is increasing, the bimodal shape disappears, net profit becomes inelastic to the fast parameters and perfectly elastic to the slow parameters.

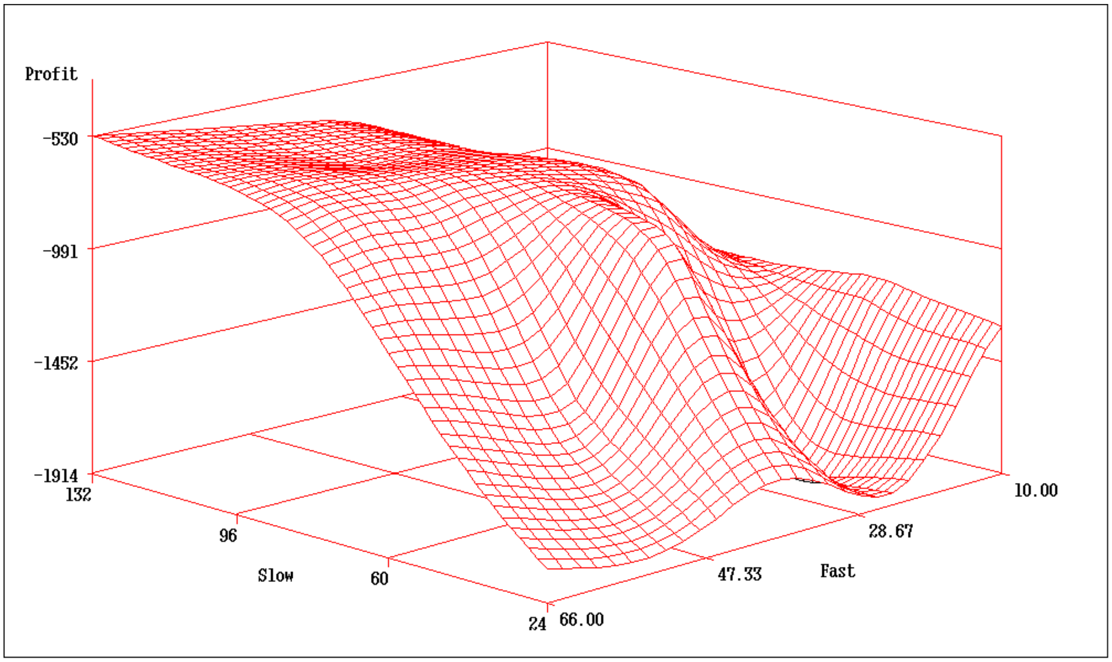

Figure 11 3D sensitivity of net profit to slow and fast parameters plain.

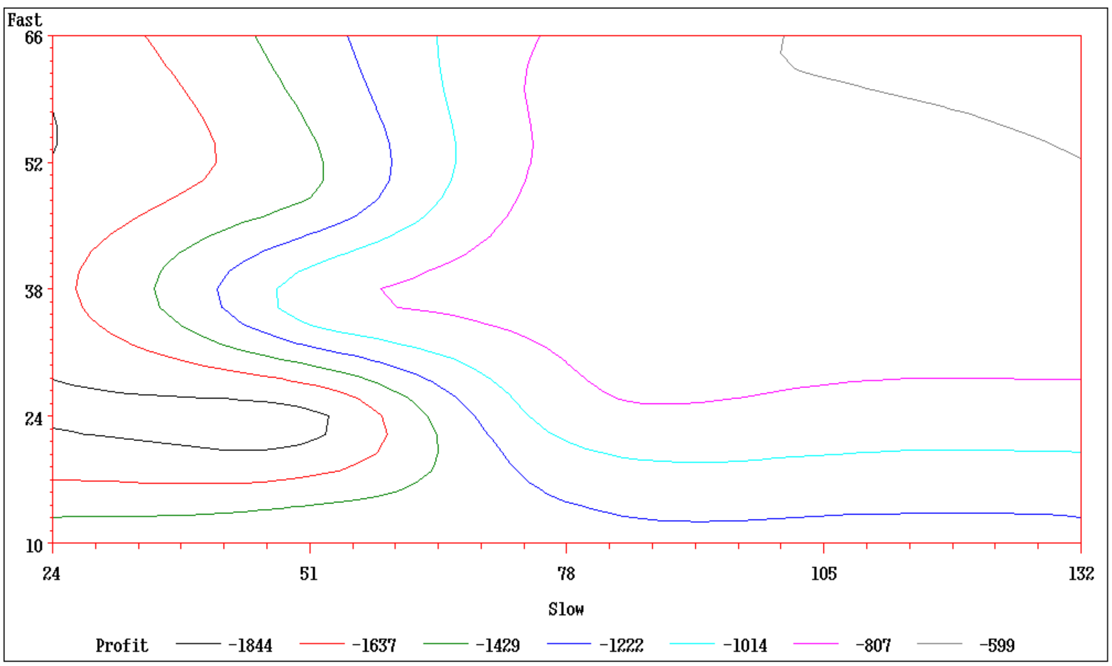

Figure 12 Contour of the sensitivity of net profit to slow and fast parameters plain.

Figure 13 and 14 demonstrate the relationship between the net profit, fast and signal parameters. The net profit presents a clear bimodal shape with respect to the fast parameter and increases as the increase of the signal parameter.

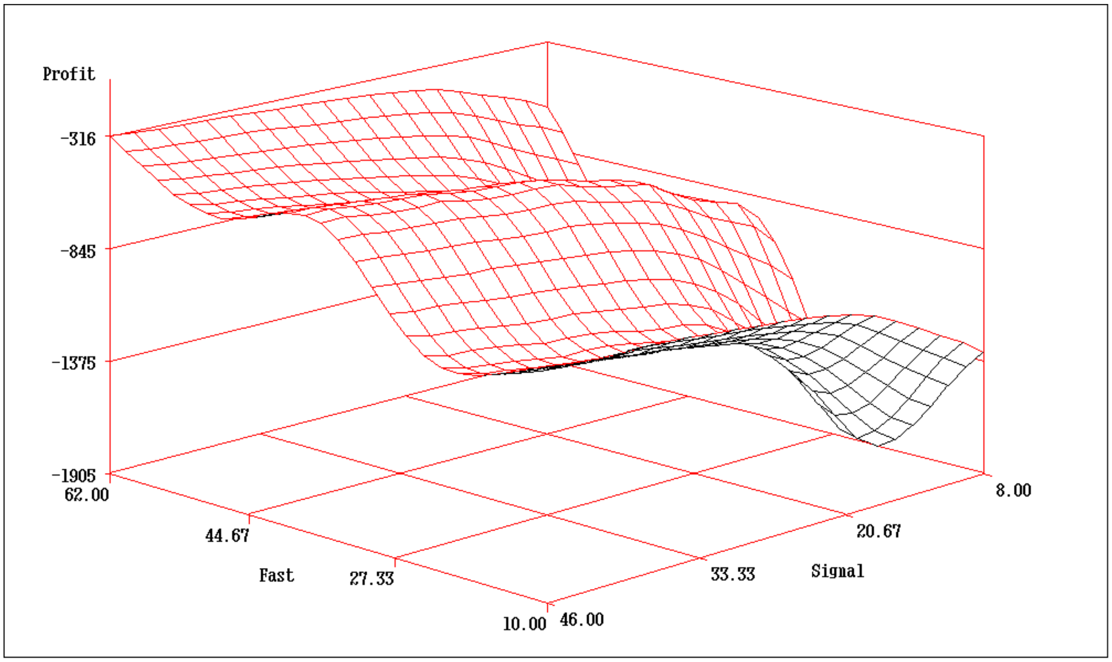

Figure 13 3D sensitivity of net profit to signal and fast parameters plain.

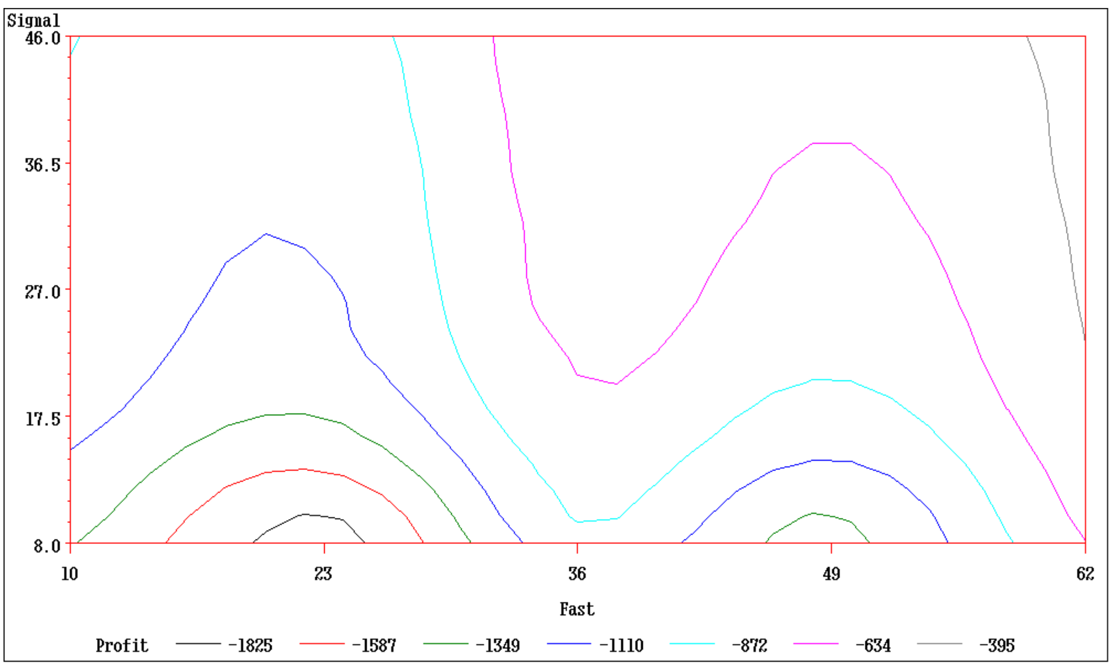

Figure 14 Contour of the sensitivity of net profit to signal and fast parameters plain.

From figure 15 and 16, we are able to obtain information of the relationship among the net profit, signal and slow parameters. When the parameter gets larger, the net profit becomes elastic to the signal parameter, and turns more and more inelastic to the slow parameter. We also observe the growing trend of raw profit with the raise of the signal parameter for a given slow parameter.

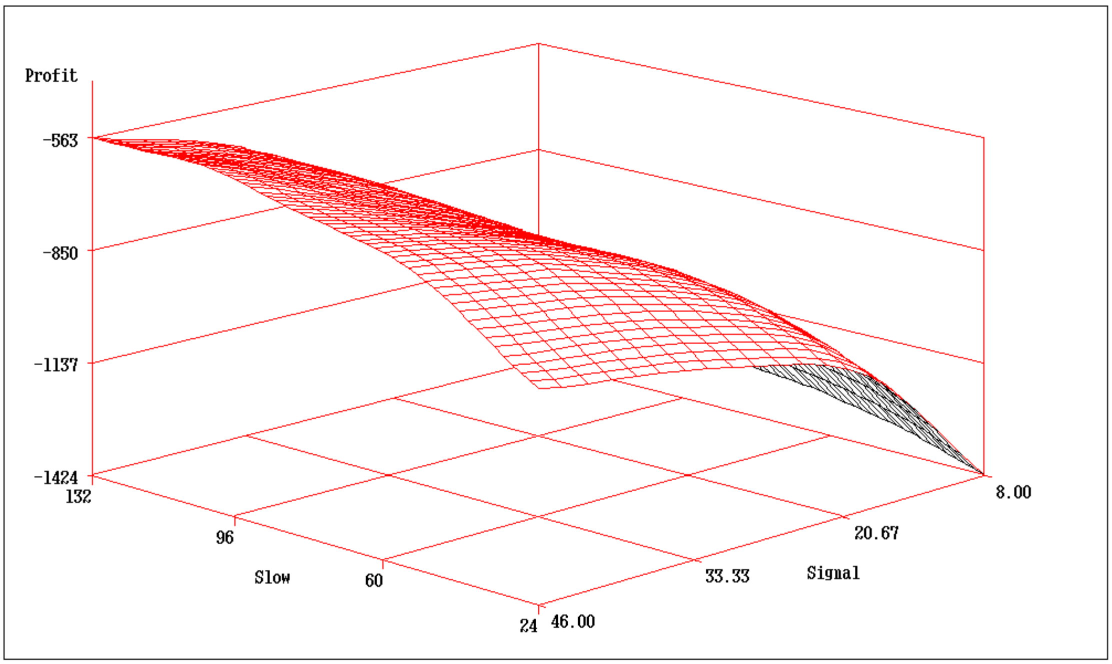

Figure 15 3D sensitivity of net profit to slow and signal parameters plain.

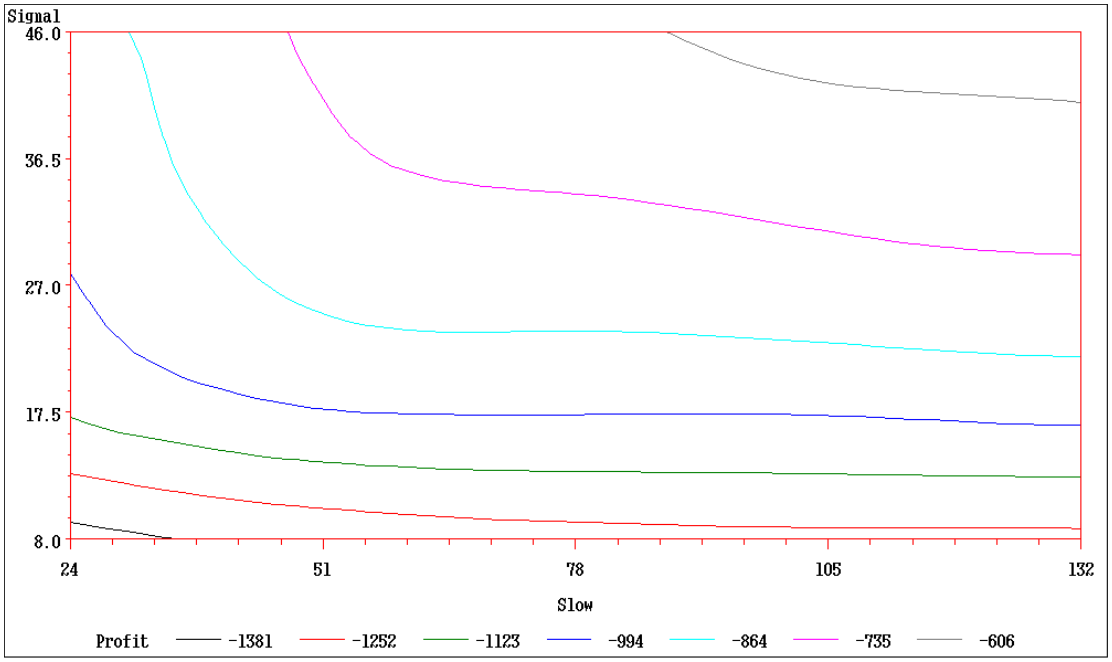

Figure 16 Contour of the sensitivity of net profit to slow and signal parameters plain.

2.1.3.2 Comparison between EMA and SMA

Exponential weighted moving average is employed in this study giving its popularity in the financial industry. Due to the high frequency nature of the data, even large MACD parameters fail to generate reasonable number of transactions per day which result in huge transaction cost that erodes the small profits. Since simple moving average usually generates slower signals than EMA due to the heavier weights on earlier price, SMA is also tested to generate fast, slow trendline as well as the signal MACD line.

The parameter test result for SMA is illustrated in figure 17. The transaction counts decreases as the lag number of parameters increases. For one particular day, the trades initiated by (60, 130, 45) are much fewer than (24, 52, 18). The issue with large lag number reside in the ability to capture the trading opportunity in very fluctuate market.

Those small market ups and downs would be ignored by the relative large smooth parameters and investors might lose potential profitable chances.

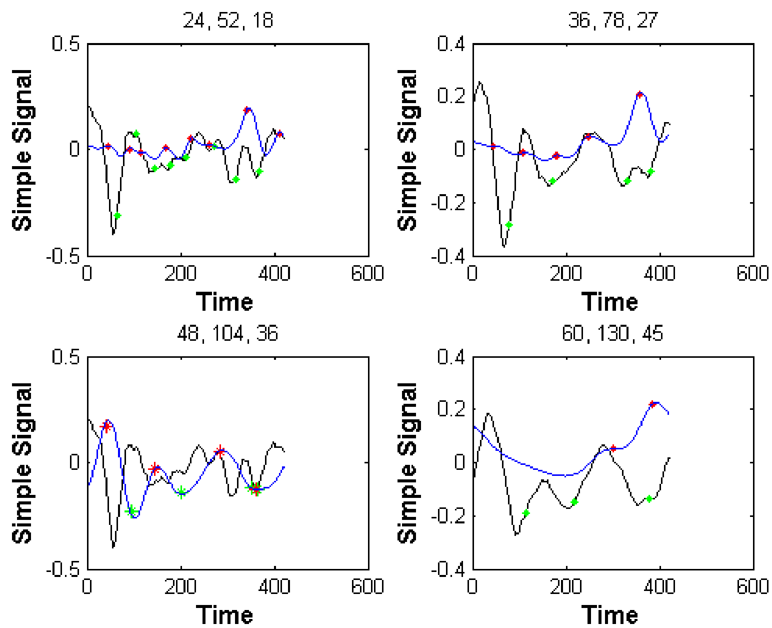

Figure 17 the transaction counts for SMA in terms of different parameter specifications.

In figure 18, the transaction schedule of both SMA and EMA are displayed in the upper panel and lower panel separately. From naked eye, we can observe a much busier trading planner for SMA method. Also for the same lag number set (48, 104, 36), SMA MACD and MACD signal line appear to be more smooth than the ones by EMA. This is also a common phenomenon for other set of lag parameters. Fewer transactions lead to less cost, however, SMA MACD fails to get rid of negative profit. Even though it smoothed the transaction schedule, it is not able to capture the optimal timing to generate profit.

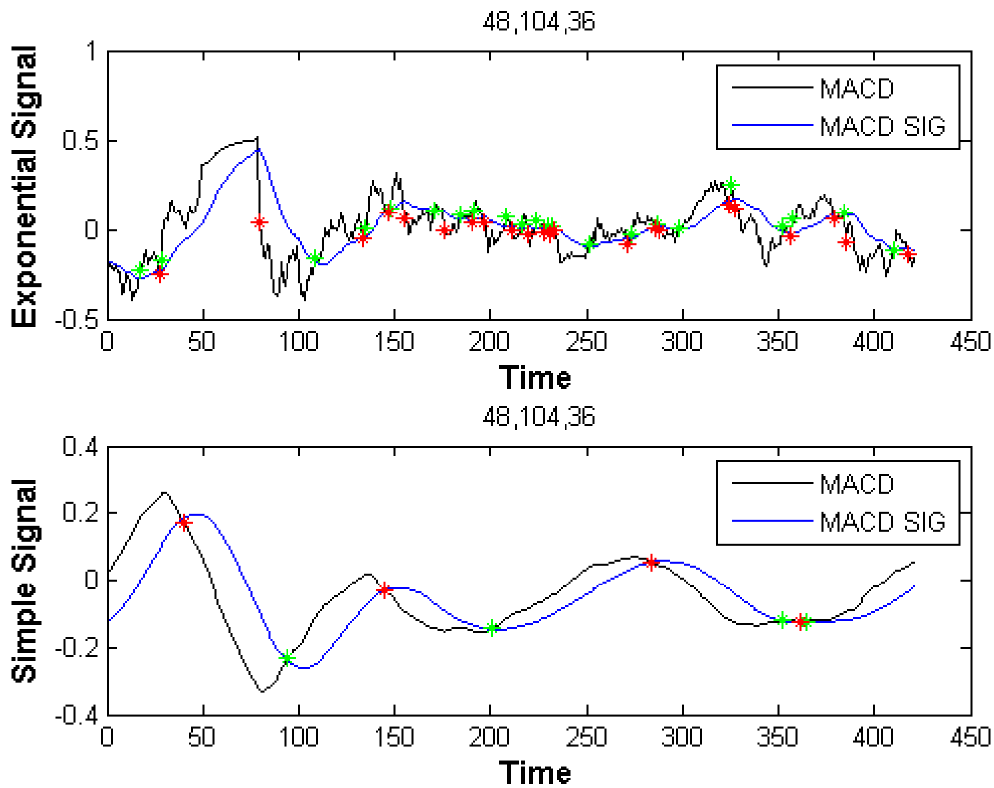

Figure 18 Transaction counts comparison between EWA and SMA. EWA is displayed in the upper panel and SWA is shown in the lower panel.

2.1.4 *Summary*

For statistical arbitragers, apparently MACD is not a good choice since taking the transaction cost into account; no positive profit is generated for any parameters set. For algorithmic traders, where the analysis should be done without taking into account the transaction cost, MACD also fails to provide any positive raw profit. Trade on this MACD strategy will turn out to be a significant loss for this one share position.

The relationship among profit, raw profit and three parameters is of great complexity. Generally, profit and raw profit both increase with the increase of lag parameters even though not monotonically. Varying the lag parameters could not guarantee positive absolute profit either in raw returns or after transaction cost, thus for both algorithmic traders and statistical arbitragers, solely looking at lag parameters would not help make investment decisions to generate positive returns.

Even though EMA is more popular among practitioners for MACD, for the high frequency trading data used in this study, it seems to be less powerful compared to SMA. EMA produces many more transactions and induces much higher transaction cost which worsens the associated net profit and return.

2.2 Bollinger Bands

2.2.1 *Background*

John Bollinger invented another very popular technical analysis tool in the 1980s, which is then named after him as ‘Bollinger Bands’. The purpose of Bollinger Bands is to provide a relative definition of high and low price to previous trades.

Bollinger Bands consist of three major bands:

- A middle band being an N-period simple moving average
- An upper band at K times an N-period standard deviation above the middle band
- A lower band at K times an N-period standard deviation below the middle band

The most popularly values for N and K used in the financial industry are 20 days and 2 standard deviations respectively (www.bollingeronbollingerbands.com). The default choice for the average is a simple moving average, but other types of averages such as exponential moving average can be employed if preferred. Usually the same lag number is used for both deriving the middle band and calculating standard deviation.

Many traders use Bollinger Bands to derive %b and Bandwidth, two major indicators based on the bands. %b, used for indentifying the position in relation to the bands, is rooted in the formula for stochastic oscillator which was introduced by George Lane to compare to closing price of a commodity to its price range over a given time span (www.bollingeronbollingerbands.com). %b is defined as follow:

$$\%b = (\text{last - lower BB}) / (\text{upper BB - lower BB})$$

%b equals 1 at the upper band and 0 at the lower band. This indicator is widely used for system building and pattern recognition.

Bandwidth offers an insight of how wide the Bollinger Bands are on a normalized basis (www.bollingeronbollingerbands.com). It is defined as follow:

$$\text{Bandwidth} = (\text{upper BB - lower BB}) / \text{middle BB}$$

Using the default parameters of a 20 - lag look back and plus/minus two standard deviations, Bandwidth is equal to four times the 20 - lag coefficient of variation. The coefficient of variation is defined as result of standard deviation of a 20 - lag index prices divided by mean of the same period index prices.

$$\text{Coefficient of Variation} = \frac{\text{STD(index)}}{\text{Mean(index)}}$$

Bandwidth is very useful for traders to identify price trends and trade opportunities arising from relative extremes in volatility. In our study, we will simply employ the number of standard deviation to represent the bandwidth.

Different traders use Bollinger Bands indicators differently. Some buy when price touches the lower BB and exit when price touches the moving average in the center of the bands. Others buy when price breaks above the upper BB or sell when price falls below the lower BB. Moreover, the use of BB is not confined to stock traders; option traders, most notably implied volatility traders, often sell options when BB are historically far apart or buy options when the BB are historically close together, in both instances, expecting volatility to revert back towards the average historical volatility level for the stock.

When the bands lie close together a period of low volatility in stock price is indicated. When they are far apart a period of high volatility in price is indicated. When the bands have only a slight slope and lie approximately parallel for an extended time the price of a stock will be found to oscillate up and down between the bands as though in a channel.

As always, traders are inclined to use BB with other indicators to see if there is confirmation. In particular, the use of an oscillator like BB will often be coupled with a non-oscillator indicator like chart patterns or a trendline. If these indicators confirm the recommendation of the BB, the trader will have greater evidence that what the bands forecast is correct.

In this section, we will focus on the sensitivity of profit to parameters of bandwidth and numbers of periods used to calculate bands with respect to algorithmic trading (before transaction costs) and statistical arbitrage (after transaction costs).

2.2.2 *Methodology*

Accompanying with the default 20-period simple moving average middle band and ±2 STD upper / lower bands, we extended the period number from 21 minutes to 130 minutes with step length equals to 1 minute, and ±1.5 STD, ±2.5 STD are also included in the test. The total specifications finally arrive at 3×110 = 330 sets.

Similar to the trading restriction applied in the MACD strategy, for trades using Bollinger Band, we also limit the position to be exact one share and no successive buy or short/ sell is allowed. Transaction cost is fixed at 3 cents each deal as well. Figure 19 illustrates when the transaction will be triggered based on Bollinger Bands. A long/ buy signal (in green) appears when the original price line comes back in the lower band after crossing over, and a short/sell signal(in red) appears when the original price line comes back in the upper band after crossing over. If the price line just crosses over the upper / lower band but never comes back in, no transaction signal is generated. The long/ buy signal always stay on or around the lower band by nature while the short/ sell signal always stay on or around the upper band.

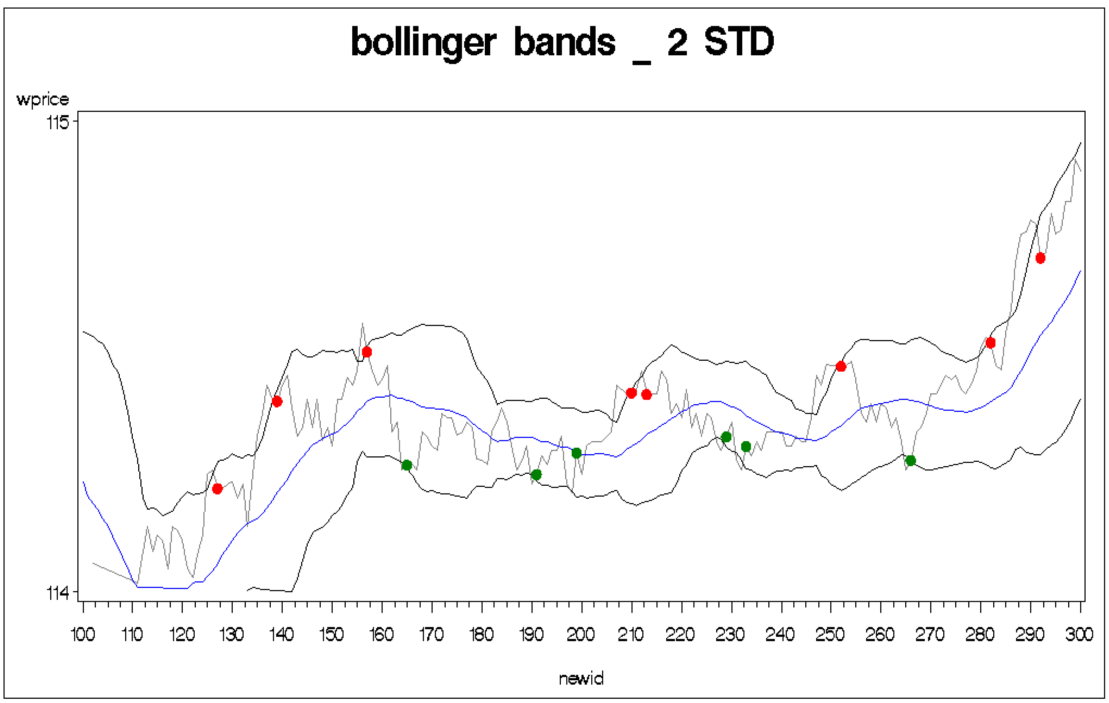

Figure 19 Demonstration of Bollinger Bands. The middle grey line is the original price. The middle blue line is the N-period moving average of price line. The upper and lower black lines are the plus and minus 2 standard deviations of the blue line. Red points represent the short / sell signals and green points represent the long signals.

2.2.3 *Results*

2.2.3.1 Bandwidth

Figure 20 and 21 provide examples of trading conditions under ±1.5 and ±2.5 STD Bandwidth with 21- period. From the naked eye observation, it is not difficult to conclude that the wider the band, the less transaction incurred and less transaction cost induced.

Figure 22 plots the net profit after transaction cost and helps statistical arbitragers to make decision. The story for these arbitragers is relative simple. The profit is monotonically increasing with increasing band width. For those who are seeking for abnormal profit, wider Bollinger bands might meet their needs. However, the severe problem here is that none positive profit is resulted from the BB strategy after taking into account of transaction fee. Purely looking at BB strategy will not necessarily generate exciting abnormal positive results.

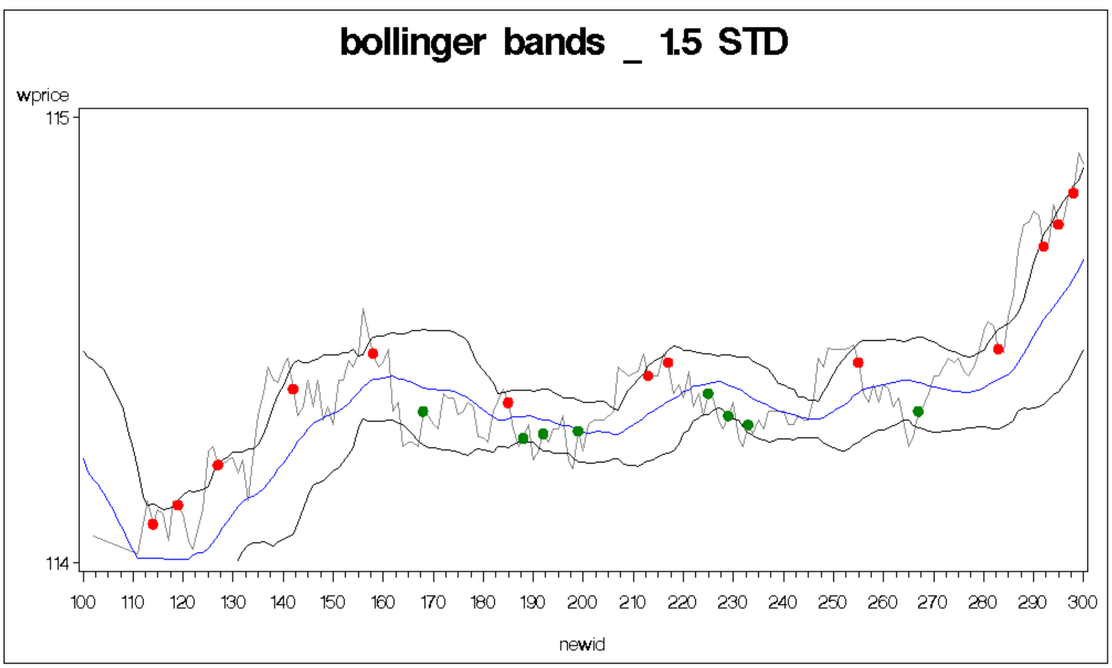

Figure 20 Demonstration of Bollinger Bands with plus and minus 1.5 STD.

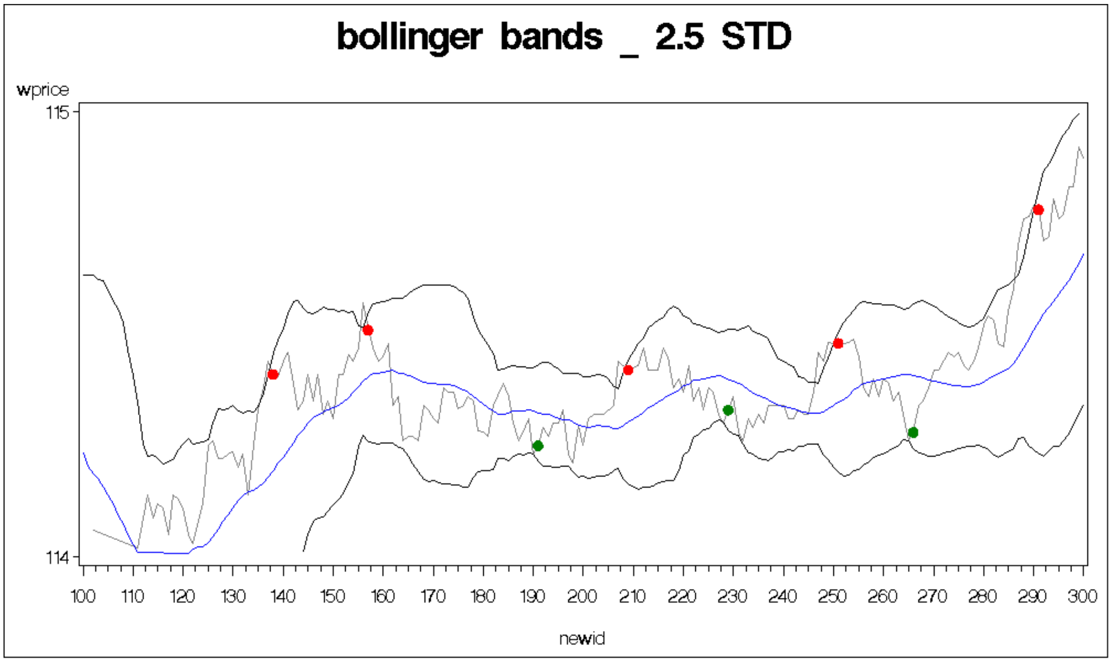

Figure 21 Demonstration of Bollinger Bands with plus and minus 2.5 STD.

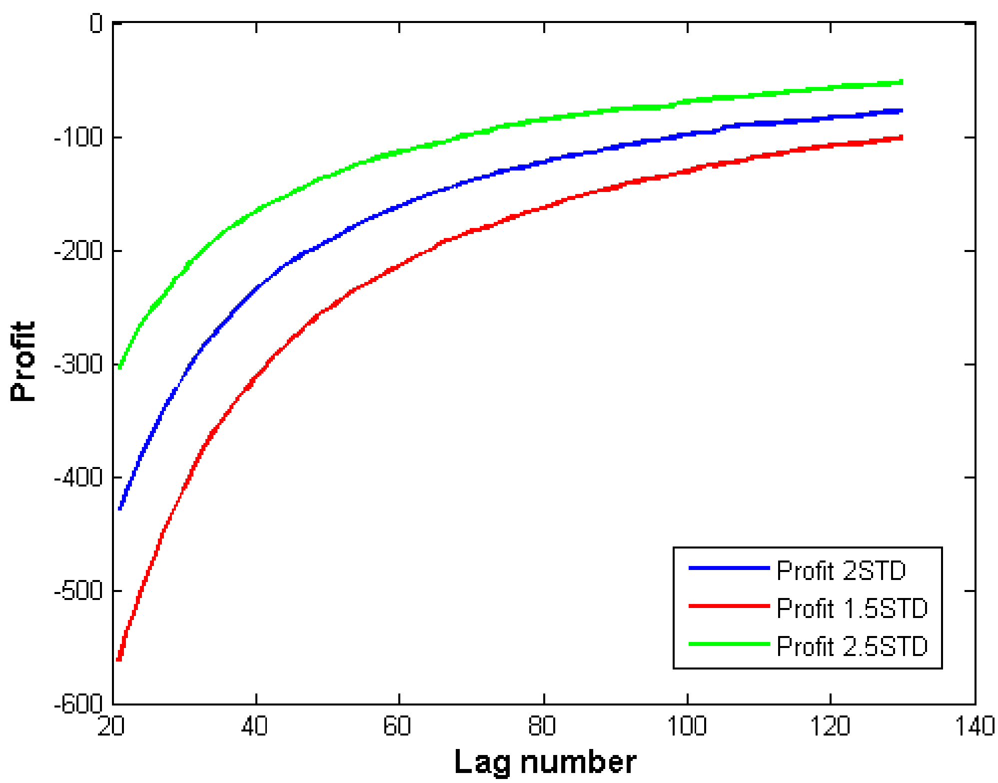

Figure 22 Relationship between net profits and lag numbers as well as bandwidth.

For algorithmic traders, the story here is very different. Looking at Figure 23 which graphs raw profit before transaction cost versus band width and lag numbers, the structure of raw profit for different band width is much more complicated. No apparent monotonic relation holds here. Even though positive raw profit is observed, the magnitude is too small for the 5 years holding period. The return is eroded by the time. Without considering lag number, it is powerless to conclude what band width would outperform others.

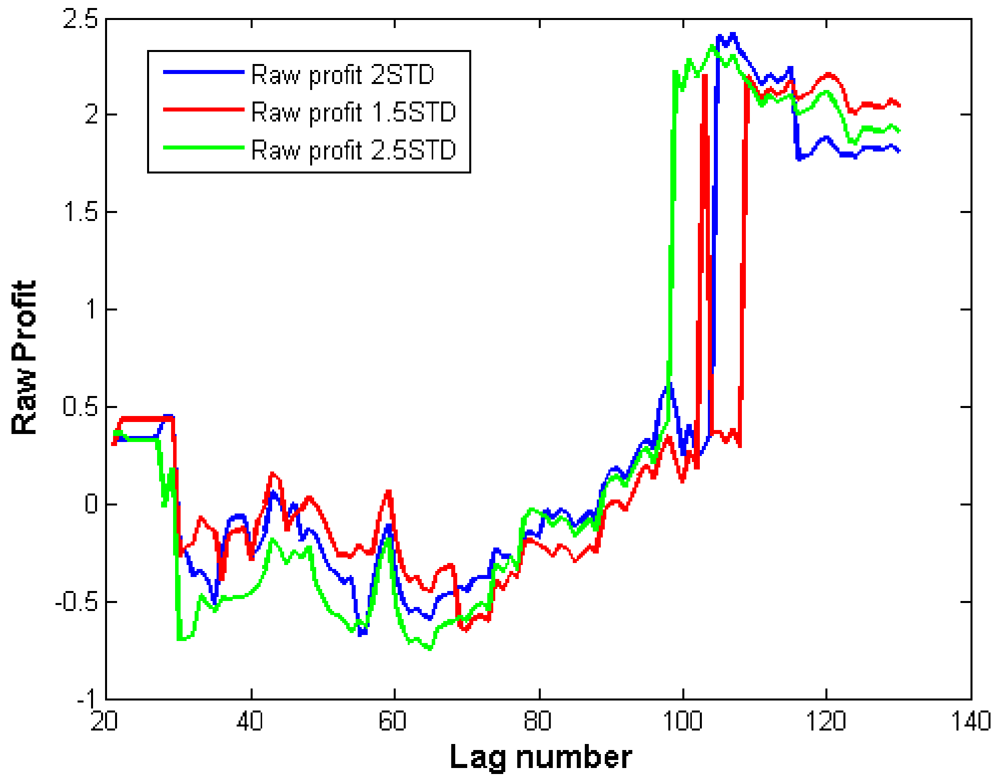

Figure 23 Relationship between gross profits and lag numbers as well as bandwidth.

Figure 24 generates the histogram chart for raw profit and net profit for different band widths. The left panels display the raw profit before transaction cost. The mean raw profit ranges from 0.37898 for ±1.5 STD to 0.42562 for ± 2.5 STD. one may argue that BB generates positive return in general, however, given the 5 years holding period, the mean index price is above 100 dollar, this less than 50 cents mean raw profit would hardly be attractive to institutional investors. Additionally, the distribution of the raw profit for one certain band width is not uniform but bimodal. Two distinct modes stand at raw profit = 0 and 2 and the weight of the negative peak is much larger than the positive one which indicate that for different lag numbers, regardless the band width, it is more likely to generate zero and negative raw profit.

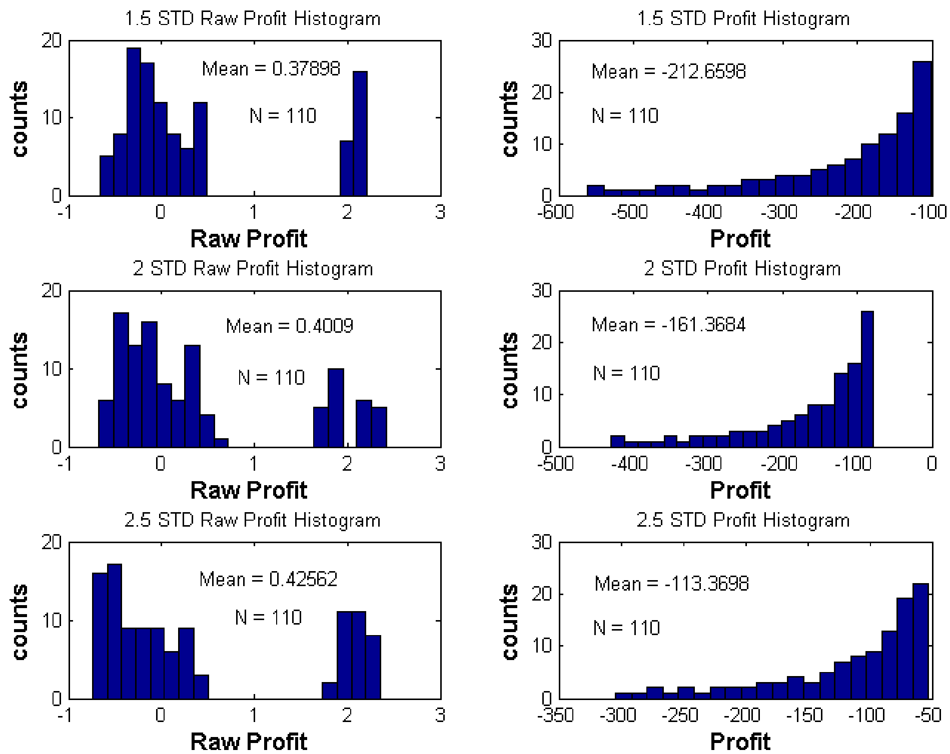

Figure 24 Histogram for net profit and gross profit. The left panels display the gross profit distributions in term of different bandwidth. The right panels display the net profit distributions in term of different bandwidth.

Taking the 3 cents transaction fee into account, the net profit turns into totally negative values as shown on the right panels of Figure 24. The choice of various band widths will not be able to generate positive profits. Figure 25 demonstrates the relationship among the numbers of transactions, lag numbers and band width. Obviously, the wider the bands, the fewer transactions incurred, and consequently less cost associated even though all three band widths fail to make money.

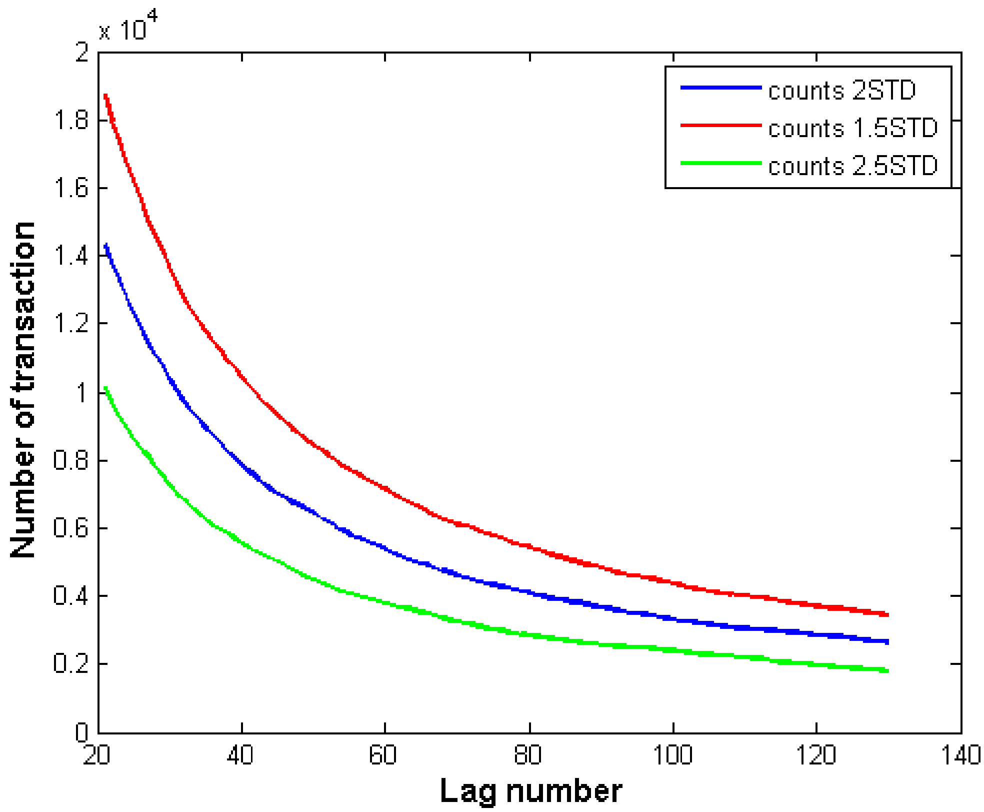

Figure 25 Relationship between numbers of transaction and lag numbers as well as bandwidth.

2.2.3.2 Number of Lags

Figure 22 and figure 23 exhibit the relationship between raw profit, profit and lag numbers. Similar to the relation with band width, the net profit monotonically increases with the increasing of lag numbers, and stays below zero. Neglecting the transaction fee, no simple monotonic relation is expressed in the raw profit figure. When number of lags is below 100, the relationship does not prove any noticeable trend, but a sudden jump happens when number of lag raises over 100 and BB starts to create somewhat larger raw profit which is still away from significance if holding period is considered.

As discussed in the band width section, referring to the histogram chart, the N value represents the total number of different lags examined. And most of time a negative or zero raw profit is generated and never a positive net profit is discovered. Purely looking at BB will not make promising trading results.

2.2.4 *Summary*

Bollinger bands are one of the most popular technical tools for traders to determine overbought and oversold conditions. Taking a range-bound market for example, Bollinger Bands serves very well as prices travel between the two bands. (Cooper 2009)

However, using the Bollinger Band as a sole buy/sell indicator is not very smart given the result of our study. It is more rational to combine other techniques together with Bollinger Bands to help call tops and bottoms. As John Bollinger says, "Tags of the bands are just that – tags, not signals. A tag of the upper BB is not in and of itself a sell signal. A tag of the lower BB is not in and of itself a buy signal."

Comparing to the result of the MACD, in our case, BB generates less negative results, but still negative. For both algorithmic trading (where raw returns matter) and statistical arbitrage (where net returns matter), neither MACD nor BB solely would be a good choice. This result is also an implication of the *Weak Form Market Efficiency*, which states that technical analysis techniques will not be able to consistently produce excess returns. The market participants are not able to systematically profit from market inefficiencies by adopting the simple indicators rooted in MACD or Bollinger Bands.

2.3 Long Only Filter Strategy

One particular long only filter statistical trading strategy is proposed and tested to see if abnormal excess profit can be generated by choosing specific buy and sell trigger parameters. A buy-and-hold benchmark for risk and profit comparison will be introduced as well in the coming section. The strategy is designed to take advantage of short term volatility. After a position has been bought, volatility will make the price oscillate around that buying price. The key is to sell when the current price is higher than the buying price. If the current price is below the buying price, then the trader waits. The major assumption underlying this strategy is that the S&P500 will not go bankrupt. This is not the case of any individual stock. So ultimately, every bought position will be sold at a higher price.

2.3.1 *Trading Rules*

In this section, the number of long position is no longer restricted to exact one share and excessive long or short/sell transactions are also allowed under this trading rule. Figure 26 thoroughly describes how different types of transactions are initiated.

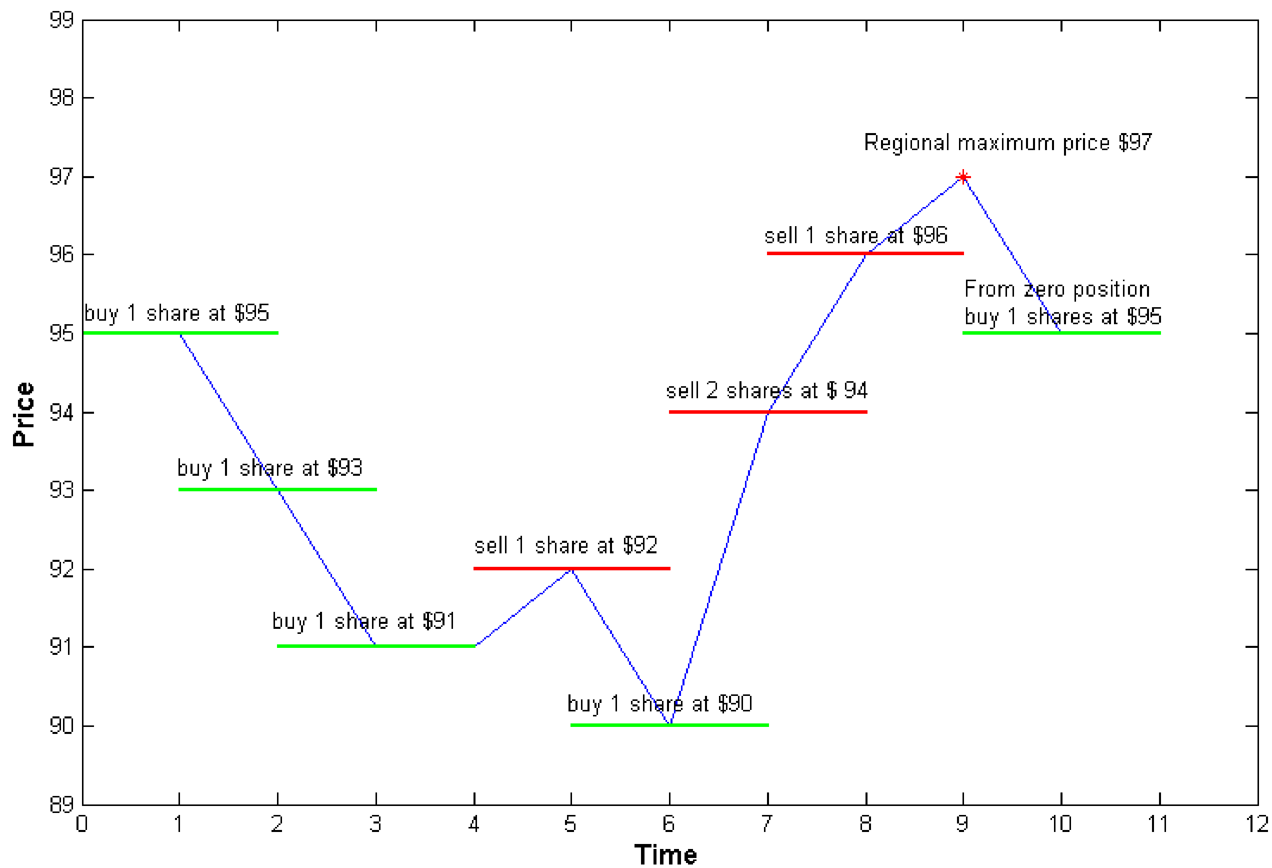

Figure 26 Illustration of long only filter strategy (LOFS). Green line represents the long action, and red line represents the sell action. Regional maximum price is marked as a red star.

Table 1 LOFS Trading Rule Schedule

Time	1	2	3	4	5	6	7	8	9	10	11
Price	95	93	91	91	92	90	94	96	97	95	95
Action	Initial Buy	Buy	Buy	None	Sell	Buy	Sell	Sell	None	Initial Buy	None
Position	1	2	3	3	2	3	1	0	0	1	1

Table 1 lists the full schedule of trading rules:

Time 1 Initiate the trade, buy one share at $95, position = 1;

Time 2 Price drops to $93, because it is $2 lower than the previous position, buy one share at $93, position = 2;

Time 3 Price drops to $91, because it is $2 lower than the previous position, buy one share at $91, position = 3;

Time 4 Price stays at $91, no transaction, position = 3;

Time 5 Price increases to $92, since it is $1 above the previous $91 position, sell one share at $92, position = 2;

Time 6 Price drops to $90, because it is $3 lower than the previous $93 position, buy one share at $90, position = 3;

Time 7 Price increases to $94, because it is $4 above the previous $90 position and $1 dollar above the previous $93 position, sell two shares at $94, position = 1;

Time 8 Price increases to $96, because it is $1 dollar above the previous $95 position, sell one share at $96, position = 0;

Time 9 Position = 0, compare current price with all previous prices, since current price is the maximum, no transaction;

Time 10 Price drops $2 below the previous maximum price where position = 0, an initial buy is triggered, buy one share at $95, position = 1;

Time 11 Price stays at $95, no transaction, position = 1;

Since it is a long only strategy, shorting is not permitted here. Once a sell signal emerges and position is not zero, a sell transaction is established; if the previous position is zero, even with a sell signal, no transaction will incur for current timing. In the illustration above, if current price is at least $2 lower than the previous positions, a buy signal is induced; if current price is at least $1 above the previous positions, a short/ sell signal is induced. Transaction cost is fixed at 3 cents. In the model analysis, two filter parameters called buy sigma and sell sigma in the format of percentage are utilized to identify buy / sell signals. If price drops by buy sigma, a buy transaction is formed; while if price increases by sell sigma, a sell transaction is formed. Both buy sigma and sell sigma range from 0.1% to 10% corresponding to the 1 dime to 10 dollar given the underlying value is $100. These ranges reliably mimic the real world conditions.

We trust C++ to carry out the strategy and track the raw profit, net profit, total transaction counts. In order to reveal the risk associated with this strategy, we also recorded maximum continuous buy counts and maximum continuous cash outflow. For simple comparison with benchmark, the profit from the LOFS is defined as:

$$\text{Profit} = \frac{\$100000 \times \text{total net profit}}{\text{maximum continuous buy counts}}$$

However, if the maximum continuous cash outflow excesses the total cash outflow generated by the maximum continuous buy counts, this definition will not well represent the profitability level. An adjusted profit is deliberated to discover the optimal profit level for a given total investment dollar amount. In our case, this amount is set as $100, 000. The adjusted profit is calculated as:

$$\text{Adj Profit} = \frac{100000 \times \text{total net profit}}{\text{maximum continuous cash outflow}}$$

In the following section, the results related to filter parameters sensitivity to profit and risks as well as the fitness of the strategy will be discussed.

2.3.2 *Benchmark*

The buy-and-hold benchmark is built on the total investment of $100,000. Once the $100,000 is used up, no more long action could be triggered. Risk of the benchmark is measured by the percentage of total investment. Since this benchmark strategy will hold the whole investment all the time, the position at any point of time before the last trading minute of 2006 is 100%. The dollar profit of the benchmark is calculated as below:

$$\text{Profit}_{\text{benchmark}} = \frac{\$100000 \times (\$141.5804 - \$114.3700)}{\$114.3700} = \$23791.55$$

In order to match the risk profile of our long only filter strategy, Treasury notes are introduced here together with the benchmark to mimic the customized strategy risk level. Treasury notes are usually considered risk free since no one is expecting U.S government is going to bankrupt. Since our trading period is from 2002 to 2006, we are using the averaged 5-year T-bond rate to calculate the interest income. The average value used for calculation is 3.34%. The interest yielded is calculated as below:

$$\text{interest} = \$100000 \times (1+3.34\%)^5 - \$100000 = \$17853.45$$

A portfolio composed of T-bonds and the benchmark will be calculated in term of dollar profit and risk to compare with selected long only filter strategy profit and risk patterns. This comparison will help to identify the optimal buy and sell parameter specifications of the long only filter strategy for statistical traders.

2.3.3 *Results*

Talking about profit without considering the associated risks is meaningless. For conservative investors, lower profit may be preferred given a lower portfolio risk; while for aggressive investors, high profit with high portfolio risk may be favored since extraordinary return is what they are looking for and they cares less about the magnitudes of risks. Since little money was made by the algorithmic strategies, the risk associated with those tools is not discussed in detail in previous sections. However, more significant profit is generated under statistical arbitrage; it is a must to consider the risk associated with such abnormal positive profit.

Since both the benchmark and our strategy are recognized as long the index, the strategy that has the largest quantity held in the index is the more risky one. For example, the buy-and-hold will keep a 100% full position in index at all times. On the other hand, a strategy that put 50% in index and 50% in cash is considered less risky. A percentage position parameter is defined to represent the average position in index at every sigma buy and sigma sell level adopted in the strategy. The parameter is derived as the weighted average counts of buys divided by the maximum consecutive number of buys, where the weighted average counts of buys is defined as the sum of positions at each time point divided by the total number of transactions.

$$\text{Weighted average number of buys} = \frac{\text{sum of long positions at every time point}}{\text{total minutes}};$$

$$\text{Position} = \frac{\text{weighted average number of buys}}{\text{maximum consecutive number of buys}}$$

Take the data in Table 1 for example, the max consecutive number of buys = 3:

Weighted average number of buys = (1+2+3+3+2+3+1+0+0+1+1)/11 = 1.55;

Position = 1.55 / 3 = 0.52 = 52% < 100% which indicates less riskiness of the strategy than benchmark.

LOFS potentially could help investors to reduce their risk level, however, if it could meanwhile maintain the profit level as the bench is yet not proofed. Figure 27 illustrates the profit distribution for all buy/sell parameters specifications. Clearly, the plain is basically yellowish and greenish color with the value around $15000. The highest profits appear where sigmabuy is large and sigmasell is relatively small. Investors could generate the highest profit level when price drops a lot to buy and rise a little bit to sell.

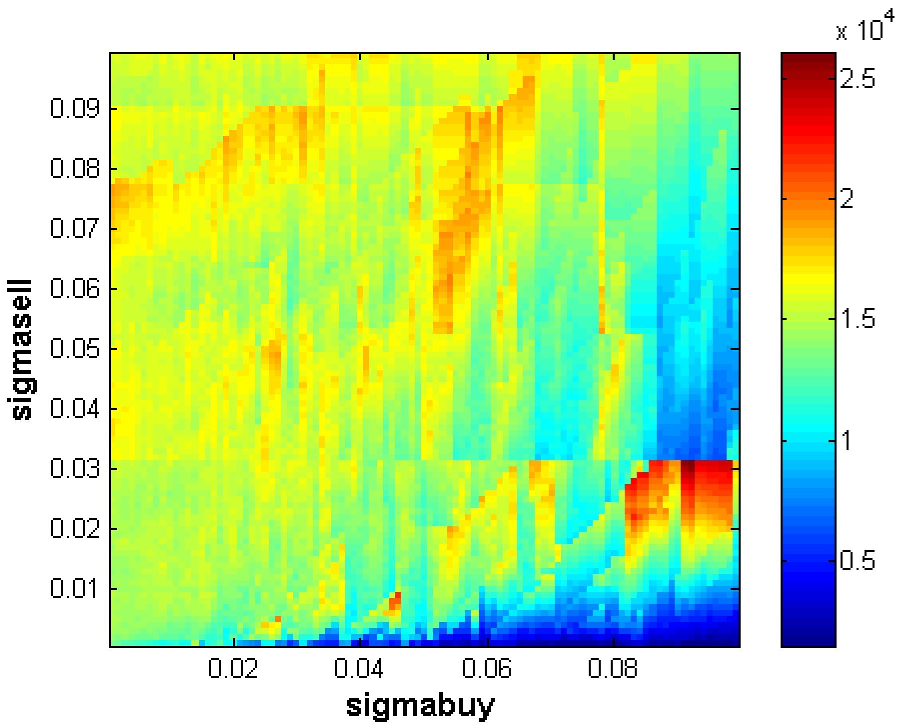

Figure 27 Profit distributions for all buy and sell parameters specifications. The highest profit resides at sigmabuy > 8% and 2% < sigmasell <3%. The lowest value stays where sigmasell is significantly small (less than 1%). Overall, the panel is yellowish which represents a mean dollar profit around $15000.

In the same figure, a clear profit kink appears at simgasell = 0.03 and sigmabuy ranges between 0.08 and 0.1. For a given period, if the price movement is less dramatic, fewer big price jumps would be observed and recorded, and fewer prices would match the sell or buy signals thus less profit is made. If the buy or sell parameter is too big that none of the price ever matches the criteria, a zero profit would be observed since no transaction is made and no profit is realized. By decomposing all adjusted profit into annual level, obviously this kink is induced in 2002. Figure 28 to Figure 32 illustrate the adjusted profit distribution in 2002 through 2006. In the example of year 2002, when sell sigma is set too large given a large big sigma buy, very few prices ever match the criterion and a sell transaction could not be easily set off and no profit could be generated which corresponds to the fact of low profit area is produced.

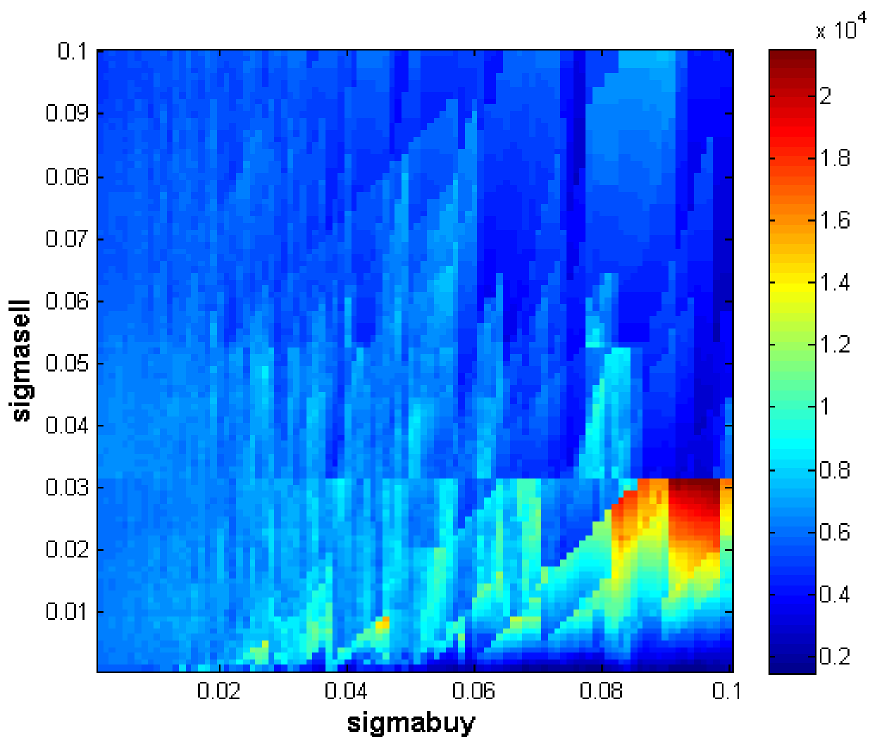

Figure 28 2002 adjusted profit distribution for all buy and sell parameters specifications.

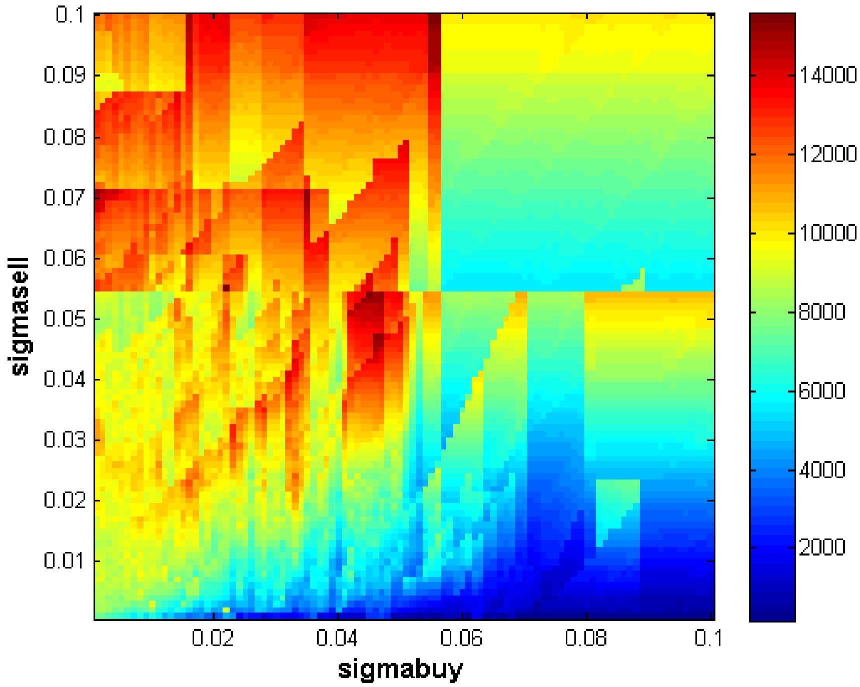

Figure 29 2003 adjusted profit distribution for all buy and sell parameters specifications.

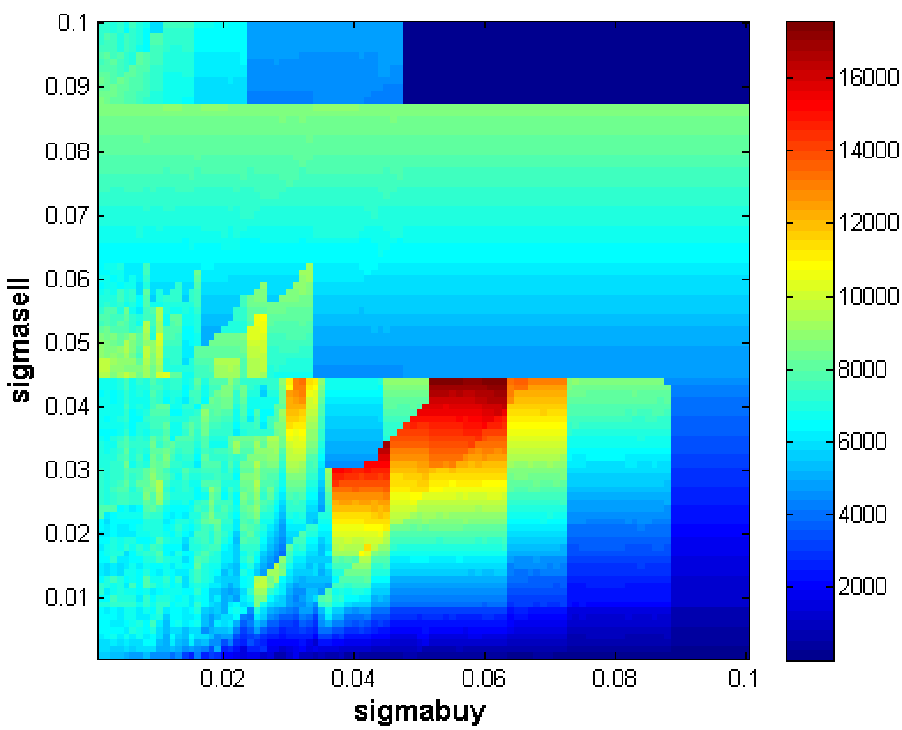

Figure 30 2004 adjusted profit distribution for all buy and sell parameters specifications.

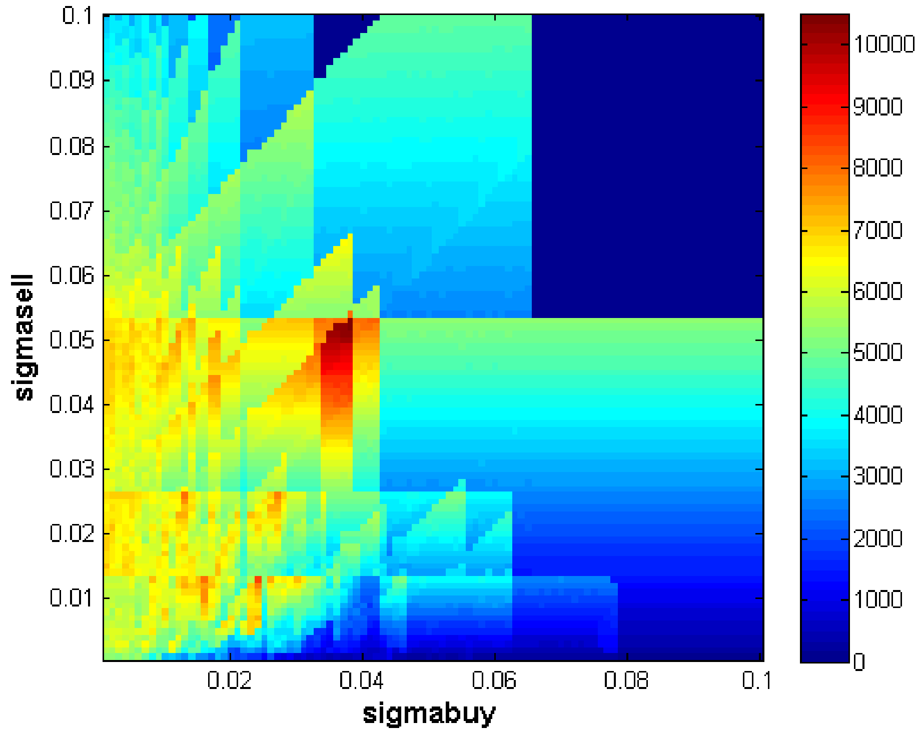

Figure 31 2005 adjusted profit distribution for all buy and sell parameters specifications.

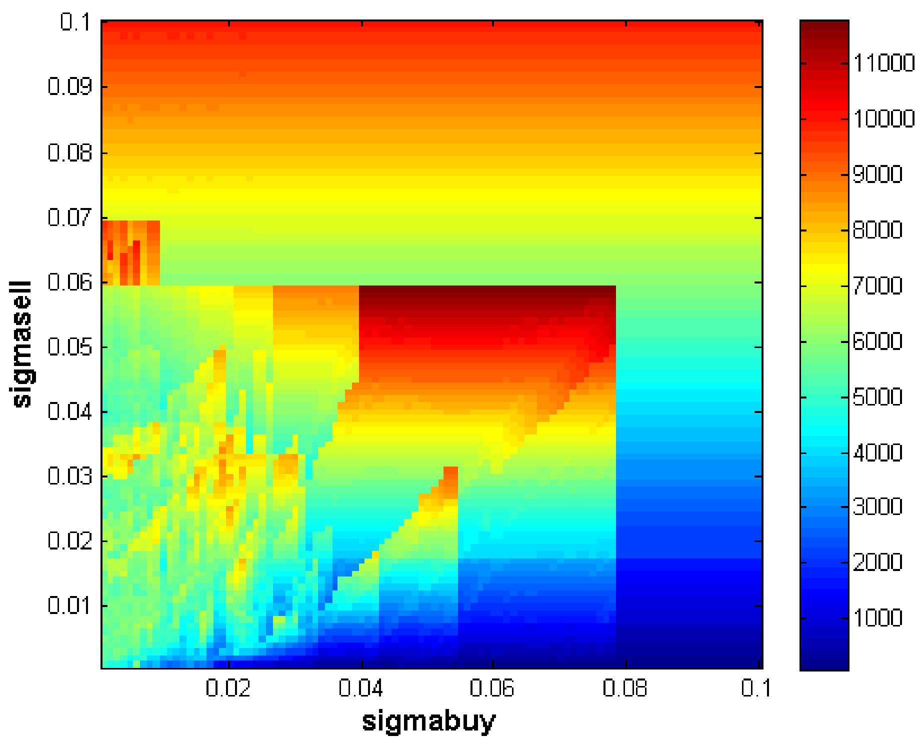

Figure 32 2006 adjusted profit distribution for all buy and sell parameters specifications.

By further decomposing 2002 annual data into monthly data, we notice that multiple months contribute to the kink. (Figure 33- Figure 35). Except for October, the rest of the months display a relatively low profitability when sigma sell increases above 5%, which indicates that the price never rise up to 5% or more for a given period in year 2002 and no sells could be executed to make profit. This result is not surprising since a bearish trend clearly displayed in index price time series in 2002 (Figure 1). It is never easy to profit from a bearish market.

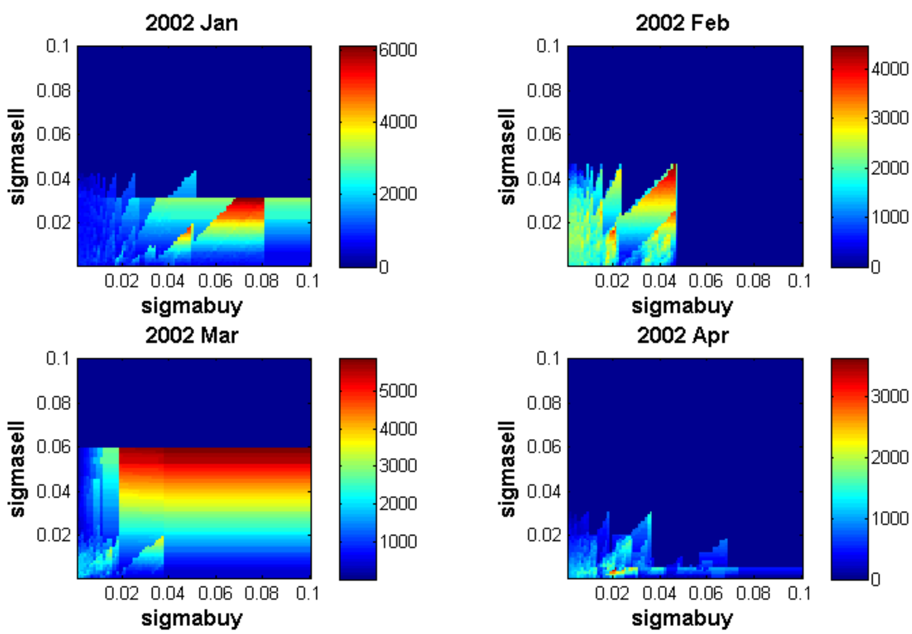

Figure 34 2002 Jan - Apr adjusted profit distribution for all buy and sell parameters specifications.

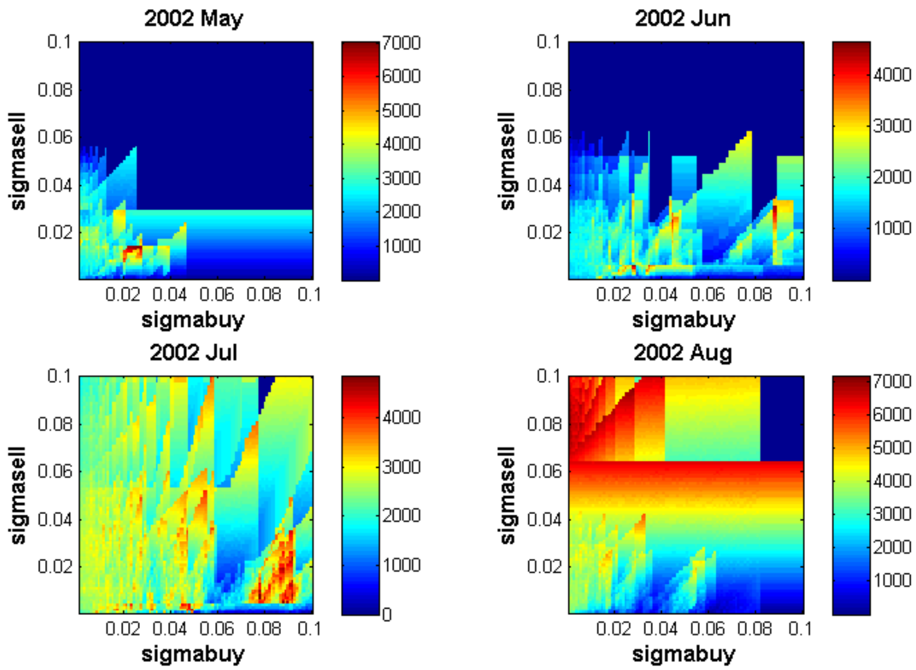

Figure 34 2002 May - Aug adjusted profit distribution for all buy and sell parameters specifications.

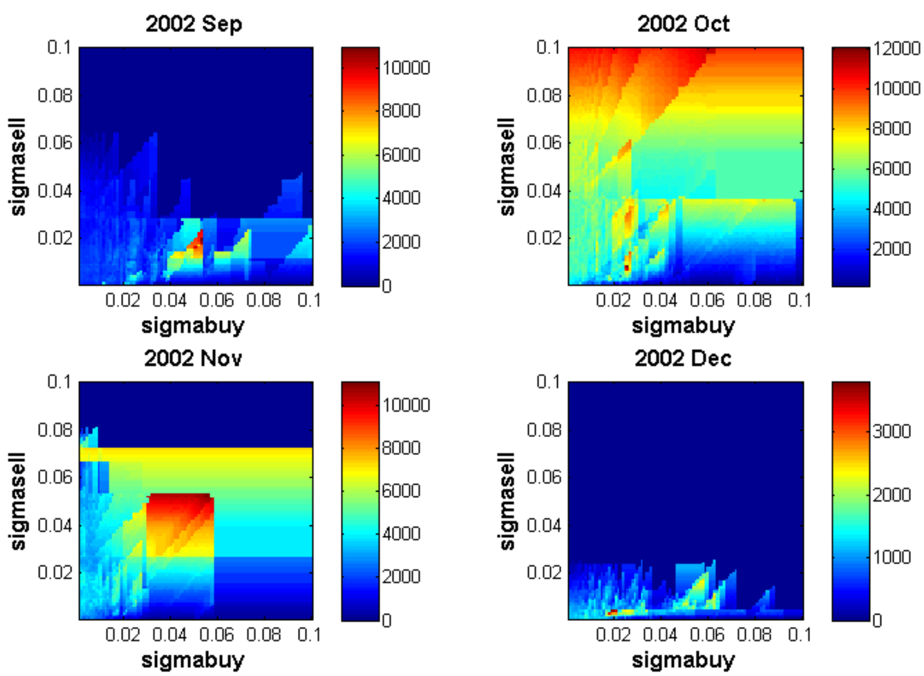

Figure 35 2002 Sep - Dec adjusted profit distribution for all buy and sell parameters specifications.

In Figure 36, the left panel graphs the histogram of dollar profit for all specifications, and the right panel displays the risk levels for all specification. The mean adjusted profit and mean position are calculated. This strategy significantly reduces the total investment risk level in term of position; however, without carefully choosing proper buy and sell parameters, this method can not generate abnormal profit comparing with benchmark for investors.

Choosing from LOFS and benchmark would be up to investors' risk and profit preferences. To further test the superiority of LOFS to benchmark, we introduce a portfolio composed of benchmark and 5-year Treasury notes to imitate the risk level of LOFS so that the profit from both could be compared at the same risk level. The mean risk level of LOFS (22.569%) will firstly be adopted to generate the portfolio.

Since Treasury notes are considered risk free, the risk of the portfolio is determined by the weight of benchmark. The following equations illustrate how the risk and dollar profit of this portfolio is calculated.

$$\text{risk}_{\text{portfolio}} = 22.569\% \times 100\% + (1\text{-}22.569\%) \times 0\% = 22.569\%$$
$$\text{profit}_{\text{portfolio}} = 22.569\% \times \$23791.55 + (1\text{-}22.569\%) \times \$17853.45 = \$19193.62$$

To compare the LOFS with portfolio, excess profit is created as adjusted profit minus portfolio profit to exam the superiority of LOFS. Figure 37 provides an insight of the relationship between LOFS and the portfolio of benchmark and Treasury Notes.
At the average level, the portfolio outperforms the LOFS strategy since it produces more dollar profit with the same risk level. Comparing the mean value of LOFS adjusted profit and portfolio profit, the latter is slightly higher than the former. The

average excess profit is -$2291.82, and its distribution displays more counts in the negative ranges.

Does this result imply that LOFS also fails? Our answer is not at all. Investors will not just randomly choose buy / sell signal but choose the optimal ones to maximize their profit or minimize cost. Same logic also applies to LOFS; investors would need to choose different buy/sell triggers to optimize their investment. For such reason, we choose one specification with the maximum profit: sigmabuy = 9.2%, sigmasell = 3.1%, adjusted profit = $26076.23. The position for this pair of parameters = 6.6604%.

Similar to the risk and profit calculation for portfolio above, in this case the risk and profit are computed as:

$$risk_{portfolio} = 6.6604\% \times 100\% + (1\text{-}6.6604\%) \times 0\% = 6.6604\%$$
$$profit_{portfolio} = 6.6604\% \times \$23791.55 + (1\text{-}6.6604\%) \times \$17853.45 = \$18248.95$$

Obviously, the portfolio profit is much less than the LOFS adjusted profit. By carefully selecting appropriate trigger parameters, LOFS is a promising strategy for statistical traders who care about transaction cost.

Distinct from the other two technical analysis tools, this long only filter strategy generates positive net adjusted profit after transaction cost at every parameter level. The maximum adjusted profit locates at where buy sigma = 9.2% and sell sigma = 3.1%.

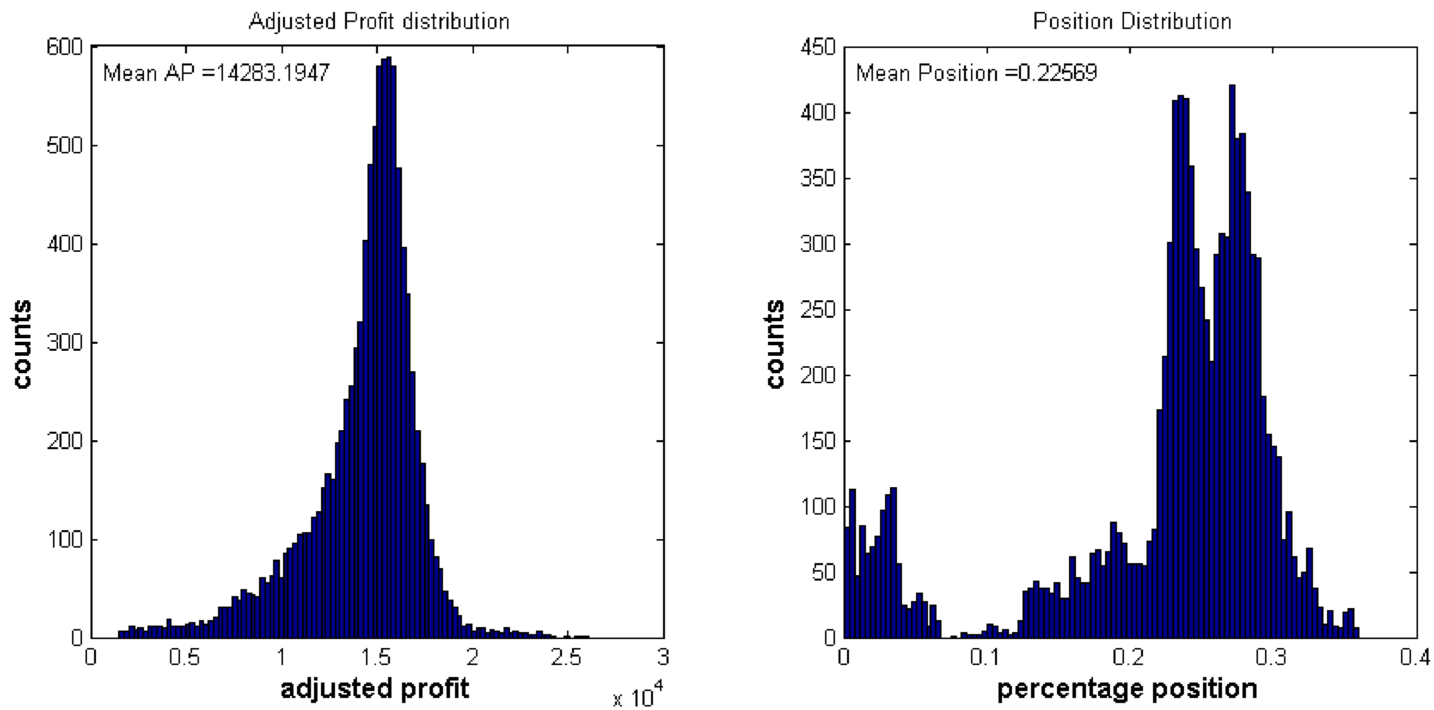

Figure 36 Risk and profit distributions of LOFS parameters specifications. On the left panel, the distribution of adjusted profit is illustrated. The mean adjusted profit (AP) equals $14283.1947. The mean risk level is measured by mean position which equals to 22.569

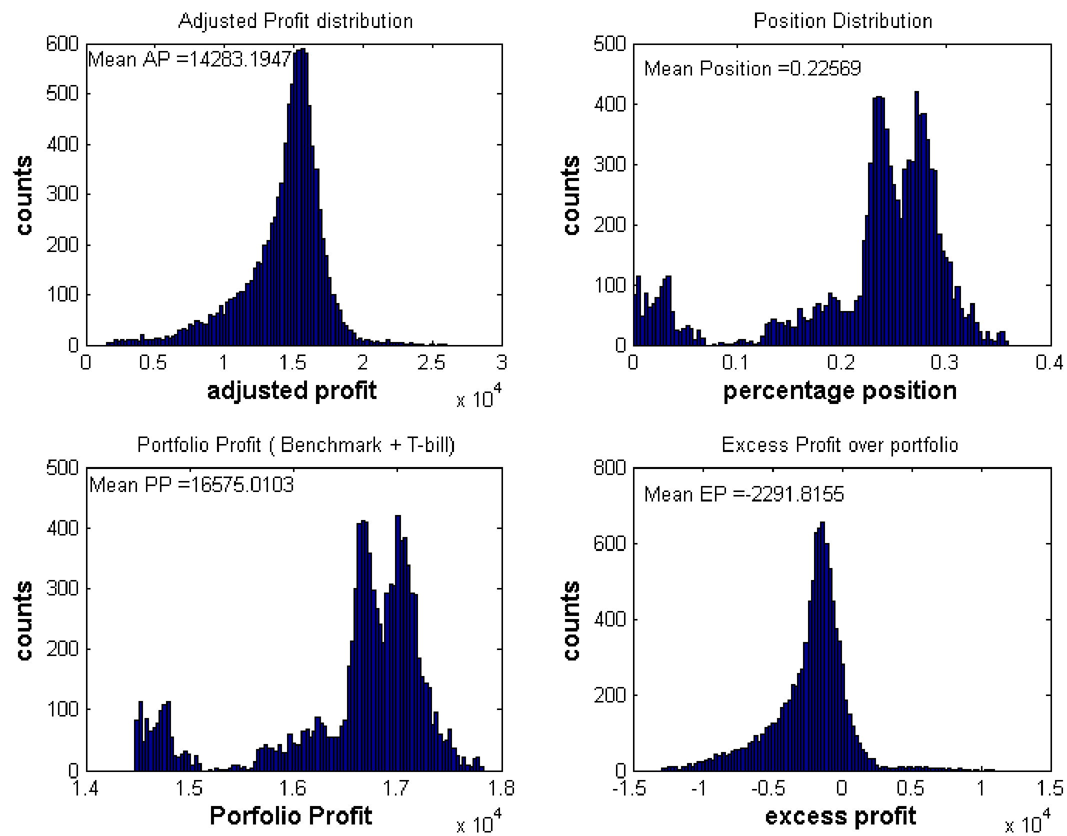

Figure 37 Comparisons between LOFS and Portfolio. The upper panel provides distribution information for LOFS adjusted profits and risk levels. The lower left panel illustrates the distribution of portfolio profit and the lower right panel tells us the distribution of excess profit which equals to adjusted profits minus portfolio profit. The mean value for each panel is also listed

2.4 Summary

Compared to the previous two algorithmic trading strategies, this long only filter strategy is able to generate large positive profit. This is expected by the nature of the strategy: sell price is always higher than buy price. The volatility of prices at different periods plays an important role for the selection buy / sells parameters. If price movement is significant, larger parameters would capture such price pattern and generate considerable revenue; while volatility is relatively low, large sell parameter will not produce sound investment. In the down trend of price (Year 2002), less profit could be driven out of large sell sigma, but strong profit is recorded where buy sigma is set to be large. When the price is upward trending, both sell sigma and the volatility of the underlying price will impact the magnitude of profit. All the profit discussed is in absolute dollar basis.

In terms of the investment risk associated with this filter strategy, the risk level is always below that of the buy-and-hold; while the absolute adjusted profit does not outperform the simple buy-and-hold strategy all the time. Nevertheless, investors with limited budget and conservative investment style would still prefer the filter strategy which presents a lower risk relative to the benchmark. Comparing to the portfolio of benchmark and 5-year Treasury Notes, on average LOFS does not outperform the portfolio. However, by carefully choosing appropriate buy and sell criteria, both lower risk and higher profit relative to the benchmark could be realized. The optimal parameter level is sigmabuy equals to 9.2% and sigmasell equals to 3.1%. If history repeats itself, statistical investors who adopt such specifications should generate abnormal profit and lower risk in later years.

3 SUMMARY

3.1 Conclusion

This work starts with the investigation of the effectiveness of the two most popular technical analysis indicators - MACD and Bollinger Bands for algorithmic trading and statistical arbitrage. Intraday SPY time series prices between 2002 and 2006 are used. The profit results from both strategies confirm that simple indicators generated from MACD and Bollinger Bands fail to capture appropriate timing and price for trades.

Algorithmic traders who concentrate on execution of large volume of trade will not be able to benefit from MACD crossover signals. The selection of lag parameters will not improve the profit before transaction cost. After transaction cost, the profit will be worsened and statistical arbitragers are unable to generate abnormal profit based on simple MACD oscillators either. In one word, MACD does not provide any useful strategy here to either optimize profit or reduce transaction cost.

Similar conclusions are also drawn for the Bollinger Bands. Careful selection of lag parameters and band width could generate some positive profit before transaction cost. However, the profit is to small to be worth the time and effort of the algorithmic traders. For statistical arbitragers, transaction costs immediately wear away the tiny positive raw profit. Basically, taking advantage of Bollinger Bands alone can not yield any excess income for them. Solely depending on Bollinger Bands to minimize transaction cost and discover investment opportunities is not realistic.

The Long Only Filter Strategy (LOFS) created in this study outperforms both MACD and Bollinger Bands in term of its effectiveness for creating sizeable positive profit before and after transaction cost. The nature of this strategy (that exit price is forced to be higher than the entrance price) guarantees that it would not produce negative profit. The sell and buy parameters play a core role in yielding high profits. Large price drops captured by big buy parameters together with small price rises captured by small sell parameters generate the optimal level of after transaction cost profit. This strategy also outperforms the benchmark buy-and-hold strategy in term of the magnitude of net profit and associated risks. LOFS that generate comparable net profit as the benchmark turns out to hold a much less average position which indicates less risk. In other words, investors could make the same amount of profit by investing a much smaller amount of cash in LOFS than in the benchmark. Meanwhile investors could invest the remaining cash in some other financial product such as Treasury notes to further grow their fortune. When comparing to the portfolio composed of benchmark and Treasury notes, by carefully choosing trading parameters, LOFS could still outperform it assuming that in the future, the history will repeat itself and the S&P 500 does not go bankrupt.

Nevertheless, investors' risk preference should be considered when employing LOFS. The average risk for all LOFS of different parameter specification is lower than the benchmark, but the average profit is also lower than that of the benchmark. It is up to the individual investor to decide whether LOFS should be applied and what kind of buy and sell thresholds should be specified for LOFS. For conservative investors, LOFS may be preferred because of its low risk characteristic. For investors who have aggressive investment style, the employment of LOFS will be an open question.

The issues with LOFS have been discussed in the earlier section. The usage of LOFS is based on the assumption that history repeats itself, and investors are at 100% investment position at the bottom of the stock cycle. If this bottom is underestimated in the future, the investors will be stuck in the market taking full position risks as the benchmark while the profit may not necessarily beat the benchmark either. If market exits the bottom faster than expected, then investors fail to fully take advantage of the strategy and the profit is not at the optimal level. Thus the estimation of maximum successive buy count is crucial in this analysis and a well understanding of the market behavior is also critical.

3.2 Future Work

The effective indicators based on MACD or Bollinger Bands are not limited to crossover signals or parameter selections. However in this study we only focused on the most widely used oscillator signals and the adjustment for their parameters. The failure of MACD and Bollinger Bands in this study does not imply they never work under any circumstances. More sophisticated MACD and Bollinger Bands indicators could be created, employed and combined to further investigate their usefulness.

Investors never talk about return without describing the associated risk. The success of LOFS makes a thorough risk analysis important. Other more advanced risk analysis using concepts such as Sharpe ratio could be computed to match the customized risk and return interests for different market players. The risk characteristics of the LOFS in terms of the specification of parameters for different financial instruments are also worth exploring. The robustness of the LOFS optimal parameters (buy and sell thresholds) over time is also an interesting extension. The time period could be divided into two halves and the optimal parameters for one half can be compared with those of the other half.

REFERENCES

Barber, Brad M., Lee, Yi-Tsung, Liu, Yu-Jane and Odean, Terrance, *Do Individual Day Traders Make Money? Evidence from Taiwan.* Electronic access: http://ssrn.com/abstract=529063, *2005.*

Bollinger Bands Tutorial. Electronic access: http://www.bollingerbands.com.

Brown, P. David and Jennings, H, Robert. *On Technical Analysis*. Review of Financial Study 1989; 2: 527-551., 1989.

Cooper, Ian. *Bollinger Bands Explained. How to profit using Bollinger bands: A widely misused technical indicator.* Wealth Daily, May 5th 2009.

Economist.com, Looking for options Derivatives drive the battle of the exchanges. April 15, 2007.

Economist.com, Technology and exchanges. Feb 2nd 2006.

Hong, Harrison and Stein, C. Jeremy. *A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets*. The Journal of Finance Vol LIV.No.6., 1999.

Kaufman, J.Perry. *A Short Course in Technical Trading*. p199 – 212, 2003.

Lo, Andrew, Mamaysky, Harry and Wang, Jiang. *Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation.* The Journal of Finance Vol. LV, No. 4., 2000.

Murphy, John. *Technical Analysis of the Financial Markets*. pp. 252–255, 1999.

Saphir, Ann. *Chaos theorists: Proprietary trading firms racking up huge profits*. Crain’s Detroit Business., 2009.

Tanompongphandh, Thanasin. *On The Impossibility of Testing Weak-Form Efficiency of Ultra-High Frequency Stock Prices.* M.S Thesis, Cornell University., 2008.

The Economist. *Algorithmic trading, Ahead of the tape*. 383: p. 85., June 23, 2007.

The Wall Street Journal Europe, *MTS to mull bond access*., p. 21., April 18, 2007.

Timmons, Heather. *A London Hedge Fund That Opts for Engineers, Not M.B.A.'s.* New York Times, August 18, 2006.

Waldemar, Stronka. *On the Analogy between Scientific Study of Technical Analysis and Ethnopharmacology*. 2006.

APPENDICES

A.1 Job Advertisements

1. Intraday Statistical Arbitrage Trading Strategist
   Location: NYC
   $150,000 + huge percentage payout

   Description:
   The hottest algorithmic proprietary trading firm in New York is feverishly expanding. They are in the process assembling the most successful ultra-high frequency trading team in the world. With their infrastructure, technical resources and capital allocation, your short-term strategies will flourish. In addition to providing one of the leading high frequency trading platforms, you will be eligible for one of the industry's highest percentage payouts.

   Requirements:
   In order to be considered as a candidate for this exclusive group, you must have developed highly efficient, scalable intraday trading strategies within a well-known quantitative trading group. The ideal candidate will also have experience developing models, algorithms and production code. An advanced degree in computer science, physics, statistics, mathematics, etc… is always preferred.

   Job Reference #: BB036
   The Hagan-Ricci Group
   Phone: (212)-681-6333
   http://www.hrg.net/index.php/job_post/view/314

2. GOLDMAN SACHS ELECTRONIC TRADING STRATEGIES

New Jersey/New York Campus
Analyst

The Goldman Sachs Electronic Trading Strategies team works directly with the electronic trading desk to enhance and optimize the Goldman Sachs suite of algorithms, develop pre-trade and post-trade analytical tools, and guide clients (institutional managers, hedge funds, and broker/dealers) to improve their trade execution performance and algo usage.

JOB DESCRIPTION:
Strategist at analyst level to work on customized execution quality analysis for key clients and on the ongoing development and maintenance of the team's trading analytics and algorithms. Job responsibilities are split evenly across quantitative analysis, writing on execution performance and hands-on coding.

JOB QUALIFICATIONS:
College graduate with strong quantitative, statistical and programming skills. A Master degree in a quantitative discipline is a plus but not required.

REQUIRED:
Experience in processing, troubleshooting and analyzing large data sets. Should be creative and able to perform well when working independently, meet tight deadlines and display strong teamwork abilities. The candidate must have strong oral and written communication skills. Desired but not required: Knowledge of finance and econometrics, relevant experience in financial services industry and/or relevant academic experience.

APPLICATION PROCEDURE:
Please email resume and cover letter to:
Email: MAILTO:gsetstratrecruiting@gs.com

A.2 SAS Code For MACD and Bollinger Bands

1. MACD

```
libname macdtest 'D:\My doc\2009 Spring\Research\macdtest';
libname bymonth 'D:\My doc\2009 Spring\Research\bymonth';
libname byday 'D:\My doc\2009 Spring\Research\byday';
options mprint mlogic fullstimer;

%macro setew;
 %do i=1 %to 130 ;
  price&i = lag&i(wprice);
 %end;
 %mend;

%macro ewma;
*-- fast ewma --*;
  mf = (1- (2/(60+1)));   wf = 0;  emaft = 0;
  %do i=1 %to 60;
   emaf&i =(mf**&i)*(price&i);  emaft = emaft +emaf&i; wf = wf + (mf**&i) ;
  %end;
  emaf = emaft / wf;

*-- slow ewma --*;
 ms= (1- (2/(130+1))); ws = 0;  emast = 0;
 %do i=1 %to 130;
  emas&i =(ms**&i)* (price&i); emast = emast +emas&i; ws = ws + (ms**&i) ;
 %end;
  emas = emast / ws;  macd = emaf - emas;
%mend;

%macro macdsignal;
macdma = 0; wsig = 0; sig = 1-(2/(45+1));
retain wsig 0;
%do i=1 %to 45 ;
   macd&i = (sig**&i)*lag&i(macd); macdma = macdma+macd&i; wsig = wsig + (sig**&i);
%end;
macdsig = macdma/wsig;
%mend;

data macdtest.eligibledata_ewma_1;
set macdtest.eligibledata2;
%setew;
%ewma;
%macdsignal;
macdbar = macd - macdsig; lmacdbar = lag1(macdbar); macdj = macdbar * lmacdbar;
run;
quit;

data macdtest.eligibledata_ewma;
set macdtest.eligibledata_ewma_1;
if macdj > 0  then change = 0;
if macdj < 0 and macdbar > lmacdbar  then change = 1  ; * buy / long signal;
if macdj < 0 and macdbar < lmacdbar  then change = 2 ; * sell/ short signal;
```

```
if change = 1 then nmacd1 = macd;
if change = 1 then nwprice1 = wprice;
if change = 2 then nmacd2 = macd;
if change = 2 then nwprice2 = wprice;
if change = 1 or change = 2 then nmacd = macd ;
if change = 1 or change = 2 then nwprice = wprice;
if macdj = . then change = 0 ;
if change = 0 then nwprice = .;
if change = 0 then nmacd = .;
if change = 0 then nwprice1 = .;
if change = 0 then nmacd1 = .;
if change = 0 then nwprice2 = .;
if change = 0 then nmacd2 = .;
run;
quit;

data macdtest.eligibledata_ewma_profit;
set macdtest.eligibledata_ewma;
if change NE 1 and change NE 2 then delete;
run;
quit;

data macdtest.eligibledata_ewma_profit_5;
set macdtest.eligibledata_ewma_profit;
lnwprice = lag1(nwprice);
if change = 1 then profit = lnwprice - nwprice;
if change = 1 then retn = profit / lnwprice; * past second is shorting;
if change = 2 then profit = -lnwprice +nwprice ; * past secod is long position;
if change = 2 then  retn = profit / nwprice;
run;
quit;

data macdtest.eligibledata_ewma_profit;
set macdtest.eligibledata_ewma;
if change NE 1 and change NE 2 then delete;
run;
quit;

data macdtest.eligibledata_ewma_profit_3;
set macdtest.eligibledata_ewma_profit;
lnwprice = lag1(nwprice);
if change = 1 then profit = lnwprice - nwprice - 0.03;
if change = 1 then retn = profit / lnwprice; * past second is shorting;
if change = 2 then profit = -lnwprice +nwprice - 0.03; * past secod is long position;
if change = 2 then  retn = profit / nwprice;
run;
quit;
```

2. Bollinger Bands

```
libname macdtest 'D:\My doc\2009 Spring\Research\macdtest';
libname bband 'D:\My doc\2009 Spring\Research\bband';
options mprint mlogic fullstimer;

%macro setma;
 %do i=1 %to 130 ;
  price&i = lag&i(wprice);
  %end;
%mend;

%macro map;
%do j = 21 %to 130;
aprice&j = 0;
%do i = 1 %to &j;
aprice&j = aprice&j + price&i;
%end;
map&j = aprice&j/&j;
%end;
%mend;

%macro stdp;
%do j = 21 %to 130;
var&j = 0;
%do i = 1 %to &j;
var&j = var&j + (price&i-map&j)*(price&i-map&j);
%end;
varp&j = var&j / &j;
stdp&j = sqrt(varp&j);
%end;
%mend;

%macro bbandonehalf;
%do j = 21 %to 130;
date&j = date; second&j = second; wprice&j = wprice; id&j = newid;
upper&j = map&j + 1.5*stdp&j;
lower&j = map&j - 1.5*stdp&j;

buysig&j = wprice&j - lower&j; lbuysig&j = lag1(buysig&j); bsig&j = buysig&j * lbuysig&j;
sellsig&j = upper&j - wprice&j; lsellsig&j = lag1(sellsig&j); ssig&j = sellsig&j*lsellsig&j;

if bsig&j < 0 and buysig&j > lbuysig&j then bprice&j = wprice&j;
if ssig&j < 0 and sellsig&j > lsellsig&j then sprice&j = wprice&j;
if bprice&j ne . then tprice&j = 1; * 1 means buy;
if sprice&j ne . then tprice&j= 2; * 2 means sell;
keep wprice&j  bprice&j sprice&j tprice&j ;

%end;
%mend;

data bband.onehalfbbprofitall;
set macdtest.eligibledata2;
%setma;
```

```
%map;
%stdp;
%bbandonehalf;
run;
quit;

%macro prebbprofitonehalf;
%do j = 21 %to 131;
data bband.onehalfprebbprofit&j;
set bband.onehalfbbprofitall;
if bprice&j ne wprice&j  and sprice&j ne wprice&j then delete;
if tprice&j = lag1(tprice&j) then delete;
keep wprice&j bprice&j sprice&j tprice&j;
%end;
%mend;

%prebbprofitonehalf;

%macro bbprofitonehalf;
%do j = 21 %to 131;
data bband.onehalfbbprofit&j;
set bband.onehalfprebbprofit&j;

if tprice&j = 1 then profit&j = lag1(wprice&j) - wprice&j - 0.03;
if tprice&j = 1 then retn&j = profit&j / lag1(wprice&j); * past second is shorting;
if tprice&j = 2 then profit&j = -lag1(wprice&j) + wprice&j - 0.03; * past secod is long position;
if tprice&j = 2 then retn&j = profit&j / wprice&j;

%end;
%mend;

%bbprofitonehalf;

%macro sumonehalf;
%do j = 21 %to 131;
proc means data = bband.onehalfbbprofit&j;
                        var wprice&j profit&j retn&j;
                        title onehalfma&j;
%end;
%mend;

%sumonehalf;

%macro bbandtwohalf;
%do j = 21 %to 130;
date&j = date; second&j = second; wprice&j = wprice; id&j = newid;
upper&j = map&j + 2.5*stdp&j;  lower&j = map&j - 2.5*stdp&j;

buysig&j = wprice&j - lower&j; lbuysig&j = lag1(buysig&j); bsig&j = buysig&j * lbuysig&j;
sellsig&j = upper&j - wprice&j; lsellsig&j = lag1(sellsig&j); ssig&j = sellsig&j*lsellsig&j;

if bsig&j < 0 and buysig&j > lbuysig&j then bprice&j = wprice&j;
if ssig&j < 0 and sellsig&j > lsellsig&j then sprice&j = wprice&j;
if bprice&j ne . then tprice&j = 1; * 1 means buy;
if sprice&j ne . then tprice&j= 2; * 2 means sell;
```

```
keep wprice&j  bprice&j sprice&j tprice&j ;

%end;
%mend;

data bband.twohalfbbprofitall2;
set macdtest.eligibledata2;
%setma;
%map;
%stdp;
%bbandtwohalf;
run;
quit;


%macro prebbprofittwohalf;
%do j = 21 %to 131;
data bband.twohalfprebbprofit2&j;
set bband.twohalfbbprofitall2;
if bprice&j ne wprice&j  and sprice&j ne wprice&j then delete;
if tprice&j = lag1(tprice&j) then delete;
keep wprice&j bprice&j sprice&j tprice&j;
%end;
%mend;

%prebbprofittwohalf;

%macro bbprofittwohalf;
%do j = 21 %to 131;
data bband.twohalfbbprofit2&j;
set bband.twohalfprebbprofit2&j;

if tprice&j = 1 then profit&j = lag1(wprice&j) - wprice&j - 0.03;
if tprice&j = 1 then retn&j = profit&j / lag1(wprice&j); * past second is shorting;
if tprice&j = 2 then profit&j = -lag1(wprice&j) + wprice&j - 0.03; * past secod is long position;
if tprice&j = 2 then retn&j = profit&j / wprice&j;

keep wprice&j bprice&j sprice&j tprice&j profit&j retn&j  ;
%end;
%mend;

%bbprofittwohalf;

%macro sumtwohalf;
%do j = 21 %to 131;

proc means data = bband.twohalfbbprofit2&j; var wprice&j profit&j retn&j; title twohalfma&j;
%end;
%mend;

%sumtwohalf;
```

A.3 C++ Code For LOFS

```
// strategy.cpp : Defines the entry point for the console application.
//
#include "stdafx.h"
#include "stdlib.h"
#include "string.h"
#include "stdio.h"
#include "math.h"
#include <iostream> using namespace std;

long loaddata(char filename[], double data[]);
long sellfunc(double currentprice, double sigmasell, double* sale, long* nsell);
long longmax(long a, long b);
double doublemax(double a, double b);
double sumbuyhist();
double price[530000], buyhist[530000];
long nbuy;

int _tmain(int argc, _TCHAR* argv[])
{  long i, N, nsell, buynum, nbuymax;
   char filename[] = "filtereddata5yr_0613.txt"; // specify your input file name
   double sigmabuy, sigmasell, sellprice, sale, maxcashpmt, totalsell, totalbuy, totalcost, trancost, retn, costfee,
profit,totalpft, sell, cost, maxprice,totalbuyct,sum;
   N = loaddata(filename, price);  // subroutine loaddata
   costfee = 0.03;
   FILE *fpwrite;
   fpwrite = fopen("resultfinal.txt", "w"); // specify your output file name
   fprintf(fpwrite, "sigmabuy sigmasell totalbuy totalsell trancost totalpft pft1 nbuymax maxcashpmt adjprofit
totalbuyct sum \n");

   //sigmabuy start from 0.001 and increase 0.001 every loop until it reaches 0.10
   for (sigmabuy = 0.001;sigmabuy <= 0.101;sigmabuy+= 0.001)
            {  for (sigmasell = 0.001;sigmasell <= 0.101;sigmasell+= 0.001)
                    {  printf("%lf %lf\n",sigmabuy,sigmasell); //initialization

         retn = 0; nbuy = 1; nbuymax = 1; maxcashpmt = price[0]; sale = 0; sellprice = 0; totalsell = 0;
         totalbuy = price[0]; trancost = costfee; buyhist[0] = price[0]; sell = 0; totalpft = 0; maxprice =0;
         totalbuyct = 1; sum = 1;

         for (i=1;i<N;i++)  { nbuymax = longmax(nbuymax, nbuy);  // subroutine longmax
            maxcashpmt = doublemax(maxcashpmt, sumbuyhist());
                    if ( nbuy == 0 )  { maxprice = doublemax (maxprice, price[i]);
                            if ( price[i] < maxprice*(1-sigmabuy))
                            { nbuy = 1;
                              buyhist[0] = price[i]; totalbuy = totalbuy + price[i];
                              trancost = trancost + costfee; maxprice = 0;
                              totalbuyct = totalbuyct ++;
                              sum = sum +totalbuyct;        }}
                                    else if (nbuy>0 && price[i]<buyhist[nbuy-1]*(1-sigmabuy))
                                            { nbuy++; buyhist[nbuy-1] = price[i];
                                              totalbuy = totalbuy + price[i];
                                              trancost = trancost + costfee;
                                              totalbuyct = totalbuyct++;
                                              sum = sum +totalbuyct;      }

            else if (nbuy > 0 && sellfunc( price[i],sigmasell, &sale, &nsell))  // subrountine sellfunc
            {  sellprice = price[i];
               retn = retn + nsell*sellprice/sale - 1; totalsell = totalsell + sellprice*nsell;
               sell = sellprice * nsell; cost = costfee*nsell;
```

```
                trancost = trancost + costfee*nsell; profit = sell - cost - sale;
                totalpft = totalpft + profit; nbuy = nbuy - nsell;
                totalbuyct = totalbuyct - nsell; sum = sum +totalbuyct;    }
                                        else
                                        {           totalbuyct = totalbuyct;
                                                    sum = sum + totalbuyct;
                                        }; }
   fprintf(fpwrite, "%lf %lf %lf  %lf %lf %lf %lf %ld  %lf %lf %lf %lf \n", sigmabuy, sigmasell, totalbuy, totalsell,
trancost,totalpft, totalsell+nbuy*price[N]-totalbuy-trancost,nbuymax, maxcashpmt,
100000*totalpft/maxcashpmt,totalbuyct, sum); }}

fclose(fpwrite);
   getchar();
   return 0;
}

long loaddata(char filename[], double data[])
{  FILE *fp;
   long i;
   char ch;
   if ((fp = fopen(filename, "r"))==NULL) {
      printf("cannot open file\n"); return -1; }
   for(i=0;!feof(fp);i++) { fscanf(fp, "%lf\n", &data[i]);    }
   fclose(fp);
   return i;}

long sellfunc(double currentprice, double sigmasell, double* sale, long* nsell)
{  long nselln=0, i;
   double salen = 0;

            if (nbuy == 0) { printf("no buyhist, quit...");   exit(0); }

 for (i=0;i<nbuy;i++) {
      if (currentprice>buyhist[i]*(1+sigmasell)) {
         nselln++;  salen = salen + buyhist[i];  }}
   *nsell = nselln; *sale = salen;
    return nselln;
}

long longmax(long a, long b) {
   if (a>b) {  return a;   }
        else {return b;    }}

double doublemax(double a, double b) {
   if (a>b) {return a; }
        else {return b;    }}

double sumbuyhist() {
   double sump=0;
   long i;
   for (i=0;i<nbuy;i++) { sump+=buyhist[i];    }
   return sump; }
```