Hong Kong University of Science of Technology

Final Project

MAFS5130-Qnautitative Analysis of Financial Time Series

Science or Myth:

Could Technical Indicators Predict Markets?

By

CONG Mingshu, 20006892
CHEN Wanjun, 20001373
ZHAO Jihong, 07627388
SHI Luyao, 20005771
ZHANG Yuting, 20005109

Instructor: Dr. Ling Shiqing

May 2012

Contents

Figure Indexes

Table Indexes

Introduction

The view on technical analysis varies from person to person. Some claim it as a science while others blame it as myth. In this article, we try to examine the effect of those popular technical indicators based on the historical data of HSI. We try to answer the question: could technical indicators predict markets?

Technical Analysis Overview

Fundamental Analysis & Technical Analysis

While investors have hundreds of methods to predict the market and finally make investment decisions, those methods fall into two very broad categories: fundamental analysis and technical analysis. Fundamental analysis involves analyzing the characteristics of a company in order to estimate its value and it is the basic method used by researchers from investment banks in their research reports. On the other hand, technical analysis takes a completely different approach; it doesn't care one bit about the so called "intrinsic value" of a company or a commodity. Technicians are only interested in the price movements in the market.

Basically, technical analyst is trying to predict the market through historical trading data, among which closing prices and the trading volumes are of the most important. Stocks can be seen as a special commodity, whose price direction or trend in the near future can be determined by the relative strength of its demand to its supply. Technical analysis attempts to understand the emotions in the market by studying the market itself, as opposed to its components.

Technical analysis is widely used in Asia, especially by those individual investors, called *Sanhu* in Mandarin. There are popular blogs online analyzing the market by technical analysis technique, and some Chinese investors believe in predictions in these blogs more than research reports from famous investment banks. Some analysis in these blogs is quite fancy. For example, one blogger believes there is some supernatural relationships between the market mood and the 24 seasonal segments in Chinese traditional lunar calendar. Notwithstanding tons of critics, all those bloggers allege their analysis is “totally scientific”. Even in Hong Kong, the most important financial center in East Asia, technical analysis is popular. Even the articles on Mingpao, one of the most influential local newspapers, mention technical analysis techniques from time to time.

Despite notorious among professional bankers as unreasonable, technical analysis is still widely used by traders for short term price prediction. The common view on technical analysis among investment bankers is that while it cannot value a stock properly, it is useful for short term trading.

Market efficiency

The efficient-market hypothesis (EMH) may be the most important hypothesis in modern finance. Asserting that financial markets are "informationally efficient", one cannot consistently achieve returns in excess of average market returns on a risk-adjusted basis, given the information available at the time the investment is made.

There are three major versions of the hypothesis: "weak", "semi-strong", and "strong". The weak-form EMH claims that prices on traded assets (e.g., stocks, bonds, or property) already reflect all past publicly available information. The semi-strong-form EMH claims both that prices reflect all publicly available information and that prices instantly change to reflect new public information. The strong-form EMH additionally claims that prices instantly reflect even hidden or "insider" information.

The market efficiency concept has lots to do with fundamental analysis and technical analysis. In fact, there is another way to define efficient market. A “weakly efficient market” is a market in which technical analysis will not work. A “semi-strong efficient market” is a market in which fundamental analysis will not work.

Since EMH is the cornerstone of almost all modern financial theories, testing the market efficiency has aroused lots of academic interests. One of the approaches for this purpose is to test the effectiveness of technical or fundamental analysis. This is also another purpose of this paper.

Technical Analysis & Behavioral Finance

Behavioral finance is a field of finance that proposes psychology-based theories to explain stock market anomalies. Within behavioral finance, it is assumed that the information structure and the characteristics of market participants systematically influence individuals' investment decisions as well as market outcomes.

In recent years, it has become more and more obvious that psychology plays an ever-more important role in financial markets and also drives back the influence on the rational actions of stock market participants. In broad terms, it argues that some financial phenomena can be better understood using models in which some agents are not fully rational. Behavioral Finance is showing that in an economy where rational and irrational traders interact, irrationality can have a substantial and long-lived impact on prices.

From this point of view, technical analysis may be effective in the sense that if every market participator believes in those well known technical techniques, those techniques will actually show its effect.

Categories of Technical Analysis

Candlestick Chart

Candlestick chart, also called K-chart in Asia, is a style of bar-chart used primarily to describe price movements of a security, derivative, or currency over time.

Candlesticks are usually composed of the body (black or white), and an upper and a lower shadow (wick): the area between the open and the close is called the real body, price excursions above and below the real body are called shadows. The wick illustrates the highest and lowest traded prices of a security during the time interval represented. The body illustrates the opening and closing trades. If the security closed higher than it opened, the body is white or unfilled, with the opening price at the bottom of the body and the closing price at the top. If the security closed lower than it opened, the body is black, with the opening price at the top and the closing price at the bottom. A candlestick need not have either a body or a wick.

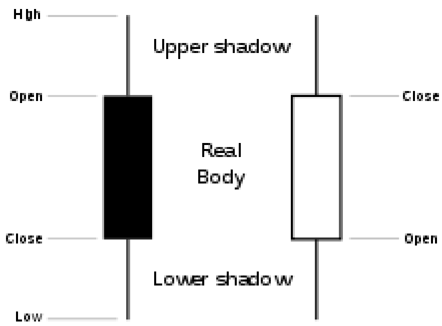

Figure 1: The basic candlestick

There are popular candlestick patterns used for predicting the markets, such as *three white soldiers*, *three black crows*, *morning star*, etc.

Elliott Wave Theory

The Elliott Wave Principle is a form of technical analysis that some traders use to analyze financial market cycles and forecast market trends by identifying extremes in investor psychology, highs and lows in prices, and other collective factors.

In Elliott's model, market prices alternate between an impulsive, or motive phase, and a corrective phase on all time scales of trend, as the illustration shows. Impulses are always subdivided into a set of 5 lower-degree waves, alternating again between motive and corrective character, so that waves 1, 3, and 5 are impulses, and waves 2 and 4 are smaller retraces of waves 1 and 3. Corrective waves subdivide into 3 smaller-degree waves starting with a five-wave counter-trend impulse, a retrace, and another impulse. In a bear market the dominant trend is downward, so the pattern is reversed—five waves down and three up. Motive waves always move with the trend, while corrective waves move against it.

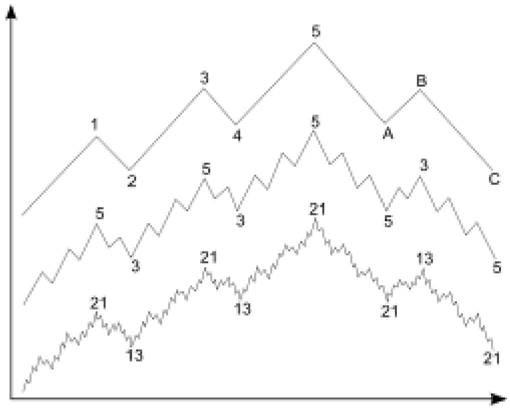

Figure 2: Wave illustration from R.N. Elliott's essay, "The Basic of Wave Principles"

Chart Pattern Theory

A chart pattern is a pattern that is formed within a chart when prices are graphed. In stock and commodity markets trading, chart pattern studies play a large role during technical analysis. When data is plotted there is usually a pattern which naturally occurs and repeats over a period. Chart patterns are used as either reversal or continuation signals.

Examples of "classical" chart patterns widely by traders and investors include: Head and shoulders, trend lines, cup and handle, double top and double bottom, triple top and triple bottom, broadening top, price channels, wedge pattern, triangle, flag and pennant patterns, etc.

Figure 3: Illustration of the pattern *head and shoulders*

Technical Indicators

The above three methods are mainly based on the patterns of the stock price chart. Technical Indicator is a result of mathematical calculations based on indications of price and/or volume. The values obtained are used to forecast probable price changes. Because these indicators have exact values for any given time, they are more suitable for quantitative analysis and computational trading.

Therefore, technical Indicators are the main focus of this paper.

Main Categories of Technical Indicators

According to the underlying trading data involved, the technical indicators can be divided into three main categories.

Price-based indicators

- *Average Directional Index* — a widely used indicator of trend strength
- *Commodity Channel Index* — identifies cyclical trends
- *MACD* — moving average convergence/divergence
- *BIAS* — the departure against the moving average

- *PSY* — the market mood indicator
- *Momentum* — the rate of price change
- *Relative Strength Index (RSI)* — oscillator showing price strength
- *Stochastic oscillator* — close position within recent trading range
- *Trix* — an oscillator showing the slope of a triple-smoothed exponential moving average
- *%C* — denotes current market environment as range expansion or contraction plus highlights ta extremes when the condition should be changing.

Volume-based indicators

- *Accumulation/distribution index* — based on the close within the day's range
- *Money Flow* — the amount of stock traded on days the price went up
- *On-balance volume (OBV)* — the momentum of buying and selling stocks

Breadth Indicators

These indicators are based on statistics derived from the broad market.

- *Advance Decline Line* — a popular indicator of market breadth
- *McClellan Oscillator* - a popular closed-form indicator of breadth
- *McClellan Summation Index* - a popular open-form indicator of breadthData Set Selection & Description

Considering their popularity in China, we choose three price-based indictors (MACD, BIAS, PSY) and one volume-based indicators (OBV) in our study.

Introduction of HSI

History & Components

The Hang Seng Index (HSI) is a free float-adjusted market capitalization-weighted stock market index in Hong Kong. It is used to record and monitor daily changes of the largest companies of the Hong Kong stock market and is the main indicator of the overall market performance in Hong Kong. These 48 constituent companies represent about 60% of capitalisation of the Hong Kong Stock Exchange.

HSI was started on November 24, 1969, and is currently compiled and maintained by Hang Seng Indexes Company Limited, which is a wholly owned subsidiary of Hang Seng Bank, one of the largest banks registered and listed in Hong Kong in terms of market capitalization. It is responsible for compiling, publishing and managing the Hang Seng Index and a range of other stock indexes, such as Hang Seng China Enterprises Index, Hang Seng China AH Index Series, Hang Seng China H-Financials Index, Hang Seng Composite Index Series, Hang Seng China A Industry Top Index, Hang Seng Corporate Sustainability Index Series and Hang Seng Total Return Index Series. Hang Seng in turn, despite being a public company, is held in majority by British financial firm HSBC

Figure 4: Candlestick chart for Hang Seng Index for the past 8 months

In our following study, we use the daily closing prices of HSI from Jan 1st, 2005 to Dec 31st, 2011.

HSI as a tradable asset

Unlike ordinary stocks, HSI cannot be traded directly. However, it can also be a tradable asset through the financial product HSI ETF. Basically, this is a fund tracking the movement of HSI by holding the component stocks. Therefore, if we can predict HSI correctly, we can make profit by trading HSI ETFs.

Unit Price of HSI ETF	Determined by the forces of market supply and demand (Not applicable to in-kind transactions conducted through Participating Dealers)
Trading Channel	Trade through stock brokers or banks on the SEHK
Trading Hours	Traded during SEHK trading hours
Minimum Trading Size	100 Units
Fees for acquisition and disposal of Units	- Brokerage# - SFC Transaction Levy - SFC Investor Compensation Levy - SEHK Trading Fee - Stamp Duty - Purchase Order Handling Fee#(if any)
Fund Operating Cost	Management Fee, Trustee Fee and Other Expenses*

Table 1: Trading details of HSI ETF offered by Hang Seng Bank

Data Description

Now we analyses the data of the HSI index from 2005 to 2011, as a preparation for the further study of the relationship between the HSI and the technical indicator. Frist we check if the In(HSI) are stationary through the DF test, and make some differentiation if the data is not stationary. Then we check if the data is serial correlated through the ACF and Q-test. At last we check if the ARMA model is fitted for the transformed data. The detail SAS result can see in the appendix A.

Stationary checking---Dickey-Fuller Unit Root Tests

Dickey-Fuller Unit Root Test is conducted to confirm two series are stationary and the results are listed in Table bellowed. The second column of Table specifies three types of models, namely Zero Mean, Single Mean, and Trend. The third column (Rho) and the fifth column (Tau) are the test statistics for the unit root tests. The remaining columns (column 2 and column 6) are the p-values for corresponding models.

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	**0**	0.0273	0.6894	0.36	0.7887		
	1	0.0270	0.6893	0.37	0.7907		
	2	0.0279	0.6895	0.38	0.7927		
	3	0.0286	0.6897	0.40	0.7999		
Single Mean	**0**	-7.9263	0.2214	-2.19	0.2114	2.47	0.4366
	1	-7.4898	0.2455	-2.13	0.2330	2.35	0.4676
	2	-7.7317	0.2319	-2.17	0.2183	2.44	0.4454
	3	-7.2252	0.2613	-2.12	0.2376	2.34	0.4701
Trend	**0**	-8.4065	0.5527	-2.02	0.5909	2.42	0.6924
	1	-7.8546	0.5958	-1.94	0.6320	2.28	0.7195
	2	-8.1162	0.5752	-1.98	0.6123	2.37	0.7025
	3	-7.4187	0.6304	-1.89	0.6594	2.24	0.7269

Table 2: DF test for the lnHSI

From the result of the DF-test of the ln(HSI), we can see that the p values of the Tau are all larger than 0.1, that mean the data is not stationary and we should take a differentiation of the data. The first order of the differentiation of the data is showed bellowed. Actually ΔlnHSI is the log return of the Heng seng index, and we believed it is more practical.

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	**0**	-1774.14	0.0001	-42.83	<.0001		
	1	-1734.94	0.0001	-29.44	<.0001		
	2	-2004.37	0.0001	-25.13	<.0001		
	3	-2390.45	0.0001	-22.32	<.0001		
Single Mean	**0**	-1774.30	0.0001	-42.82	<.0001	916.71	0.0010
	1	-1735.45	0.0001	-29.43	<.0001	433.11	0.0010
	2	-2005.72	0.0001	-25.13	<.0001	315.67	0.0010
	3	-2393.59	0.0001	-22.32	<.0001	249.00	0.0010
Trend	**0**	-1775.08	0.0001	-42.83	<.0001	917.01	0.0010
	1	-1737.78	0.0001	-29.44	<.0001	433.43	0.0010
	2	-2011.63	0.0001	-25.14	<.0001	316.09	0.0010
	3	-2407.77	0.0001	-22.34	<.0001	249.49	0.0010

Table 3: DF test for the ΔlnHSI

From the Table 4, we can see that all the p-values of the log return of the HIS are less than 0.05, that mean the ΔlnHSI series is stationary (no drift and trend component) at the significant level of 5%.

Serial correlation checking

<table>
<tr><th colspan="5">Autocorrelations</th></tr>
<tr><th>Lag</th><th>Covariance</th><th>Correlation</th><th>-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1</th><th>Std Error</th></tr>
<tr><td>0</td><td>3.257820</td><td>1.00000</td><td>| |********************|</td><td>0</td></tr>
<tr><td>1</td><td>-0.108728</td><td>-.03337</td><td>| *|. |</td><td>0.024126</td></tr>
<tr><td>2</td><td>0.038679</td><td>0.01187</td><td>| .|. |</td><td>0.024153</td></tr>
<tr><td>3</td><td>-0.155472</td><td>-.04772</td><td>| *|. |</td><td>0.024156</td></tr>
<tr><td>4</td><td>-0.115761</td><td>-.03553</td><td>| *|. |</td><td>0.024211</td></tr>
<tr><td>5</td><td>-0.011145</td><td>-.00342</td><td>| .|. |</td><td>0.024242</td></tr>
<tr><td>6</td><td>-0.025084</td><td>-.00770</td><td>| .|. |</td><td>0.024242</td></tr>
<tr><td>7</td><td>0.060837</td><td>0.01867</td><td>| .|. |</td><td>0.024243</td></tr>
<tr><td>8</td><td>0.115599</td><td>0.03548</td><td>| .|* |</td><td>0.024252</td></tr>
<tr><td>9</td><td>-0.169436</td><td>-.05201</td><td>| *|. |</td><td>0.024282</td></tr>
<tr><td>10</td><td>-0.217853</td><td>-.06687</td><td>| *|. |</td><td>0.024347</td></tr>
<tr><td>11</td><td>0.073055</td><td>0.02242</td><td>| .|. |</td><td>0.024453</td></tr>
<tr><td>12</td><td>0.090969</td><td>0.02792</td><td>| .|* |</td><td>0.024465</td></tr>
</table>

Table 5: ACF of the ΔlnHSI

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	8.38	6	0.2114	-0.033	0.012	-0.048	-0.036	-0.003	-0.008
12	25.80	12	0.0115	0.019	0.035	-0.052	-0.067	0.022	0.028

Table 6: Q-test of the ΔlnHSI

From the ACF, we can see that the sample has no serial correlations. But the Ljung-Box statistics give that Q(12)=25.80 with a p-value 0.0115, suggesting that we can reject the H0 that the first 12 lags of ACF are zero. That mean the data is serial correlated and we can try to fit an ARMA model for the percentage log return of the HSI.

Fitting the ARMA model.

We try many model, the AR(1), AR(2), MA(1), MA(2), ARMA(1,1), ARMA(2,2), etc. But the results are all not fitted as the p-value of the corresponding parameters are all larger than 0.1, which mean at the significant level of 10%, the parameters are all not significantly different from zero. As a conclusion, we cannot model the HSI with the ARMA model.

On Balance Volume indicator(OBV)

Overview of OBV indicator

On Balance Volume (OBV) measures buying and selling pressure as a cumulative indicator that adds volume on up days and subtracts volume on down days. OBV was developed by Joe Granville and introduced in his 1963 book, Granville's New Key to Stock Market Profits. It was one of the first indicators to measure positive and negative volume flow. We can look for divergences between OBV and price to predict price movements or use OBV to confirm price trends.

The On Balance Volume (OBV) line is simply a running total of positive and negative volume. A period's volume is positive when the close is above the prior close. A period's volume is negative when the close is below the prior close.

$$\mathrm{OBV(t)} = \begin{cases} \mathrm{OBV(t)} = \mathrm{OBV(t-1)} + \mathrm{Volume(t)}, & when\ \Delta lnHSI(t) > 0 \\ \mathrm{OBV(t)} = \mathrm{OBV(t-1)} - \mathrm{Volume(t)}, & when\ \Delta lnHSI(t) < 0 \\ \mathrm{OBV(t)} = \mathrm{OBV(t-1)}, & when\ \Delta lnHSI(t) = 0 \end{cases}$$

OBV rises when volume on up days outpaces volume on down days. OBV falls when volume on down days is stronger. A rising OBV reflects positive volume pressure that can lead to higher prices. Conversely, falling OBV reflects negative volume pressure that can foreshadow lower prices. OBV would often move before price. Expect prices to move higher if OBV is rising while prices are either flat or moving down. Expect prices to move lower if OBV is falling while prices are either flat or moving up.

On Balance Volume (OBV) is a simple indicator that uses volume and price to measure buying pressure and selling pressure. Buying pressure is evident when positive volume exceeds negative volume and the OBV line rises. Selling pressure is present when negative volume exceeds positive volume and the OBV line falls. We can use OBV to confirm the underlying trend or look for divergences that may foreshadow a price change.

Data Checking

Obviously the OBV is a serial correlated and unstationary time series, so when we study the relationship of the HSI and OBV indicator, we use the transformed OBV as:

$$\Delta\mathrm{OBV}(t) = \mathrm{OBV}(t) - \mathrm{OBV}(t-1) = \mathrm{sgn}(\mathrm{OBV}(t) - \mathrm{OBV}(t-1)) \times \text{Volume(t)}$$

As the order of magnitude between the ΔlnHSI and ΔOBV is quite a big different, we transformed the dOBV further as:

$$\Delta\mathrm{OBV}'(t) = \Delta\mathrm{OBV}(t)/10^8$$

Next we check if the data of $\Delta\mathrm{OBV}'(t)$ is stationary through the DF-test is SAS:

Dickey-Fuller Unit Root Tests					
Variable	**Type**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**
dlnHSI	**Zero Mean**	-1743.0	0.0001	-29.51	<.0001
	Single Mean	-1743.4	0.0001	-29.50	<.0001
	Trend	-1745.2	0.0001	-29.51	<.0001
dOBV	**Zero Mean**	-1671.2	0.0001	-28.89	<.0001
	Single Mean	-1676.5	0.0001	-28.93	<.0001
	Trend	-1677.5	0.0001	-28.93	<.0001

Table 7: DF test for the ΔOBV

From the Table6, we can see that all the p-values of the dOBV' are less than 0.05, that mean the ΔOBV series is stationary (no drift and trend component) at the significant level of 5%.

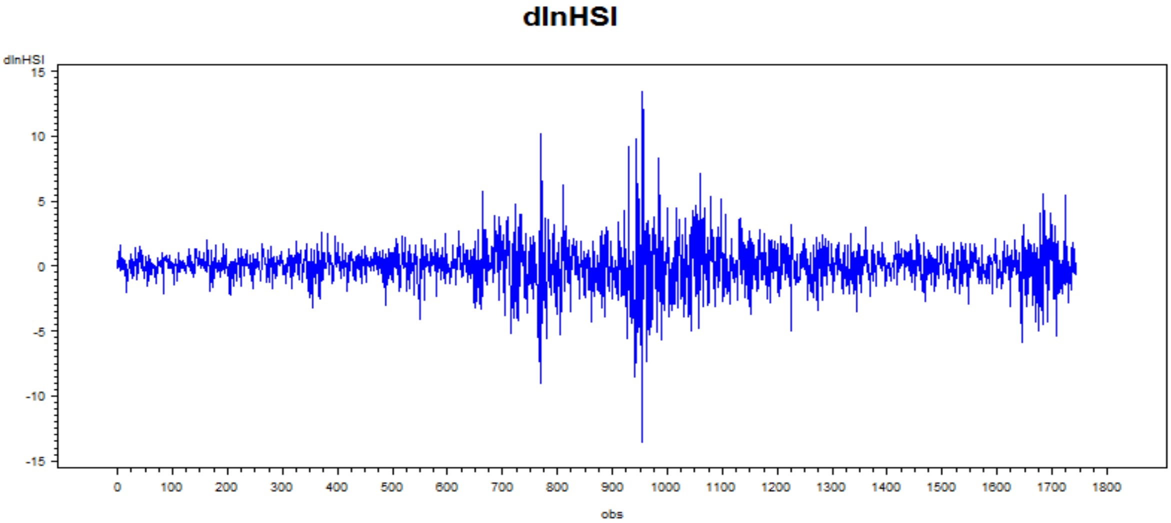

Figure 5: the time series of dlnHSI

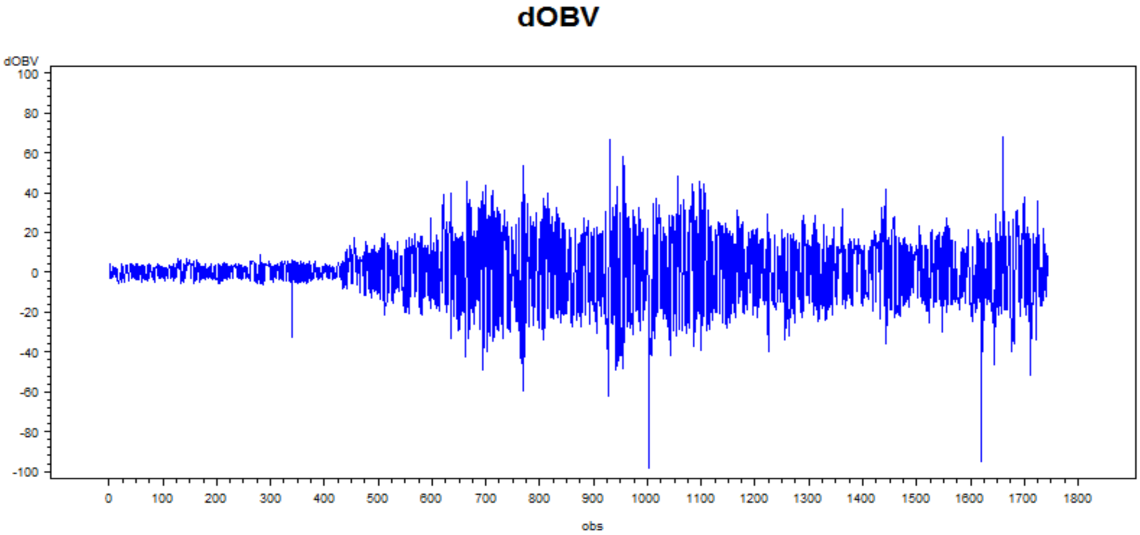

Figure 6: the time series of dOBV

Model fitting

SAS Variable Name	Description
dlnHSI	The number of holidays in Spain taken by U.K. residents
d_OBV	U.K. real personal disposable income

Table 8: variable used in the SAS test

We tried different parameters in VARMA model to fit our log return of HSI and ΔOBV. We tried VAR(1), VAR(2), VAR(3), VARMA(1,1), VARMA(2,2), models and found that VARMA(1,1) is the best one. VARMA(11,) shows that there is significant relationship between log return of HSI and ΔOBV.

The VARMAX Procedure

Type of Model	VARMA(1,1)
Estimation Method	Maximum Likelihood Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
dlnHSI	**CONST1**	1.09004	0.64374	1.69	0.0906	1
	AR1_1_1	3.69917	0.47270	7.83	0.0001	dlnHSI (t-1)
	AR1_1_2	-1.83658	0.13425	-13.68	0.0001	dOBV(t-1)
	MA1_1_1	3.78239	0.47464	7.97	0.0001	e1(t-1)
	MA1_1_2	-1.84108	0.13419	-13.72	0.0001	e2(t-1)
dOBV	**CONST2**	2.66670	1.57101	1.70	0.0898	1
	AR1_2_1	6.83499	1.76920	3.86	0.0001	dlnHSI (t-1)
	AR1_2_2	-3.49872	0.41650	-8.40	0.0001	dOBV (t-1)
	MA1_2_1	7.00441	1.79618	3.90	0.0001	e1(t-1)
	MA1_2_2	-3.50858	0.41773	-8.40	0.0001	e2(t-1)

Table 9: Parameter Estimates of the VARMA(1,1) Model

From the above result, we see that the VARMA(1,1) model for ΔlnHIS(percentage) and the transformed indicator ΔOBV is

$$\Delta \ln HSI(t) = 1.09 + 3.70\Delta \ln HIS\%(t-1)$$

$$-1.84 \times \Delta OBV(t-1) + 3.78a_1(t-1) - 1.84a_2(t-1)$$

Since the p-values for all coefficients in the above equation are less than 0.1, which confirm that the coefficients are significant at the level of 10%. The result shows that when the trading volume of stocks of the HSI increase, the Heng Seng Index tend to has a smaller log return.

Findings

As a normal sense, if the OBV is down means money is drawing out from that market, and if OBV is up means money is pumping into that market. So the log return of the HSI should increase or decrease along with the return of the OBV in a common sense. Also we know that price follows the money flow. But what we get from the VARMA model is against our common sense, which suggests that When the OBV increases, it leads to a decrease in the log return of the HSI.

Moving Average Convergence/Divergence (MACD)

Overview of the indicator

The standard setting for MACD is the difference between the 12 and 26-period EMAs. Mathematically,

$$MACD_t = EMA_t\,(12) - EMA_t(26)$$

where EMA is a type of moving average that is similar to a simple moving average, except that more weight is given to the latest data. In terms of formula,

$$EMA_t(n) = EMA_{t-1}(n) + \frac{2}{n+1} \times (HSI_t - EMA_{t-1}(n))$$

By taking difference of the longer moving average from the shorter moving average, the MACD actually turns two trend-following indicators into a momentum oscillator, which measures the rate of change of our HSI Index. Positive values of it increase as $EMA_t\,(12)$ diverges further from $EMA_t\,(26)$, meaning an increasing upside momentum.

Since MACD is calculated from HSI, its value should be a good reflection of HSI itself. Meanwhile due to its own popularity in technical analysis for stock trading, we have good reasons to believe that it actually reversely affect the level of HSI, thus forming a two-way interaction between the two.

Data Checking

Before performing any specific modeling for the purpose of analysis, we ran a Dickey-Fuller Test on the log transformation of MACD to see if it is stationary.

We choose a log transformation of MACD instead of MACD itself only to make it better compatible with the log transformation we made to our HSI data. Yet unfortunately, as suggested by a p-value greater than 0.05 in the following table, the null hypothesis that the log transformation of MACD is non-stationary can hardly be rejected.

Dickey-Fuller Unit Root Tests					
Variable	**Type**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**
Ln(MACD)	**Zero Mean**	-0.05	0.6712	-0.15	0.6333
	Single Mean	-78.32	0.0019	-6.26	<.0001
	Trend	-79.53	0.0008	-6.31	<.0001

Table 10： DF Test for log transformation of MACD

A natural way to address the problem is to take a one-time difference of the log MACD, which is what we did in our case.

Similarly Dickey-Fuller Unit Root Test is performed to test for the stationarity of this new series. With all the p-values smaller than 0.05, the stationarity of the one-time difference of log transformation of MACD is guaranteed.

Dickey-Fuller Unit Root Tests					
Variable	**Type**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**
Dif(Ln(MACD))	**Zero Mean**	-618.89	0.0001	-17.58	<.0001
	Single Mean	-618.89	0.0001	-17.58	<.0001
	Trend	-618.91	0.0001	-17.57	<.0001

Table 11： DF Test for one-time difference of log MACD

Model Specification - VARMA (1, 1)

As explained earlier, we expect a mutual interaction between HSI and MACD, or at least we are here to test about such correlation. So a simple regression model of HSI return on the one time difference of log MACD and other terms will not be likely to capture the general picture due to its hidden assumption of causality of the dependent and independent variables.

A general VARMA model on the other hand, will be more relevant in the literature. It is able to capture the two-way effect of HSI and MACD level on each other, if such effect actually exists.

The cross correlations of residuals are presented as following. Notice that the correlations are only significant up to first lag, suggesting a model specification of VARMA (1, 1)

Schematic Representation of Cross Correlations of Residuals

Variable/Lag	0	1	2	3	4	5	6	7	8	9	10	11	12
R	++	++	..	..	..	..	..	..	..	--	--	..	..
Dif(Ln(MACD))	++	++	..	--	..	..	..	..	++	--	--	.+	..

+ is > 2*std error, - is < -2*std error, . is between

Table 12： Cross Correlation of Residuals for VARMA(1,1) Model

Different specifications are still performed and VARMA(1, 1) Model indeed has a best fit to our data. Main results of the estimation using this model are presented in the following tables, which include Model Parameter Estimates, Covariances of Innovations and Information Criteria.

P-values for all parameter estimates are less than 0.05, suggesting significance of the estimates at 5% significance level.

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
R	AR1_1_1	-0.30710	0.08485	-3.62	0.0003	R(t-1)
	AR1_1_2	-6.88926	1.54814	-4.45	0.0001	Dif(Ln(MACD)) (t-1)
	MA1_1_1	-0.39308	0.10144	-3.87	0.0001	e1(t-1)
	MA1_1_2	10.88581	2.60234	4.18	0.0001	e2(t-1)
Dif(Ln(MACD))	AR1_2_1	-0.02488	0.00536	-4.64	0.0001	R(t-1)
	AR1_2_2	0.35097	0.07800	4.50	0.0001	Dif(Ln(MACD)) (t-1)
	MA1_2_1	-0.02548	0.00513	-4.97	0.0001	e1(t-1)
	MA1_2_2	-0.23914	0.09496	-2.52	0.0119	e2(t-1)

Table 13： Parameter estimates for VARMA(1,1) Model

Covariances of Innovations		
Variable	R	dlnmacd
R	3.22425	0.03452
dlnmacd	0.03452	0.00047

Table 14： Covariances of Innovation for VARMA(1,1) Model

Information Criteria	
AICC	-8.04577
HQC	-8.0364
AIC	-8.04579
SBC	-8.02041
FPEC	0.00032

Table 15： Information Criteria for VARMA(1,1) Model

Model checking: (undecided yet)

Findings:

(Fitted model, financial intuition, possible application)

With the above analysis, we are able to obtain the following fitted model for HSI return and one time difference of log MACD:

$$R_t = -0.30710\, R_{t-1} - 6.88926\, \mathbf{Dif}(\mathbf{Ln}(\mathbf{MACD}))_{t-1} + a_t^1 - 0.39308\, a_{t-1}^1 + 10.88581\, a_{t-1}^2$$

Coefficient before $\mathbf{Dif}(\mathbf{Ln}(\mathbf{MACD}))_{t-1}$ is significantly different from 0, suggesting that **Dif(Ln(MACD))** in the previous period does make an impact on the HSI return at the current time. This coincides with our prediction that the collateral use of MACD for technical analysis in the market place actually creates a reversal impact on the stock price itself.

The negative sign of coefficient appears to be controversial against financial intuition.

From intuition when **Dif(Ln(MACD))** takes a positive value, MACD is increasing and we should expect to see an increase in Rt as well. However the negative coefficient before **Dif(Ln(MACD))** right gives an opposite direction.

A possible explanation could be that since MACD is a relatively long-term indicator, it is somehow lagging behind the stock price movement. Thus in market place the traders may take an opposite move to capture such arbitrage opportunity. Again, it suggests that the use of financial indicator for technical analysis, although make full sense, can be very complicated and subjective with individuals in real practice.

BIAS

Overview of the indicator

BIAS, or Y Value, measures the divergence of current stock price (in our case, the HSI) from an N-day simple moving average of the stock prices. Mathematically,

$$BIAS_t = \left(HSI_t - \frac{1}{N} \sum_{i=t-N+1}^{t} HSI_i \right) \div \frac{1}{N} \sum_{i=t-N+1}^{t} HSI_i \times 100\%$$

where N is chosen to be 6 days in our case.

Technical analysis using BIAS is based on the assumption that stock price (in our case the log HSI), may it being above or below the moving average, if happens to be too far away from that average, will always tend to converge back from a longer perspective. This therefore can provide useful inferences for trading strategies of the stock.

Again since BIAS is calculated from HSI, its value should be a good reflection of HSI itself. Meanwhile given its popularity in technical analysis for stock trading, we have good reasons to believe that it actually reversely influence the level of HSI, thus forming a two-way interaction between the two.

Data Checking

Before performing any specific modeling for the purpose of analysis, we ran a Dickey-Fuller Test on BIAS to see if it is stationary. Relevant results are provided in the following table. With all the p-values smaller than 0.05, we can assume stationarity of the data.

Dickey-Fuller Unit Root Tests					
Variable	**Type**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**
BIAS	**Zero Mean**	**-649.29**	**0.0001**	**-18.01**	**<.0001**
	Single Mean	**-649.79**	**0.0001**	**-18.01**	**<.0001**
	Trend	**-651.67**	**0.0001**	**-18.03**	**<.0001**

Table 16： DF Test for BIAS

Model Specification - VARMA (1, 1)

As explained, we expect mutual interaction between HSI and BIAS. So for similar reason as that in the case of MACD, a simple regression model of HSI return on BIAS and other variables will not give an accurate estimate. A general VARMA model is instead adopted to capture the two-way effect of HSI and BIAS, if it really exists.

Different specifications are performed before we find a best fitted model, the VARMA(1, 1) Model. Main results of the estimation using this model are presented in the following tables, which include Model Parameter Estimates, Covariances of Innovations and Information Criteria.

P-values for all parameter estimates are less than 0.1, suggesting significance of the estimates at 10% significance level.

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
R	AR1_1_1	0.31229	0.07005	4.46	0.0001	R(t-1)
	AR1_1_2	-0.06859	0.01970	-3.48	0.0005	BIAS(t-1)
	MA1_1_1	0.28023	0.05043	5.56	0.0001	e1(t-1)
	MA1_1_2	-0.04069	0.02321	-1.75	0.0797	e2(t-1)
BIAS	AR1_2_1	2.28695	0.40050	5.71	0.0001	R(t-1)
	AR1_2_2	0.70266	0.04961	14.16	0.0001	BIAS(t-1)
	MA1_2_1	1.61577	0.40686	3.97	0.0001	e1(t-1)
	MA1_2_2	0.70853	0.05386	13.16	0.0001	e2(t-1)

Table 17： Model Parameter Estimates for VARMA(1,1) Model

Covariances of Innovations		
Variable	R	BIAS
R	0.00033	0.00027
BIAS	0.00027	0.00022

Table 17： Covariances of Innovations for VARMA(1,1) Model

Information Criteria	
AICC	**-21.1231**
HQC	**-21.1137**
AIC	**-21.1231**
SBC	**-21.0976**
FPEC	**6.7E-10**

Table 188: Information Criteria for VARMA (1,1) Model

Conclusion:

(Fitted model, financial intuition, possible application)

With the above analysis, we are able to obtain the following fitted model for HSI return and BIAS:

$$R_t = 0.31229\, R_{t-1} - 0.06859\, BIAS_{t-1} + a_t{}^1 + 0.28023\, a_{t-1}^1 - 0.04069\, a_{t-1}^2$$

Notice that the coefficient before $BIAS_{t-1}$ is significantly different from 0, suggesting that BIAS in the previous period does make an impact on the HSI return at the current time. This coincides with our prediction that the collateral use of BIAS for technical analysis in the market place actually create a reversal impact on the stock price itself.

In addition, this coefficient is estimated to be negative, which is greatly in line with the theory of mean-reversing, suggesting that when the stock price is sufficiently lower than average, the negative BIAS will serve to push the price up in the next period, so as to make it gradually converge to average.

Other Unsuccessful Attempts

Besides above three indicators, we have tried many other technical indexes to find a good shot; however the others' performances are not satisfying. In another way, choice of an appropriate index is the key in technical analysis. The failed ones include KDJ and PSY, of which the similar analysis procedures are shown.

KDJ

The Chinese name of KDJ index is random index, was first proposed by George Lane. KDJ is most often used in short-term analysis by calculating Raw Stochastic Value (RSV), K line, D line and J line. J line comprehends the K line and D line, thus, is the index we choose to use.

(1) Computation Formulae:

$$RSV_t = \frac{P_t - LP_n}{HP_n - LP_n} \times 100$$

$$K_t = \frac{2}{3}K_{t-1} + \frac{1}{3}\mathrm{RSV}_t$$

$$D_t = \frac{2}{3}D_{t-1} + \frac{1}{3}\mathrm{K}_t$$

$$J_t = 3D_t - 2\mathrm{K}_t$$

P_t=Today's closing price

LP_n= The lowest closing price of the last n days, n=9 in this study

HP_n= The highest closing price of the last n days

$$\mathrm{K}_0 = D_0 = 50$$

(2) Stationary Test

Figure 7 obviously shows that J line fluctuate steady around a mean greater than zero. So, we have an intuition that j line is stationary.

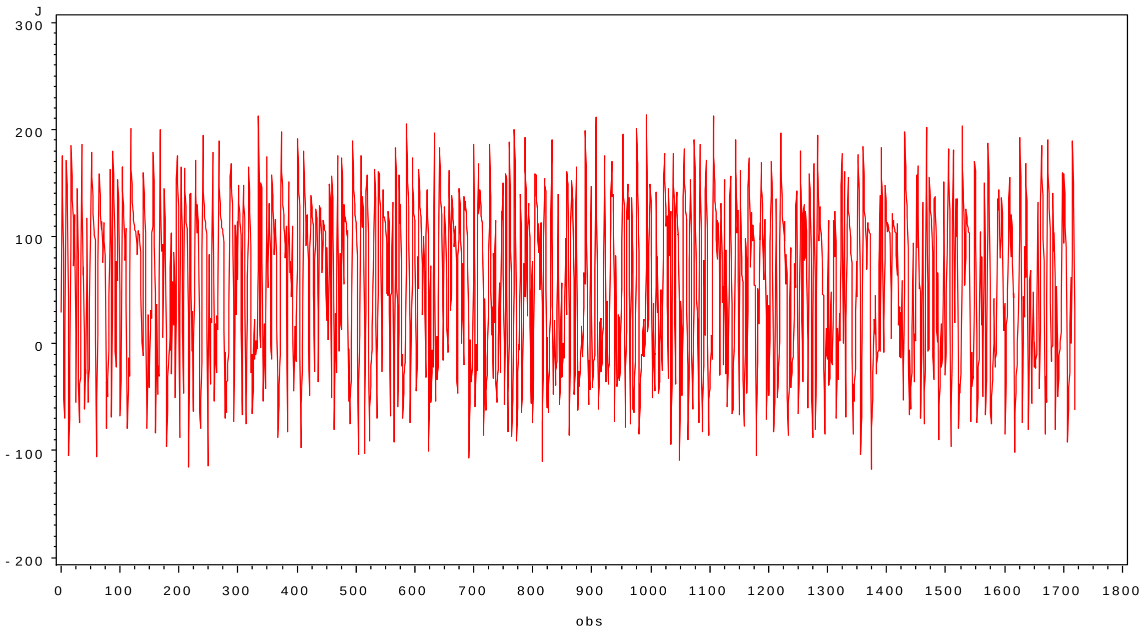

Figure 7 Plot of J Index

Also, referring to Table 19, all p-values are significantly smaller than 0.05. By the Dickey-Fuller test, all the three tests on equations respectively with trend term, single mean term or zero mean reject the null hypothesis that the time series has a unit root. Hence, we can consider the J index is a stationary series.

Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
J	Zero Mean	-730.88	0.0001	-19.13	<.0001
	Single Mean	-1383.3	0.0001	-26.26	<.0001
	Trend	-1390.3	0.0001	-26.32	<.0001

Table 19 Unit Root Test of J index

(3) Parameters Estimation

After extensive trials, we still cannot find a good enough model of HSI log return and J index. One result of them is like Table 20. The estimators in equation 1 which HSI log return is the dependent variable in are not as significant as the ones in the equation 2 where J index as the dependent variable.

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	AR1_1_1	-0.03814	0.02472	-1.54	0.1231	r(t-1)
	AR1_1_2	0.00000	0.00000	0.90	0.3668	J(t-1)
J	AR1_2_1	1801.67572	69.42511	25.95	0.0001	r(t-1)
	AR1_2_2	0.69148	0.01341	51.55	0.0001	J(t-1)

Table 20Estimation of VAR(1) model of HSI log return and J index

It should be understood in the way that J index is calculated from HSI log return through some straight forward transforms while the return cannot be expressed by the simple way.

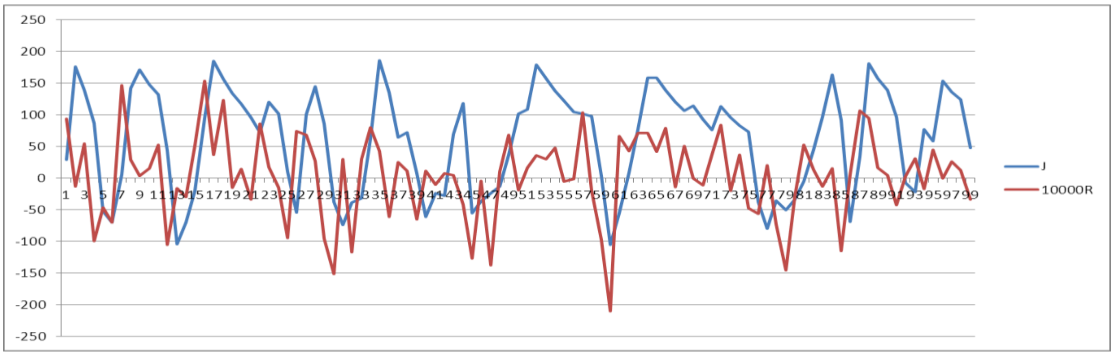

Figure 8 Plot of J Index and 10000 times HSI Log Return (A Cut Period)

Also, from Figure 8, it is observed that J line almost follows the volatility of the log return in some lagged time periods. Under this circumstance, lagged return fits the J line very well, while having the two exchanged is not the case.

PSY

The Psychological Line is a "sentiment indicator". As for indicators of this type, an attempt is made to look behind the obvious mood of the market and to detect undertones for a trend change. PSY is the ratio of the number of rising periods over the total number of periods. It reflects the buying power in relation to the selling power.

(1) Computation Formulae

$$PSY(N) = A \div N \times 100$$

The parameter N describes the observation period. It can assume any value between 1 and 500. The standard setting in china is N = 12. A is the number of rising periods over the total number of periods.

Although, similar with the study of KDJ, we cannot get an adequate VAR model for the PSY index to HSI log return, we would like to show the simplified procedure below.

(2) Stationary Test

Looking through Figure 9 and Table 21, both of the plot of PSY index and the unit root test indicate that the series of PSY is stationary, so we can take it into the VAR model building.

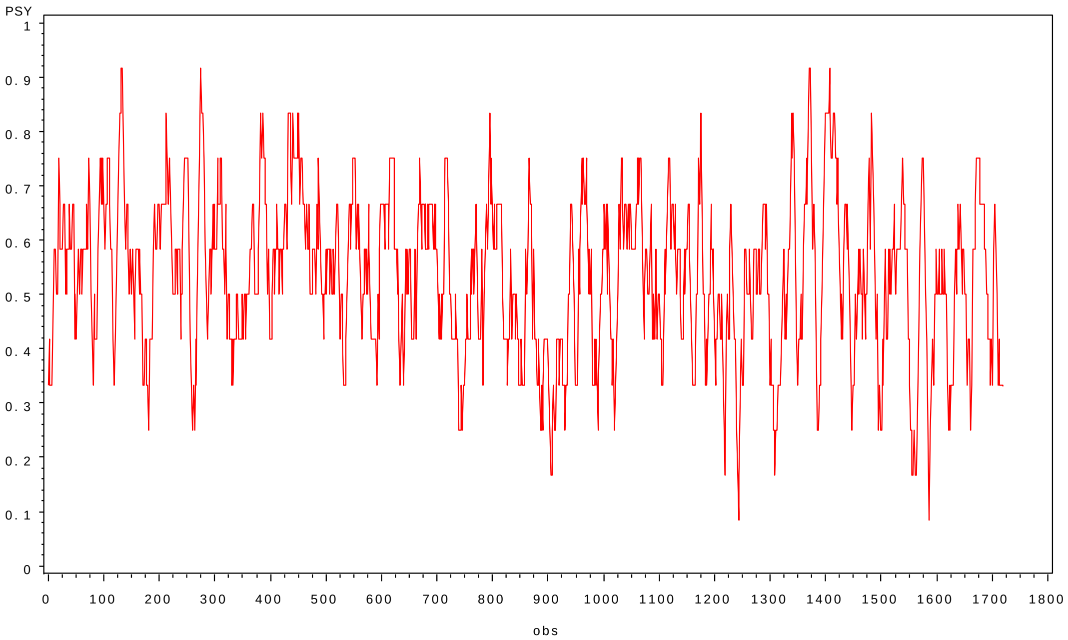

Figure 9 Plot of PSY index

Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
PSY	Zero Mean	-9.96	0.0284	-2.24	0.0242
	Single Mean	-154.68	0.0001	-8.77	<.0001
	Trend	-161.69	0.0001	-8.99	<.0001

Table 21 Unit Root test of PSY index

(3) Parameters Estimation

Realizing that significance of the estimators in Table 22 is alike with those of J line. Estimators in equation 1 are insignificant while in equation 2 are significant.

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
R	AR1_1_1	-0.03338	0.02412	-1.38	0.1667	R(t-1)
	AR1_1_2	0.00031	0.00079	0.39	0.6962	PSY(t-1)
PSY	AR1_2_1	1.53075	0.06870	22.28	0.0001	R(t-1)
	AR1_2_2	0.99397	0.00224	443.19	0.0001	PSY(t-1)

Table 22 Estimation of VAR(1) model of HSI log return and PSY index

AR-EGARCH Model

Since having introduced earlier, the log return of HSI series has different density over different periods, which is a clear Conditional heteroscedastic effect, we should further try a GARCH model to fit the log return of HSI. Then we see how the GARCH is.

Mean Equation Determination

Firstly, we detect whether the mean equation has some components besides zero mean.

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	8.38	6	0.2114	-0.033	0.012	-0.048	-0.036	-0.003	-0.008
12	25.80	12	0.0115	0.019	0.035	-0.052	-0.067	0.022	0.028

Table 23 Serial Correlation Check of HSI Log Return

In above table, the none zero autocorrelation can be found from the small enough p-value of Q(12) statistics. Then we get the best fitted one for the mean Equation--AR(3). The estimation results are shown in the Table 24.

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
AR1,1	-0.04762	0.02411	-1.98	0.0484	3

Table 24 Parameter Estimation of AR Model for the Mean Equation

ARCH Test

Next, we check for the ARCH effect in the residual series. The Q statistics and Lagrange Multiplier (LM) test are applied, and both of them significantly indicate that there is ARCH effect in the residual.

Q and LM Tests for ARCH Disturbances				
Order	Q	Pr > Q	LM	Pr > LM
1	277.3571	<.0001	276.9538	<.0001
2	535.8133	<.0001	381.8590	<.0001
3	744.9837	<.0001	417.2769	<.0001
4	837.5226	<.0001	417.3712	<.0001
5	907.5929	<.0001	417.4302	<.0001
6	955.5060	<.0001	417.6084	<.0001
7	1011.5380	<.0001	424.3382	<.0001
8	1121.0869	<.0001	457.5337	<.0001
9	1178.1687	<.0001	457.5963	<.0001
10	1310.6641	<.0001	478.3504	<.0001
11	1455.0924	<.0001	490.9698	<.0001
12	1515.1186	<.0001	494.9347	<.0001

Table 25 ARCH Test for Residual

Conditional Heteroscedastic Equation Determination

As we known, the return in stock market does not observe normal distribution, and the convention assumes innovations observe student's t distribution or generalized error distribution. In this case, we found that standardized residual pass the T distribution test, which will be presented later. On another hand, having tried GARCH, IGARCH, TGARCH, and EGARCH, we found that EGARCH(3,3) model could fit the series to a relatively good extent. The estimation result is shown in Table 26.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
AR3	1	-0.0408	0.009439	-4.32	<.0001
EARCH0	1	-0.5125	0.1244	-4.12	<.0001
EARCH1	1	0.1713	0.0249	6.88	<.0001
EARCH2	1	0.2629	0.0324	8.11	<.0001
EARCH3	1	0.1743	0.0217	8.05	<.0001
EGARCH1	1	-0.4068	0.005728	-71.02	<.0001
EGARCH2	1	0.3649	0.006055	60.27	<.0001
EGARCH3	1	0.9800	0.004481	218.70	<.0001
THETA	1	-0.3498	0.1065	-3.28	0.0010

Table 26 Parameter Estimation of AR(3)-EGARCH(3,3) Model of HSI Log Return

Finally, we get the complicated EGARCH model as following:

$$\mathrm{HSIlgr}_t = -0.0408\,\mathrm{HSIlgr}_{t-3} + a_t, \quad a_t = \sigma_t \epsilon_t,$$

$$\begin{aligned}\ln(\sigma_t^2) = {} & -0.5125 + 0.1713g(\epsilon_{t-1}) + 0.2629g(\epsilon_{t-2}) + 0.1743g(\epsilon_{t-3}) \\ & - 0.4068\ln(\sigma_{t-1}^2) + 0.3649\ln(\sigma_{t-2}^2) + 0.9800\ln(\sigma_{t-3}^2),\end{aligned}$$

$$g(\epsilon_t) = -0.3498\epsilon_t + [|\epsilon_t| - E|\epsilon_t|]$$

Model Checking

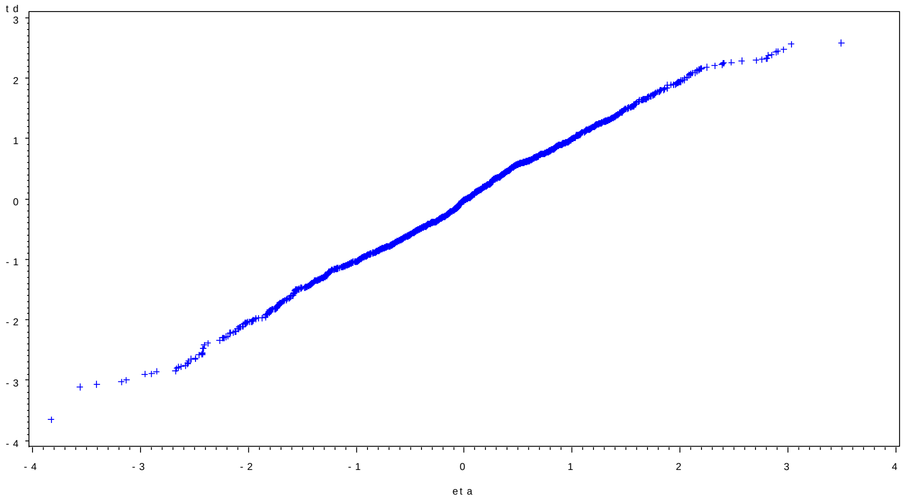

Figure 10 T Distribution QQ Plot of Standardized Residual

Figure 10 is t distribution QQ plot of Standardized Residual (ε_t), which is almost a straight line, verifying our assumption about the T distribution of standardized residual.

Looking through Table 27 and Table 28, the p-value of Q(12) is smaller than 5% significant level, thus, we can say that the AR(3)-EGARCH(3,3) model is adequate.

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	3.22	6	0.7808	0.011	0.002	0.004	-0.017	-0.003	-0.038
12	8.34	12	0.7582	0.027	-0.024	-0.005	-0.016	-0.018	0.032

Table 27 Check for Standardized Residual

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	15.72	6	0.0154	-0.049	0.015	0.061	0.010	-0.010	0.051
12	20.44	12	0.0593	-0.030	0.005	0.011	0.019	-0.032	0.017

Table 28 Check for Square Standardized Residual

Multifactor model

Factor selection

Previous study given us pretty good models, for a further step, we want to build a financial time series multi factor model employs more than only one index parameter, a multifactor model. However, when constructing a multi-factor model, it is difficult to decide how many and which factors to include. As we state before, various technical indicators obtain various properties and therefore reflect totally different but highly related aspects of the stock market. OBV is indicators to measure positive and negative volume flow. MACD returns rate of change of our HSI, just as a good reflection of HSI itself. BIAS measures the divergence of current HIS. KDJ calculates Raw Stochastic Value. PSY is a ratio reflecting the buying power in relation to the selling power. OBV, MACD and BIAS are three effective factors out of the five as we analyzed before, so it is seems to be optimal for us to consider three factors BIAS, MACD and OBV and thus we shall try 4 possible models showing below:

log return of the Heng Seng Index with BIAS, OBV and lnMACD;

log return of the Heng Seng Index with BIAS and lnMACD;

log return of the Heng Seng Index withOBV and lnMACD;

log return of the Heng Seng Index with BIAS and OBV

Intuitively, I feel the model of three indicators and the model with OBV and lnMACD may be the best two among these four possible models. That is because OBV and lnMACD represent two different aspects of the stock markets, say the volume flow and return of HSI.

Model Specification

In the previous section, we decide to use three factors to calibrate the log return of HIS, and just one step left for model fitting is which model shall we use. It is able to capture the two-way effect of HSI and various indicators level on each other, if such effect actually exists. In this section, we will look into the statistical performance of the cross correlations of these three indicators

.

```
                          The VARMAX Procedure

            Schematic Representation of Cross Correlations
Variable/
Lag        0    1    2    3    4    5    6    7    8    9    10   11   12

r          ++++ .+.- .+.- .+.- .+.- ...- .... .... .... -..- --.- ...+ ....
BIAS       ++++ .+.- .+.- -+.- ...- ...- .... ...- ...- --.- .-.. ...+ ....
d_OBV      ++++ .+.- .+.- .+.- .+.- ...- .... .... .... ...- .... .... ....
lnMACD     ++++ -+.- ++.. -+-- .+.. .-.- .... .... +..+ -..- --.- ...+ ....

        + is > 2*std error,  - is < -2*std error,  . is between
```

Table 29: The VARMAX Procedure

According to this SAS output of Schematic Representation of Cross Correlations, noticing the correlations are only significant up to fourth lag, suggesting a series of models specification. Here, we list four models with lag one.

$$\log r_t = \alpha \log r_{t-1} + \beta BIAS(t-1) + \gamma OBV(t-1) + \delta \ln MACD\,(t-1)$$

$$\log r_t = \alpha \log r_{t-1} + \beta BIAS(t-1) + \gamma OBV(t-1)$$

$$\log r_t = \alpha \log r_{t-1} + \beta BIAS(t-1) + \gamma \ln MACD\,(t-1)$$

$\log r_t = \alpha \log r_{t-1} + \beta \ln MACD\ (t-1) + \gamma OBV(t-1)$

Models Fitting

Thanks to the previous analysis, there is no need for data checking, and we directly go to the Parameter Estimates in SAS Output.

Model with three factors

For the model with three different factors, $\log r_t = \alpha \log r_{t-1} + \beta BIAS + \gamma OBV + \delta \ln MACD$ we tried VAR models up to fourth lag, we find the VAR(1) is the best among all, and the result of VAR(1) is below:

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	AR1_1_1	0.12272	0.07389	1.66	0.0969	r(t-1)
	AR1_1_2	-7.65374	3.16370	-2.42	0.0157	BIAS(t-1)
	AR1_1_3	0.00192	0.00339	0.57	0.5716	d_OBV(t-1)
	AR1_1_4	-9.80087	3.82987	-2.56	0.0106	lnMACD(t-1)
BIAS	AR1_2_1	0.00158	0.00062	2.55	0.0108	r(t-1)
	AR1_2_2	0.62517	0.02654	23.55	0.0001	BIAS(t-1)
	AR1_2_3	0.00002	0.00003	0.62	0.5344	d_OBV(t-1)
	AR1_2_4	-0.04368	0.03213	-1.36	0.1742	lnMACD(t-1)
d_OBV	AR1_3_1	0.81405	0.81481	1.00	0.3179	r(t-1)
	AR1_3_2	-29.56235	34.88819	-0.85	0.3969	BIAS(t-1)
	AR1_3_3	0.01504	0.03733	0.40	0.6870	d_OBV(t-1)
	AR1_3_4	-55.48351	42.23441	-1.31	0.1891	lnMACD(t-1)
lnMACD	AR1_4_1	0.00313	0.00095	3.30	0.0010	r(t-1)
	AR1_4_2	-0.48224	0.04053	-11.90	0.0001	BIAS(t-1)
	AR1_4_3	0.00006	0.00004	1.28	0.2008	d_OBV(t-1)
	AR1_4_4	-0.22948	0.04907	-4.68	0.0001	lnMACD(t-1)

Table 31： The VARMAX Procedure

Based on the significant level of 0.1, the P-value of index OBV is 0.5456, quite large, thus we can not reject the null hypothesis, and conclude the coefficient of index OBV is not significant, which means the OBV parameter vanishes in the model. If we neglect the insignificance, we get the model with three factors as below:

$$\log r_t = 0.12272 * \log r_{t-1} - 7.65374 * BIAS(t-1) + 0.00192 * \mathrm{d}\_OBV(t-1) - 9.80087 * \ln MACD(t-1)$$

Model with two factors

Due to the cross correlation, we may have to testify each model up to fourth lag. Actually, when P becomes larger in the VAR model, the Parameter Estimates perform insignificance. In other words, the VAR(1) model is adequate. Thus, we tried models in the following form:

$$\log r_t = \alpha \log r_{t-1} + \beta BIAS(t-1) + \gamma OBV(t-1)$$

$$\log r_t = \alpha \log r_{t-1} + \beta OBV(t-1) + \gamma LnMACD(t-1)$$

$$\log r_t = \alpha \log r_{t-1} + \beta IAS(t-1) + \gamma LnMACD(t-1)$$

According to the statistical performance of each model, we find the VAR(1) model with respect to BIAS and MACD is the best among all. In fact, we tried VAR(2) model, $\log r_t = \alpha \log r_{t-2} + \beta BIAS(t-2) + \gamma LnMACD(t-2)$, for BIAS and MACD, while it turns to be make nonsense, which corresponds to our statement at the beginning of this section. The SAS output of VAR(1) shows below.

Dickey-Fuller Unit Root Tests

Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
r	**Zero Mean**	-1734.9	0.0001	-29.44	<.0001
	Single Mean	-1735.4	0.0001	-29.43	<.0001
	Trend	-1737.8	0.0001	-29.44	<.0001
BIAS	**Zero Mean**	-655.31	0.0001	-18.09	<.0001
	Single Mean	-655.85	0.0001	-18.09	<.0001
	Trend	-657.55	0.0001	-18.11	<.0001
lnMACD	**Zero Mean**	-1936.7	0.0001	-31.10	<.0001
	Single Mean	-1936.7	0.0001	-31.09	<.0001
	Trend	-1936.7	0.0001	-31.08	<.0001

Model Parameter Estimates						
Equation	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variable**
r	**AR1_1_1**	0.14324	0.06436	2.23	0.0262	r(t-1)
	AR1_1_2	-7.71461	3.16124	-2.44	0.0148	BIAS(t-1)
	AR1_1_3	-10.16665	3.77415	-2.69	0.0071	lnMACD(t-1)
BIAS	**AR1_2_1**	0.00177	0.00054	3.28	0.0011	r(t-1)
	AR1_2_2	0.62461	0.02652	23.55	0.0001	BIAS(t-1)
	AR1_2_3	-0.04705	0.03166	-1.49	0.1375	lnMACD(t-1)
lnMACD	**AR1_3_1**	0.00372	0.00082	4.51	0.0001	r(t-1)
	AR1_3_2	-0.48400	0.04051	-11.95	0.0001	BIAS(t-1)
	AR1_3_3	-0.24008	0.04837	-4.96	0.0001	lnMACD(t-1)

Cointegration Rank Test Using Trace						
H0: Rank=r	**H1: Rank>r**	**Eigenvalue**	**Trace**	**5% Critical Value**	**Drift in ECM**	**Drift in Process**
0	**0**	0.9678	7588.1384	24.08	NOINT	Constant
1	**1**	0.5591	1690.1700	12.21		
2	**2**	0.1524	283.9763	4.14		

Long-Run Parameter Beta Estimates			
Variable	**1**	**2**	**3**
r	1.00000	1.00000	1.00000
BIAS	-34.32319	-31.55515	-338.67190
lnMACD	8.16656	-64.86788	32.97914

Adjustment Coefficient Alpha Estimates			
Variable	**1**	**2**	**3**
r	-1.06460	0.08510	0.12274
BIAS	-0.00047	0.00120	0.00104
lnMACD	-0.01546	0.01785	0.00133

Table 302： The VARMAX Procedure

The output of Cointegration Rank Test Using Trace indicates that the series are not cointegrated with rank 0,1 and2, because the trace statistics are greater than the critical values.

Till now, we tried all the possible combination of two factor models, and we can see that only the model with respect to the indexes of MACD and BIAS can we get the P-values of parameter estimates to be small (0.0262, 0.0148, 0.0071 correspond to r_t-1, BIAS, LnMACD). Therefore, we can conclude at the significant level of 0.05, the coefficients of the VAR model with respect to BIAS and MACD are significant. The corresponding model is:

$$\log r_t = 0.14324 * \log r_{t-1} - 7.71461 * BIAS(t-1) - 10.16665 * LnMACD(t-1)$$

Two factors model & three factors model

Our model construction is based on the statistical performance of return to the combination of different indexes. We tried a lot models to find an optimal model employs multi index factor. Up to now, we get two models with different numbers of factors:

$$\log r_t = 0.12272 * \log r_{t-1} - 7.65374 * BIAS(t-1) + 0.00192 * \mathrm{d\_}OBV(t-1) - 9.80087 * \ln MACD(t-1)$$

$$\log r_t = 0.14324 * \log r_{t-1} - 7.71461 * BIAS(t-1) - 10.16665 * LnMACD(t-1)$$

According to the AIC and BIC of each model, we choose the.

	Three factor model	Two factor model
AIC	-14.6506	-19.7589
SBC	-14.5998	-19.7304
AIC+ SBC	-29.2504	-39.4893

Table33： AIC and BIC of multi factor models

According to the criteria of AIC and BIC, we pick the least AIC and BIC, and therefore conclude the two factor model obtains better statistic performance. Although we expect the model with respect to OBV and MACD to be better fitted one, the OBV parameter seems to be very insignificant. I think it is because the OBV index is based on volume flow, which is quite large. However, numerically the log return is a tiny number. In order to fit a model of the log return employing OBV, the corresponding coefficient becomes quite small.

To sum up, in this section, we picked effective factors to construct multifactor models, and after all, we find an optimal multifactor model:

$$\log r_t = 0.14324 * \log r_{t-1} - 7.71461 * BIAS(t-1) - 10.16665 * LnMACD(t-1)$$

Conclusion

Best Model

Till now, we construct both single factor models with multi factors models. Comparing their AIC and BIC below, we can see which one is the model we are looking for.

	Two Factor	MACD	BIAS	OBV	EGARCH
AIC	-19.7589	-8.04597	-21.1231	6.317855	-9854.3675
SBC	-19.7304	-8.04021	-21.0976	6.349421	-9805.3272

Table33： AIC and BIC of all the models

So apparently, the EGARCH model is the best one among all the models we tried before.

Weakness

For the data, we only focus on model construction and analysis, but without data pre-processing. Because we downloaded our data from the internet via Bloomberg directly, it may result in out-of-range values, impossible data combinations, missing values, etc. Although we double checked the data to be reasonable, we did not process normalization and data cleaning to exclude the outliers, which resulted from some specific scenarios. Thus, for later study, one may look forward to extracting unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining).

Model is judged on historical numbers, which might not accurately predict future values.

In our model construction and factor selection, we only considered five technical indicators. While, indicated by the result of GARCH model, one may interested in models involving volatilities and other parameters for later development.

Conclusion

In this paper, we tried to examine the effect of five popular technical indicators based on the historical data of HIS. We constructed optimal time series models for every single technical indicator and HIS, and also EGARCH model. Picked three effective indicators from five technical indicators, we considered to find out a multifactor model then. Among all the models we examined, the model of BIAS and two-factor-model are of well performance. This coincides with our motivation that the technical analysis in the market place actually creates a reversal impact on the stock price itself, and also the behavior finance theory.

Appendix

Appendix A

The SAS System

The ARIMA Procedure

Name of Variable = lnPRICE	
Mean of Working Series	9.857491
Standard Deviation	0.206085
Number of Observations	1718

Autocorrelations				
Lag	**Covariance**	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**	**Std Error**
0	0.042471	1.00000	| |********************|	0

<table>
<tr><th colspan="5">Autocorrelations</th></tr>
<tr><th>Lag</th><th>Covariance</th><th>Correlation</th><th>-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1</th><th>Std Error</th></tr>
<tr><td>1</td><td>0.042274</td><td>0.99537</td><td>| .|********************|</td><td>0.024126</td></tr>
<tr><td>2</td><td>0.042089</td><td>0.99101</td><td>| . |********************|</td><td>0.041659</td></tr>
<tr><td>3</td><td>0.041899</td><td>0.98652</td><td>| . |********************|</td><td>0.053654</td></tr>
<tr><td>4</td><td>0.041723</td><td>0.98238</td><td>| . |********************|</td><td>0.063338</td></tr>
<tr><td>5</td><td>0.041557</td><td>0.97847</td><td>| . |********************|</td><td>0.071661</td></tr>
<tr><td>6</td><td>0.041395</td><td>0.97466</td><td>| . |******************* |</td><td>0.079056</td></tr>
<tr><td>7</td><td>0.041235</td><td>0.97091</td><td>| . |******************* |</td><td>0.085765</td></tr>
<tr><td>8</td><td>0.041070</td><td>0.96701</td><td>| . |******************* |</td><td>0.091941</td></tr>
<tr><td>9</td><td>0.040894</td><td>0.96286</td><td>| . |******************* |</td><td>0.097682</td></tr>
<tr><td>10</td><td>0.040735</td><td>0.95912</td><td>| . |******************* |</td><td>0.103058</td></tr>
<tr><td>11</td><td>0.040596</td><td>0.95585</td><td>| . |******************* |</td><td>0.108129</td></tr>
<tr><td>12</td><td>0.040450</td><td>0.95241</td><td>| . |******************* |</td><td>0.112940</td></tr>
</table>

"." marks two standard errors

<table>
<tr><th colspan="3">Inverse Autocorrelations</th></tr>
<tr><th>Lag</th><th>Correlation</th><th>-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1</th></tr>
<tr><td>1</td><td>-0.47483</td><td>| *********|. |</td></tr>
<tr><td>2</td><td>-0.05358</td><td>| *|. |</td></tr>
<tr><td>3</td><td>0.03340</td><td>| .|* |</td></tr>
<tr><td>4</td><td>0.00278</td><td>| .|. |</td></tr>
<tr><td>5</td><td>-0.00830</td><td>| .|. |</td></tr>
<tr><td>6</td><td>0.01447</td><td>| .|. |</td></tr>
</table>

Inverse Autocorrelations		
Lag	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**
7	-0.00655	| .|. |
8	-0.04661	| *|. |
9	0.03709	| .|* |
10	0.03935	| .|* |
11	-0.04838	| *|. |
12	0.01116	| .|. |

Partial Autocorrelations		
Lag	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**
1	0.99537	| .|********************|
2	0.02792	| .|* |
3	-0.01585	| .|. |
4	0.03444	| .|* |
5	0.02552	| .|* |
6	0.00879	| .|. |
7	0.00619	| .|. |
8	-0.01521	| .|. |
9	-0.02977	| *|. |
10	0.04281	| .|* |
11	0.05235	| .|* |
12	-0.02180	| .|. |

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	9999.99	6	<.0001	0.995	0.991	0.987	0.982	0.978	0.975
12	9999.99	12	<.0001	0.971	0.967	0.963	0.959	0.956	0.952

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	**0**	0.0273	0.6894	0.36	0.7887		
	1	0.0270	0.6893	0.37	0.7907		
	2	0.0279	0.6895	0.38	0.7927		
	3	0.0286	0.6897	0.40	0.7999		
Single Mean	**0**	-7.9263	0.2214	-2.19	0.2114	2.47	0.4366
	1	-7.4898	0.2455	-2.13	0.2330	2.35	0.4676
	2	-7.7317	0.2319	-2.17	0.2183	2.44	0.4454
	3	-7.2252	0.2613	-2.12	0.2376	2.34	0.4701
Trend	**0**	-8.4065	0.5527	-2.02	0.5909	2.42	0.6924
	1	-7.8546	0.5958	-1.94	0.6320	2.28	0.7195
	2	-8.1162	0.5752	-1.98	0.6123	2.37	0.7025

The SAS System

The ARIMA Procedure

Name of Variable = dlnPRICE	
Mean of Working Series	0.017611
Standard Deviation	1.804943
Number of Observations	1718

Autocorrelations				
Lag	**Covariance**	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**	**Std Error**
0	3.257820	1.00000	| |********************|	0
1	-0.108728	-.03337	| *|. |	0.024126
2	0.038679	0.01187	| .|. |	0.024153
3	-0.155472	-.04772	| *|. |	0.024156
4	-0.115761	-.03553	| *|. |	0.024211
5	-0.011145	-.00342	| .|. |	0.024242
6	-0.025084	-.00770	| .|. |	0.024242
7	0.060837	0.01867	| .|. |	0.024243
8	0.115599	0.03548	| .|* |	0.024252
9	-0.169436	-.05201	| *|. |	0.024282
10	-0.217853	-.06687	| *|. |	0.024347
11	0.073055	0.02242	| .|. |	0.024453
12	0.090969	0.02792	| .|* |	0.024465

"." marks two standard errors

Inverse Autocorrelations		
Lag	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**
1	0.03713	| .|* |
2	-0.02088	| .|. |
3	0.04604	| .|* |
4	0.03926	| .|* |

Inverse Autocorrelations		
Lag	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**
5	0.00752	| .|. |
6	0.01831	| .|. |
7	-0.01378	| .|. |
8	-0.03728	| *|. |
9	0.05329	| .|* |
10	0.06876	| .|* |
11	-0.02527	| *|. |
12	-0.02879	| *|. |

Partial Autocorrelations		
Lag	**Correlation**	**-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1**
1	-0.03337	| *|. |
2	0.01077	| .|. |
3	-0.04703	| *|. |
4	-0.03890	| *|. |
5	-0.00493	| .|. |
6	-0.00954	| .|. |
7	0.01467	| .|. |
8	0.03520	| .|* |
9	-0.05136	| *|. |
10	-0.07075	| *|. |
11	0.02360	| .|. |
12	0.02925	| .|* |

Autocorrelation Check for White Noise									
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Autocorrelations**					
6	8.38	6	0.2114	-0.033	0.012	-0.048	-0.036	-0.003	-0.008
12	25.80	12	0.0115	0.019	0.035	-0.052	-0.067	0.022	0.028

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	**0**	-1774.14	0.0001	-42.83	<.0001		
	1	-1734.94	0.0001	-29.44	<.0001		
	2	-2004.37	0.0001	-25.13	<.0001		
	3	-2390.45	0.0001	-22.32	<.0001		
Single Mean	**0**	-1774.30	0.0001	-42.82	<.0001	916.71	0.0010
	1	-1735.45	0.0001	-29.43	<.0001	433.11	0.0010
	2	-2005.72	0.0001	-25.13	<.0001	315.67	0.0010
	3	-2393.59	0.0001	-22.32	<.0001	249.00	0.0010
Trend	**0**	-1775.08	0.0001	-42.83	<.0001	917.01	0.0010
	1	-1737.78	0.0001	-29.44	<.0001	433.43	0.0010
	2	-2011.63	0.0001	-25.14	<.0001	316.09	0.0010
	3	-2407.77	0.0001	-22.34	<.0001	249.49	0.0010

Appendix B

The SAS System

The VARMAX Procedure

Number of Observations	1729
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
return	**Dependent**	1729	0.01497	1.80112	-13.58202	13.40681
index	**Dependent**	1729	0.64973	19.78895	-204.33285	153.76688

Dickey-Fuller Unit Root Tests

Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
return	**Zero Mean**	-1743.0	0.0001	-29.51	<.0001
	Single Mean	-1743.4	0.0001	-29.50	<.0001
	Trend	-1745.2	0.0001	-29.51	<.0001
index	**Zero Mean**	-1671.2	0.0001	-28.89	<.0001
	Single Mean	-1676.5	0.0001	-28.93	<.0001
	Trend	-1677.5	0.0001	-28.93	<.0001

The SAS System

The VARMAX Procedure

Type of Model	VARMA(1,1)
Estimation Method	Maximum Likelihood Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
return	**CONST1**	1.09004	0.64374	1.69	0.0906	1
	AR1_1_1	3.69917	0.47270	7.83	0.0001	return(t-1)
	AR1_1_2	-1.83658	0.13425	-13.68	0.0001	index(t-1)
	MA1_1_1	3.78239	0.47464	7.97	0.0001	e1(t-1)
	MA1_1_2	-1.84108	0.13419	-13.72	0.0001	e2(t-1)
index	**CONST2**	2.66670	1.57101	1.70	0.0898	1
	AR1_2_1	6.83499	1.76920	3.86	0.0001	return(t-1)
	AR1_2_2	-3.49872	0.41650	-8.40	0.0001	index(t-1)
	MA1_2_1	7.00441	1.79618	3.90	0.0001	e1(t-1)
	MA1_2_2	-3.50858	0.41773	-8.40	0.0001	e2(t-1)

Covariances of Innovations

Variable	return	index
return	3.23642	26.88117
index	26.88117	392.59404

Information Criteria

Information Criteria	
AICC	6.317888
HQC	6.329531
AIC	6.317855
SBC	6.349421
FPEC	554.3823

Appendix D

VAR1 MODEL with all the three index

The SAS System

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
d_OBV	**Dependent**	1718	0.66501	19.84924	-204.33285	153.76688
lnMACD	**Dependent**	1718	0.00000	0.02456	-0.26575	0.41128

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	**CONST1**	0.01839	0.04354	0.42	0.6728	1
	AR1_1_1	0.12337	0.07392	1.67	0.0953	r(t-1)
	AR1_1_2	-7.68188	3.16516	-2.43	0.0153	BIAS(t-1)
	AR1_1_3	0.00186	0.00339	0.55	0.5830	d_OBV(t-1)
	AR1_1_4	-9.80716	3.83081	-2.56	0.0106	lnMACD(t-1)
BIAS	**CONST2**	0.00016	0.00037	0.43	0.6671	1
	AR1_2_1	0.00159	0.00062	2.56	0.0106	r(t-1)
	AR1_2_2	0.62493	0.02655	23.53	0.0001	BIAS(t-1)
	AR1_2_3	0.00002	0.00003	0.60	0.5456	d_OBV(t-1)
	AR1_2_4	-0.04373	0.03214	-1.36	0.1738	lnMACD(t-1)
d_OBV	**CONST3**	0.66006	0.47993	1.38	0.1692	1
	AR1_3_1	0.83727	0.81477	1.03	0.3043	r(t-1)
	AR1_3_2	-30.57203	34.88684	-0.88	0.3810	BIAS(t-1)

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
	AR1_3_3	0.01308	0.03735	0.35	0.7262	d_OBV(t-1)
	AR1_3_4	-55.70910	42.22374	-1.32	0.1872	lnMACD(t-1)
lnMACD	**CONST4**	0.00016	0.00056	0.29	0.7690	1
	AR1_4_1	0.00313	0.00095	3.31	0.0010	r(t-1)
	AR1_4_2	-0.48249	0.04055	-11.90	0.0001	BIAS(t-1)
	AR1_4_3	0.00006	0.00004	1.27	0.2052	d_OBV(t-1)
	AR1_4_4	-0.22953	0.04908	-4.68	0.0001	lnMACD(t-1)

Covariances of Innovations

Variable	r	BIAS	d_OBV	lnMACD
r	3.24833	0.02687	27.00423	0.03627
BIAS	0.02687	0.00023	0.22489	0.00030
d_OBV	27.00423	0.22489	394.63094	0.27585
lnMACD	0.03627	0.00030	0.27585	0.00053

Information Criteria

AICC	-14.6475
HQC	-14.6241
AIC	-14.6476
SBC	-14.5841

Information Criteria	
FPEC	4.351E-7

So we drop intercept term to run VAR

The SAS System

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
d_OBV	**Dependent**	1718	0.66501	19.84924	-204.33285	153.76688
lnMACD	**Dependent**	1718	0.00000	0.02456	-0.26575	0.41128

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	**AR1_1_1**	0.12272	0.07389	1.66	0.0969	r(t-1)
	AR1_1_2	-7.65374	3.16370	-2.42	0.0157	BIAS(t-1)
	AR1_1_3	0.00192	0.00339	0.57	0.5716	d_OBV(t-1)
	AR1_1_4	-9.80087	3.82987	-2.56	0.0106	lnMACD(t-1)
BIAS	**AR1_2_1**	0.00158	0.00062	2.55	0.0108	r(t-1)
	AR1_2_2	0.62517	0.02654	23.55	0.0001	BIAS(t-1)
	AR1_2_3	0.00002	0.00003	0.62	0.5344	d_OBV(t-1)
	AR1_2_4	-0.04368	0.03213	-1.36	0.1742	lnMACD(t-1)
d_OBV	**AR1_3_1**	0.81405	0.81481	1.00	0.3179	r(t-1)
	AR1_3_2	-29.56235	34.88819	-0.85	0.3969	BIAS(t-1)
	AR1_3_3	0.01504	0.03733	0.40	0.6870	d_OBV(t-1)
	AR1_3_4	-55.48351	42.23441	-1.31	0.1891	lnMACD(t-1)
lnMACD	**AR1_4_1**	0.00313	0.00095	3.30	0.0010	r(t-1)

Model Parameter Estimates						
Equation	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variable**
	AR1_4_2	-0.48224	0.04053	-11.90	0.0001	BIAS(t-1)
	AR1_4_3	0.00006	0.00004	1.28	0.2008	d_OBV(t-1)
	AR1_4_4	-0.22948	0.04907	-4.68	0.0001	lnMACD(t-1)

Covariances of Innovations				
Variable	**r**	**BIAS**	**d_OBV**	**lnMACD**
r	3.24677	0.02686	27.00061	0.03625
BIAS	0.02686	0.00023	0.22486	0.00030
d_OBV	27.00061	0.22486	394.83631	0.27579
lnMACD	0.03625	0.00030	0.27579	0.00053

Information Criteria	
AICC	-14.6506
HQC	-14.6318
AIC	-14.6506
SBC	-14.5998
FPEC	4.338E-7

Next step, we try two of the index with the return

So we drop MACD, using OBV and BIAS to run VAR(1)

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
d_OBV	**Dependent**	1718	0.66501	19.84924	-204.33285	153.76688

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	**AR1_1_1**	-0.03355	0.04166	-0.81	0.4207	r(t-1)
	AR1_1_2	-3.52284	2.72530	-1.29	0.1963	BIAS(t-1)
	AR1_1_3	0.00338	0.00334	1.01	0.3123	d_OBV(t-1)
BIAS	**AR1_2_1**	0.00088	0.00035	2.53	0.0113	r(t-1)
	AR1_2_2	0.64358	0.02283	28.19	0.0001	BIAS(t-1)
	AR1_2_3	0.00002	0.00003	0.86	0.3882	d_OBV(t-1)
d_OBV	**AR1_3_1**	-0.07065	0.45878	-0.15	0.8776	r(t-1)
	AR1_3_2	-6.17697	30.01146	-0.21	0.8370	BIAS(t-1)
	AR1_3_3	0.02332	0.03680	0.63	0.5264	d_OBV(t-1)

Covariances of Innovations

Variable	r	BIAS	d_OBV
r	3.25728	0.02690	27.05508
BIAS	0.02690	0.00023	0.22504

Covariances of Innovations			
Variable	**r**	**BIAS**	**d_OBV**
d_OBV	27.05508	0.22504	395.00351

Information Criteria	
AICC	-5.62945
HQC	-5.61891
AIC	-5.62947
SBC	-5.60091
FPEC	0.00359

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics						
Variable	**Type**	**N**	**Mean**	**Standard Deviation**	**Min**	**Max**
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
lnMACD	**Dependent**	1718	0.00000	0.02456	-0.26575	0.41128

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	**AR1_1_1**	0.14324	0.06436	2.23	0.0262	r(t-1)
	AR1_1_2	-7.71461	3.16124	-2.44	0.0148	BIAS(t-1)
	AR1_1_3	-10.16665	3.77415	-2.69	0.0071	lnMACD(t-1)
BIAS	**AR1_2_1**	0.00177	0.00054	3.28	0.0011	r(t-1)
	AR1_2_2	0.62461	0.02652	23.55	0.0001	BIAS(t-1)
	AR1_2_3	-0.04705	0.03166	-1.49	0.1375	lnMACD(t-1)
lnMACD	**AR1_3_1**	0.00372	0.00082	4.51	0.0001	r(t-1)
	AR1_3_2	-0.48400	0.04051	-11.95	0.0001	BIAS(t-1)
	AR1_3_3	-0.24008	0.04837	-4.96	0.0001	lnMACD(t-1)

Covariances of Innovations

Variable	r	BIAS	lnMACD
r	3.24548	0.02685	0.03625
BIAS	0.02685	0.00023	0.00030

Covariances of Innovations			
Variable	**r**	**BIAS**	**lnMACD**
lnMACD	0.03625	0.00030	0.00053

Information Criteria	
AICC	-19.7589
HQC	-19.7484
AIC	-19.7589
SBC	-19.7304
FPEC	2.623E-9

OBV and MACD

The SAS System

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Model Parameter Estimates						
Equation	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variable**
r	**AR1_1_1**	0.00634	0.05616	0.11	0.9101	r(t-1)
	AR1_1_2	0.00219	0.00339	0.65	0.5174	d_OBV(t-1)
	AR1_1_3	-5.07338	3.29848	-1.54	0.1242	lnMACD(t-1)

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
d_OBV	**AR1_2_1**	0.36454	0.61840	0.59	0.5556	r(t-1)
	AR1_2_2	0.01612	0.03731	0.43	0.6657	d_OBV(t-1)
	AR1_2_3	-37.22373	36.32014	-1.02	0.3056	lnMACD(t-1)
lnMACD	**AR1_3_1**	-0.00421	0.00075	-5.63	0.0001	r(t-1)
	AR1_3_2	0.00007	0.00005	1.62	0.1054	d_OBV(t-1)
	AR1_3_3	0.06839	0.04389	1.56	0.1194	lnMACD(t-1)

Covariances of Innovations

Variable	r	d_OBV	lnMACD
r	3.25596	27.02767	0.03693
d_OBV	27.02767	394.77135	0.27833
lnMACD	0.03693	0.27833	0.00058

Information Criteria

AICC	-2.46049
HQC	-2.44994
AIC	-2.46051
SBC	-2.43195
FPEC	0.085391

One step further, I try a VAR(2) model with index OBV and MACD

The SAS System

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
lnMACD	**Dependent**	1718	0.00000	0.02456	-0.26575	0.41128

The VARMAX Procedure

Type of Model	VAR(2)
Estimation Method	Least Squares Estimation

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
r	**AR2_1_1**	-0.07380	0.06436	-1.15	0.2517	r(t-2)
	AR2_1_2	-2.43753	3.16126	-0.77	0.4408	BIAS(t-2)
	AR2_1_3	9.21170	3.77408	2.44	0.0148	lnMACD(t-2)
BIAS	**AR2_2_1**	0.00153	0.00068	2.26	0.0240	r(t-2)
	AR2_2_2	0.32111	0.03337	9.62	0.0001	BIAS(t-2)
	AR2_2_3	0.06403	0.03983	1.61	0.1081	lnMACD(t-2)
lnMACD	**AR2_3_1**	-0.00106	0.00084	-1.26	0.2095	r(t-2)
	AR2_3_2	-0.30328	0.04138	-7.33	0.0001	BIAS(t-2)
	AR2_3_3	0.19373	0.04940	3.92	0.0001	lnMACD(t-2)

Covariances of Innovations

Variable	r	BIAS	lnMACD
r	3.24521	0.02592	0.03640
BIAS	0.02592	0.00036	0.00023

Covariances of Innovations			
Variable	**r**	**BIAS**	**lnMACD**
lnMACD	0.03640	0.00023	0.00056

Information Criteria	
AICC	-16.5934
HQC	-16.5829
AIC	-16.5935
SBC	-16.5649
FPEC	6.217E-8

To check this model, we add some options in the SAS statement to get the output below:

```
proc varmax data=test;
model   r      BIAS lnMACD / p=1 noint
print=(corry) minic=(p=5 q=5)
dftest cointtest=(johansen=(normalize=r))
ecm=(rank=1 normalize=r);
output lead=8;
run;
```

The SAS System

The VARMAX Procedure

Number of Observations	1718
Number of Pairwise Missing	0

Simple Summary Statistics

Variable	Type	N	Mean	Standard Deviation	Min	Max
r	**Dependent**	1718	0.01761	1.80547	-13.58202	13.40681
BIAS	**Dependent**	1718	0.00053	0.02140	-0.19421	0.11628
lnMACD	**Dependent**	1718	0.00000	0.02456	-0.26575	0.41128

Dickey-Fuller Unit Root Tests

Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
r	**Zero Mean**	-1734.9	0.0001	-29.44	<.0001
	Single Mean	-1735.4	0.0001	-29.43	<.0001
	Trend	-1737.8	0.0001	-29.44	<.0001
BIAS	**Zero Mean**	-655.31	0.0001	-18.09	<.0001
	Single Mean	-655.85	0.0001	-18.09	<.0001
	Trend	-657.55	0.0001	-18.11	<.0001
lnMACD	**Zero Mean**	-1936.7	0.0001	-31.10	<.0001
	Single Mean	-1936.7	0.0001	-31.09	<.0001
	Trend	-1936.7	0.0001	-31.08	<.0001

Cross Correlations of Dependent Series

Lag	Variable	r	BIAS	lnMACD
0	r	1.00000	0.66360	0.83772
	BIAS	0.66360	1.00000	0.34675
	lnMACD	0.83772	0.34675	1.00000
1	r	-0.03337	0.51856	-0.20759
	BIAS	-0.04459	0.70481	-0.32370
	lnMACD	-0.04999	0.28772	-0.15722
2	r	0.01187	0.40394	-0.09092
	BIAS	-0.03469	0.43212	-0.24888
	lnMACD	0.05344	0.29327	0.03694
3	r	-0.04772	0.23485	-0.22859
	BIAS	-0.05057	0.18385	-0.23528
	lnMACD	-0.09280	0.14928	-0.27657
4	r	-0.03553	0.08341	-0.09485
	BIAS	-0.01968	0.02190	-0.10550
	lnMACD	0.00217	0.07528	-0.02758
5	r	-0.00342	-0.04391	-0.06543
	BIAS	0.00378	-0.04726	-0.06504
	lnMACD	-0.03065	-0.05033	-0.06890
6	r	-0.00770	-0.03581	-0.01003
	BIAS	-0.00319	-0.02893	-0.04538
	lnMACD	0.01186	-0.02804	0.03469

Cross Correlations of Dependent Series

Lag	Variable	r	BIAS	lnMACD
7	r	0.01867	-0.01117	0.00133
	BIAS	-0.00289	-0.01510	-0.05257
	lnMACD	0.01575	-0.00949	0.02002
8	r	0.03548	0.02644	0.01644
	BIAS	-0.02422	-0.02040	-0.06627
	lnMACD	0.05949	0.04727	0.06010
9	r	-0.05201	-0.00970	-0.12538
	BIAS	-0.05088	-0.04987	-0.07845
	lnMACD	-0.05350	0.00371	-0.13998
10	r	-0.06687	-0.05518	-0.07100
	BIAS	-0.01211	-0.04934	0.01571
	lnMACD	-0.11026	-0.07312	-0.11151
11	r	0.02242	-0.03323	0.06925
	BIAS	0.04563	-0.00626	0.08036
	lnMACD	0.03100	-0.04544	0.08568
12	r	0.02792	-0.00551	0.03670
	BIAS	0.04479	0.03229	0.03720
	lnMACD	0.02659	-0.01628	0.04772

Schematic Representation of Cross Correlations

Variable/Lag	0	1	2	3	4	5	6	7	8	9	10	11	12
r	+++	.+-	.+-	.+-	.+-	..-	...	...	...	-.-	---	..+	...
BIAS	+++	.+-	.+-	-+-	..-	..-	...	..-	..-	---	.-.	..+	...
lnMACD	+++	-+-	++.	-+-	.+.	.--	...	...	+.+	-.-	---	..+	...

+ is > 2*std error, - is < -2*std error, . is between

Minimum Information Criterion Based on AICC

Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-16.0162	-16.78822	-17.64968	-18.77797	-19.47009	-19.92062
AR 1	-19.7589	-19.86948	-20.21459	-20.74225	-20.75814	-22.51578
AR 2	-20.62184	-20.76311	-20.84758	-21.60711	-21.96659	-23.08732
AR 3	-21.28509	-21.47607	-21.59555	-21.79298	-22.56885	-23.31807
AR 4	-23.52613	-23.70278	-23.71868	-23.94437	-23.96019	-23.9878
AR 5	-23.6479	-23.73176	-23.79204	-23.97993	-23.99847	-24.06152

Cointegration Rank Test Using Trace

H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process
0	**0**	0.9678	7588.1384	24.08	NOINT	Constant
1	**1**	0.5591	1690.1700	12.21		
2	**2**	0.1524	283.9763	4.14		

Long-Run Parameter Beta Estimates

Variable	1	2	3
r	1.00000	1.00000	1.00000
BIAS	-34.32319	-31.55515	-338.67190
lnMACD	8.16656	-64.86788	32.97914

Adjustment Coefficient Alpha Estimates			
Variable	1	2	3
r	-1.06460	0.08510	0.12274
BIAS	-0.00047	0.00120	0.00104
lnMACD	-0.01546	0.01785	0.00133

The VARMAX Procedure

Type of Model	VECM(1)
Estimation Method	Maximum Likelihood Estimation
Cointegrated Rank	1

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
D_r	**AR1_1_1**	-1.06460	0.02930			r(t-1)
	AR1_1_2	36.54062	1.00564			BIAS(t-1)
	AR1_1_3	-8.69415	0.23927			lnMACD(t-1)
D_BIAS	**AR1_2_1**	-0.00047	0.00025			r(t-1)
	AR1_2_2	0.01618	0.00846			BIAS(t-1)
	AR1_2_3	-0.00385	0.00201			lnMACD(t-1)
D_lnMACD	**AR1_3_1**	-0.01546	0.00042			r(t-1)
	AR1_3_2	0.53069	0.01437			BIAS(t-1)
	AR1_3_3	-0.12627	0.00342			lnMACD(t-1)

Covariances of Innovations

Variable	r	BIAS	lnMACD
r	3.80846	0.03167	0.04317

Covariances of Innovations			
Variable	**r**	**BIAS**	**lnMACD**
BIAS	0.03167	0.00027	0.00036
lnMACD	0.04317	0.00036	0.00078

Information Criteria	
AICC	-18.7815
HQC	-18.778
AIC	-18.7815
SBC	-18.772
FPEC	6.971E-9

Forecasts					
Variable	**Obs**	**Forecast**	**Standard Error**	**95% Confidence Limits**	
r	**1719**	-0.08276	1.95153	-3.90768	3.74217
	1720	-0.03370	1.97755	-3.90964	3.84223
	1721	-0.04227	2.06181	-4.08335	3.99881
	1722	-0.04078	2.13011	-4.21571	4.13415
	1723	-0.04104	2.19834	-4.34970	4.26763
	1724	-0.04099	2.26416	-4.47867	4.39669
	1725	-0.04100	2.32819	-4.60417	4.52217
	1726	-0.04100	2.39049	-4.72627	4.64427

Forecasts					
Variable	**Obs**	**Forecast**	**Standard Error**	**95% Confidence Limits**	
BIAS	**1719**	-0.00165	0.01642	-0.03384	0.03054
	1720	-0.00162	0.02272	-0.04615	0.04290
	1721	-0.00163	0.02768	-0.05589	0.05263
	1722	-0.00163	0.03188	-0.06410	0.06085
	1723	-0.00163	0.03558	-0.07136	0.06811
	1724	-0.00163	0.03893	-0.07793	0.07468
	1725	-0.00163	0.04202	-0.08398	0.08072
	1726	-0.00163	0.04489	-0.08961	0.08636
lnMACD	**1719**	-0.00242	0.02789	-0.05709	0.05224
	1720	-0.00171	0.03178	-0.06400	0.06058
	1721	-0.00184	0.03537	-0.07117	0.06750
	1722	-0.00181	0.03858	-0.07743	0.07380
	1723	-0.00182	0.04155	-0.08325	0.07961
	1724	-0.00182	0.04431	-0.08867	0.08503
	1725	-0.00182	0.04692	-0.09377	0.09014
	1726	-0.00182	0.04938	-0.09861	0.09497