J Syst Sci Complex (2012) 25: 641–674

MODELING AND FORECASTING OF STOCK MARKETS UNDER A SYSTEM ADAPTATION FRAMEWORK

Xiaolian ZHENG · Ben M. CHEN

DOI: 10.1007/s11424-012-1034-0
Received: 10 February 2011 / Revised: 20 August 2011

Abstract This paper adopts the concept of dynamic feedback systems to model the behavior of financial markets, or more specifically, the stock market from a dynamic system point of view. Based on a feedback adaptation scheme, the authors model the movement of a stock market index within a framework that is composed of an internal dynamic model and an adaptive filter. The output-error model is adopted as the internal model whereas the adaptive filter is a time-varying state space model with instrumental variables. Its input-output behavior, and internal as well as external forces are then identified. Special attention has also been paid to the recent financial crisis by examining the movement of Dow Jones Industrial Average (DJIA) as an example to illustrate the advantage of the proposed framework. Supported by time-varying causality tests, five influential factors from economic and sentiment aspects are introduced as the input of this framework. Testing results show that the proposed framework has a much better prediction performance than the existing methods, especially in complicated economic situations. An application of this framework is also presented with focuses on forecasting the turning periods of the market trend. Realizing that a market trend is about to change when the external force begins to exhibit clear patterns in its frequency responses, the authors develop a set of rules to recognize this kind of clear patterns. These rules work well for stock indexes from US, China and Singapore.

Key words Complex systems, financial modeling, financial systems, market forecasting, system economics.

1 Introduction

A financial market is a complex system involving various interacting factors from social, political to psychological aspects[1]. Although a great stride has been made in discovering its properties by many economic models, the emerged essential limitations of these methods still prompt people to search for new approaches especially those that can perform well in complicated situations. Viewing the stock market as a highly complex system, we propose a dynamic feedback adaptation framework originated from systems and control theory to model the essential behavior of the market.

Xiaolian ZHENG · Ben M. CHEN (Corresponding author)
The Department of Electrical & Computer Engineering, The National University of Singapore, Singapore 117576, *Singapore.* Email: xiaolianzheng@nus.edu.sg; bmchen@nus.edu.sg.
◇ *This paper was recommended for publication by Editor Shouyang WANG.*

As a significant role in the financial system, the stock market attracts large attentions and efforts in understanding its behavior. The stock market is based on the economic theory of demand and supply; therefore, the market price of a stock is determined by the interaction of aggregate demand and supply schedules[2]. Traditional models in time series analysis are expanded to investigate how the stock market correlates to other economic components. Popularly used models include the multivariate autoregressive moving average model in forecasting marketing time series with explanatory variables[3] by expanding the famous ARMA model[4], the vector auto regressive (VAR) framework in measuring market responses to exogenous shocks[5–6], the autoregressive conditional heteroscedasticity (ARCH) and the generalized autoregressive conditional heteroskedasticity (GARCH) models[7] as well as their family in the market risk management[8], and so on. However, these models are inadequate for simulating the behavior of the whole market in terms of their basic assumptions and structures. Benefiting from the development of multidisciplinary fields, theories and methodologies from physics, engineering, and even social science are integrated with economics in building new models for the stock market in recent decades. One famous cluster is the computational intelligence in finance represented by artificial neural networks (ANN) and support vector machine (SVM). With their ability of flexibly simulating complex nonlinear relationships between the input influential factors and the output, they perform well in stock price prediction[9–10]. Although SVM has been shown to be resistant to some inherent limitations that ANN has, the results provided by these non-parametric methods still lack transparency.

Another breakthrough is contributed by system economics, a group of methods that analyze the financial market as a complex system. In 1980, Michael[11] pointed out some potential areas where economics may interact with systems theory, one of which is the behavior finance, a very popular area today. It aims to study psychological biases of investors and the consequent influences on the market. These ideas lead to models combining the knowledge from economics, psychology, neuroscience, and systems science. The agent-based model is one such useful tool for understanding the market microstructure. Poggio, et al.[12] proposed a four-component repeated double-auction market to conduct six experiments about the market dynamics and several properties such as the market efficiency, price deviation and the distribution of wealth. Chen and Yeh[13] focused on the beliefs and behavior of traders, while Chen and Liao[14] showed that the stock price-volume causal relation exists without any explicit assumptions. All of these results are based on an agent-based artificial stock market. As Orrell and McSharry[15] presented in their survey paper, system dynamics is another powerful and promising approach in analyzing the financial market as a complex system. Considered as a "tool for learning a complex world", system dynamics has found successful applications in a wide range of areas including in the financial market analysis[16]. Cao and Wang[17] showed how information technology, control, and computer technology can contribute to financial engineering. We also note that Gerencsér and Mátyás[18] proposed a behavioral finance model based on the systems theory. In their model, the behavior of agents in the stock market is depicted by a closed loop system where the plant is the market and the controller is the belief and behavior of agents. Although their model provides a new perspective for understanding the behavior of stock market, it mainly focuses on the online regression of an autoregression (AR) model rather than the structure or the dynamics of the system.

In this work, following our preliminary result reported earlier in [19–20], we develop a complete framework for modeling financial markets from a system dynamics view point. More specifically, we propose a feedback adaptation structure to systematically model the stock market so that the market dynamics and properties can be better understood and captured. Under our framework, the modeling process is considered as a system identification problem, in which the real stock market is treated as an unknown plant and the proposed identification model is tuned by feeding back the matching error. It turns out that our proposed framework gives the best one-step-ahead prediction results as compared to the traditional methods such as the well known ARMA model with exogenous input (ARMAX). We have also applied the proposed approach to the forecasting of major market turning periods. When analyzing the external force in frequency domain by Fourier transform, the frequency contents exhibit significant peaks in certain specific time. As evidenced by the results of statistical tests, the appearing time of these significant frequency peaks provides information on the major turnings in the trend of stock price movement. The forecasting rules have successfully been tested in the DJIA, the Composite Index of Shanghai Stock Exchange (SSE), and the Straits Times Index of Singapore Stock Exchange (STI).

The utilization of systems theory in modeling financial market is relatively new. As pointed out by Orrell and McSharry[15], a framework that can tie many of the foundations of systems economics together remains to be established. Our work contributes to this area by presenting a unique system adaptation framework to model the financial market as a complex system. Although we put our attention to the stock market, with their particular influential factors as input, the structure of our framework as well as the methods to establish its every component could be used to analyze other financial markets. In contrast to single dynamic models, this framework provides not only more accurate prediction results but also a better description of the real market. The information flow of the real market is clearly reflected in the hierarchy of this system based framework.

Market force is a fundamental and key factor in our model, representing the dynamic characteristics of the system. In the theory of nonlinear dynamic economics, both endogenous and exogenous forces cause the economic fluctuation. Like the physical system, a financial market also has its fast and slow dynamics which are corresponding to these two types of forces. As such, we have two subsystems in the identification model, an internal model and an adaptive filter, respectively taking the fast and slow dynamics of the market prices into consideration. In other words, these two subsystems reflect the working schemes of endogenous and exogenous forces. Simulation results support that this framework works well in the prediction of both stock prices and turning periods with this structure. Besides the models selected in this paper, different ones could also be used for these two subsystems as long as they can catch these two types of forces. From the aspect of variables involved, this framework integrates both fundamental and technical analyses, taking the complementary nature of both analyses as an advantage. In this way, market behavior like information feedback and dynamics of the system could be better captured and analyzed in this system adaptation framework.

The outline of the paper is organized as follows. In Section 2, we present the framework of dynamic adaptation system for modeling the stock market and discuss its working scheme. Section 3 discusses the details of framework input selection, the estimation algorithms and results for both internal model and adaptive filter, and related analysis based on the selected influential factors. Section 4 assesses the forecasting capability of our proposed framework together with a comparison with some traditional approach. An application of this framework that concerns the forecasting of turning periods of the market trend is given in Section 5. Finally, we draw some concluding remarks in Section 6.

2 A System Adaptation Framework of Stock Market

Viewing the stock market as a highly complex system, we propose a closed-loop adaptation framework based on the well established systems theory. It provides a systematic way to model the market behavior. Inspired by ideas used in identifying engineering systems or systems in general, the stock market modeling process is treated as the problem of identifying a dynamic plant shown in Figure 1. The input-output behavior of the stock market is represented by the identification model $\widehat{S}$ with its output $\widehat{p}$, the estimated stock price. The actual stock price p is the output of the real stock market S. Both S and $\widehat{S}$ have the same input r which consists of external influential factors of the stock market. The structure and parameters of $\widehat{S}$ can be determined to minimize the identification error e, the difference between the actual and estimated stock price.

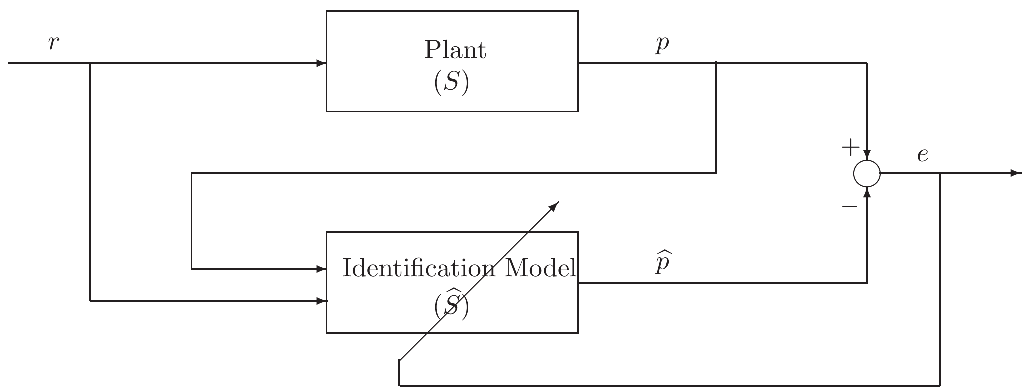


Figure 1 The system adaptation framework

The purpose of our framework is to capture the dynamics of the real stock market, so that with the appropriate input, the identification model can generate an accurate prediction result. Force and feedback are the two fundamental factors in this framework. It is believed that information inside and outside the stock market would act as forces to regulate the share price. As pointed out earlier in the introduction, both endogenous and exogenous forces act upon the stock market, making the market have slow and fast dynamics, respectively. Therefore, we establish the identification model with an internal model I and an adaptive filter A (as depicted in the block diagram of Figure 2) to capture these two types of dynamics.

The internal model, I, is considered as a trend generator in which the internal force is assumed to exist, making the estimated price have the same trend as the actual price. It is independent of the adaptive filter, mainly taking the dynamics of actual prices into consideration. In this way, it is rational to assume that the dynamics of the internal model change at a relatively slower pace, so that it can be approximated by some time-invariant systems. An output-error (OE) model is employed in this internal part, using only the historical actual prices to estimate the internal stock price $\widehat{p}_i$ of the next step. The differences between actual and internal prices are then estimated and defined as the external force:

$$e_i(n) = p(n) - \widehat{p}_i(n). \tag{1}$$

In the literature, information outside the stock market always accounts for the external forces to the market. The adaptive filter A is thus introduced to capture the influence of these information. It generates the estimated external force $\widehat{e}_i$ of the next step by analyzing major influential factors of the stock market as well as the historical information of the external force itself. Working as a cycle generator, this adaptive filter compensates the identification error e by catching the fast dynamics of the market. Young[21] described many approaches

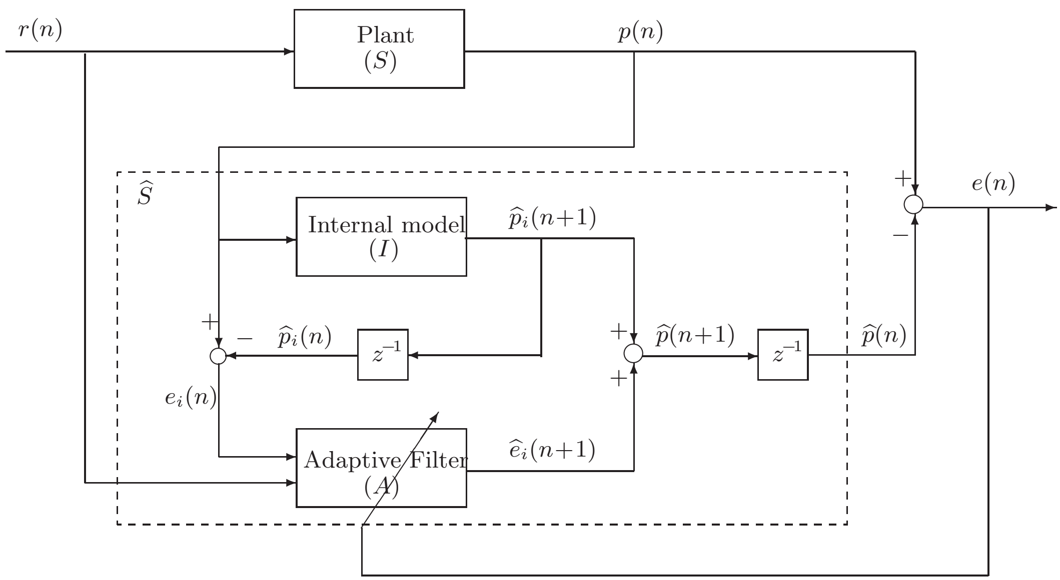


Figure 2 Block diagram of structure of the proposed system adaptation framework

of modeling economic time series based on time-varying parameter estimation. Binder and Merges[22] found that using time-varying coefficients based on the cluster regression can make a very high proportion of market volatility explainable. It is also rational that the impact of each influential factor on the stock price movement is time-dependent, so that their significances should vary at different stages. As such, it is natural for us to choose a time-varying adaptive filter linked to the external inputs. We employ a state space model with instrumental variables for the adaptive filter. Adding up the outputs of these two parts, we have the estimated stock price of the next step:

$$\widehat{p}(n+1) = \widehat{p}_i(n+1) + \widehat{e}_i(n+1). \tag{2}$$

Regarding to feedback, the other key factor, it exists in the stock market both in terms of information and psychology. It helps to select information and adjust trading strategies. That is how feedback works in a system. In the identification process, e is fed back to adjust the adaptive scheme of A. At the same time, feedback also exists inside the internal model to represent the effects of the historical data.

2.1 Internal Model

The internal model is composed of three parts, the block diagram of which is depicted in Figure 3. First, the historical stock price is preprocessed by the exponential moving average (EMA) model:

$$p_{\text{ema}}(n) = \frac{2}{N+1}p(n) + \left(1 - \frac{2}{N+1}\right)p_{\text{ema}}(n-1), \tag{3}$$

where N is an adjustable parameter, p and p_{ema} denote the actual and EMA prices, which are the input and output of this EMA part, respectively. Through (3), the transfer function from

$p(n)$ to $p_{\rm ema}(n)$, denoted by $H_{\rm ema}(z)$, can be expressed as

$$H_{\rm ema}(z) = \frac{P_{\rm ema}(z)}{P(z)} = \frac{\frac{2}{N+1}}{1-\left(1-\frac{2}{N+1}\right)z^{-1}}. \tag{4}$$

Based on the EMA data, an output-error (OE) model of multi-inputs and single-output (MISO) is employed to model the inherent evolution of stock prices. Its input $u_{\rm oe}(n)$ consists of current and $k-1$ previous samples of the EMA prices, i.e.,

$$u_{\rm oe}(n) = \begin{bmatrix} u_{{\rm oe},1}(n) \\ u_{{\rm oe},2}(n) \\ \vdots \\ u_{{\rm oe},k}(n) \end{bmatrix} = \begin{bmatrix} p_{\rm ema}(n) \\ p_{\rm ema}(n-1) \\ \vdots \\ p_{\rm ema}(n-k+1) \end{bmatrix}. \tag{5}$$

The transfer function of this MISO OE model, $H_{\rm oe}(z)$, is characterized by

$$H_{\rm oe}(z) = \begin{bmatrix} H_1(z) & H_2(z) & \cdots & H_k(z) \end{bmatrix}, \tag{6}$$

where for $j = 1, 2, \cdots, k$. $H_j(z)$ is the transfer function in the jth channel which is a rational function of z given as

$$H_j(z) = \frac{\Phi_j(z)}{\Psi_j(z)} \tag{7}$$

with

$$\Phi_j(z) = \phi_{j,1} + \phi_{j,2}z^{-1} + \cdots + \phi_{j,n_\Phi}z^{-n_\Phi+1}, \tag{8}$$
$$\Psi_j(z) = 1 + \psi_{j,1}z^{-1} + \cdots + \psi_{j,n_\Psi}z^{-n_\Psi}. \tag{9}$$

In this way, considering $d_i(n)$, the disturbance of the system which is assumed to be a white noise, the output of this part is the estimated EMA price of the next step:

$$\widehat{P}_{{\rm ema},n+1}(z) = H_{\rm oe}(z)U_{\rm oe}(z) + D_i(z), \tag{10}$$

where $\widehat{P}_{{\rm ema},n+1}(z)$ and $U_{\rm oe}(z)$ are the Z-transform of $\widehat{p}_{\rm ema}(n+1)$ and $u_{\rm oe}(n)$, respectively. After all the parameters are identified, the estimated EMA price is transformed back to the internal price by $H_{\rm ema}^{-1}(z)$, so that we have the internal price $\widehat{p}_i(n+1)$ as the output of the entire internal model.

2.2 Adaptive Filter

As depicted in Figure 4, in order to identify the dynamic influence of exogenous influential factors on the stock market, a time-varying state space model with instrumental variables is used as the adaptive filter. We adopt the following time-varying model[23−24] to represent the input-output relationship of the adaptive filter:

$$\begin{aligned}\widehat{e}_i(n+1) = & -\sum_{j=1}^{n_{\rm a}} a_j(n)e_i(n-j+1) + \sum_{j=1}^{n_1} b_{1,j}(n)r_1(n-\delta_1-j+1) \\ & + \cdots + \sum_{j=1}^{n_m} b_{m,j}(n)r_m(n-\delta_m-j+1) + \sum_{j=1}^{n_{\rm a}} a_j(n)\varepsilon(n-j+1),\end{aligned} \tag{11}$$

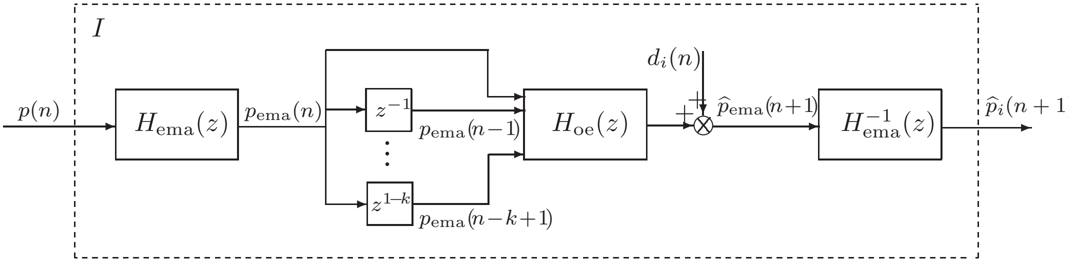


Figure 3 The internal model

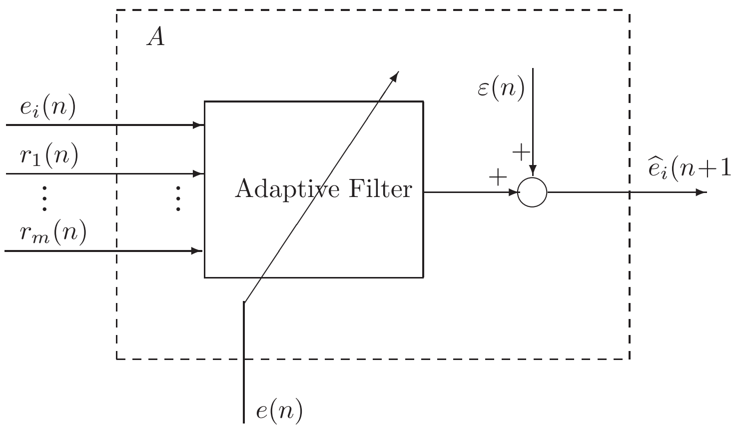


Figure 4 The adaptive filter

where $\widehat{e}_i(n+1)$ is the output of the filter denoting the estimated external force in the next step; r_j denotes a selected influential factor with a lag order of δ_j; ε is a Gaussian white noise representing uncertainties; a_j, $j=1,2,\cdots,n_{\mathrm{a}}$, $b_{i,j}$, $i=1,2,\cdots,m$, $j=1,2,\cdots,n_i$, are time-dependent parameters to be identified. It is fair to assume that all these time-varying parameters are statistically independent. We note that such a model is an extension of the autoregressive exogenous variable (ARX) model by allowing the noise to be colored. The dynamics of such a force generation system are represented by those time-varying parameters, a_j, $j=1,2,\cdots,n_{\mathrm{a}}$, $b_{i,j}$, $i=1,2,\cdots,m$, $j=1,2,\cdots,n_i$, which are assumed to evolve according to a random walk. For each of these time-varying parameters, it is assumed to be characterized by the following general stochastic model:

$$x(n)=x(n-1)+\eta_x(n), \tag{12}$$

where x represents a time-varying parameter to be identified, and η_x is a white noise input associated with x, which is assumed to have a normal distribution $N(0,Q_x)$.

By defining these parameters as state variables, we convert the model in (11) into a state-space form, which can be easily identified by some well established approaches, e.g., the instrumental variable methods. Constructing vectors $X(n)$ and $H(n)$ as

$$X(n) = [\ a_1(n)\ \ a_2(n)\ \ \cdots\ \ a_{n_{\rm a}}(n)\ \ b_{1,1}(n)\ \ \cdots\ \ b_{1,n_1}(n)\ \ \cdots\ \ b_{m,1}(n)\ \ \cdots\ \ b_{m,n_m}(n)\ ]^{\rm T}, \tag{13}$$

$$\begin{aligned} H(n) = [\ & -e_i(n)\ \ -e_i(n-1)\ \ \cdots\ \ -e_i(n-n_{\rm a}+1)\ \ r_1(n-\delta_1)\ \ \cdots\ \ r_1(n-\delta_1-n_1+1) \\ & \cdots\ \ r_m(n-\delta_m)\ \ \cdots\ \ r_m(n-\delta_m-n_m+1)\ ], \end{aligned} \tag{14}$$

and we have the state space model:

$$X(n) = X(n-1) + \eta(n), \qquad \eta(n) \sim N(0, Q), \tag{15}$$

$$\widehat{e}_i(n+1) = H(n)X(n) + \mu(n), \qquad \mu(n) \sim N(0, \sigma^2), \tag{16}$$

where η consists of all the white noise inputs associated with the parameters to be identified,

$$\mu(n) = \sum_{j=1}^{n_{\rm a}} a_j(n)\varepsilon(n-j+1). \tag{17}$$

Obviously, Q is a diagonal matrix with its diagonal elements being the variance of the related input noise. We aim to estimate $X(n)$, denoted by $\widehat{X}(n)$, such that the resulting system identification error $e(n)$ is minimized.

Since the regressors include lagged terms of the output of the adaptive filter, which are correlated to $\varepsilon(n)$, we introduce an instrumental variable $u(n)$ for eliminating possible estimation bias:

$$\begin{aligned} u(n) = & -\sum_{i=1}^{n_{\rm a}} \widehat{a}_i(n)e_i(n-i+1) + \sum_{j=1}^{n_1} \widehat{b}_{1,j}(n)r_1(n-\delta_1-j+1) \\ & + \cdots + \sum_{j=1}^{n_m} \widehat{b}_{m,j}(n)r_m(n-\delta_m-j+1) \end{aligned} \tag{18}$$

with

$$\widehat{X}(n) = [\,\widehat{a}_1(n)\widehat{a}_2(n)\ \cdots\ \widehat{a}_{n_{\rm a}}(n)\ \widehat{b}_{1,1}(n)\ \cdots\ \widehat{b}_{1,n_1}(n)\ \cdots\ \widehat{b}_{m,1}(n)\ \cdots\ \widehat{b}_{m,n_m}(n)\,]^{\rm T}. \tag{19}$$

It can be shown that such an instrumental variable is highly correlated to the original regressor vector but uncorrelated to $\varepsilon(n)$. Under this setting, we construct the instrumental vector $\widehat{H}(n)$ as

$$\begin{aligned} \widehat{H}(n) = [& -u(n)\ -u(n-1)\ \cdots\ -u(n-n_{\rm a}+1)\ r_1(n-\delta_1)\ \cdots\ r_1(n-\delta_1-n_1+1) \\ & \cdots\ r_m(n-\delta_m)\ \cdots\ r_m(n-\delta_m-n_m+1)\,]. \end{aligned} \tag{20}$$

Descriptions of the estimation process using this instrumental vector are given in detail in Subsection 3.4.

3 Model Parameter Estimation

Before estimating the parameters of both internal and external models, some preliminary works are carried out including the selection of influential factors and causality investigations between the market external force and these input factors. We would like to point out that the external influential factors, which behave as the inputs of our adaptation framework, play an essential role in our model. Many results can be found in the literatures concerning with the forces that determine the stock price, most of which focus on economic factors. A famous and also one of the earliest studies of this topic is Chen, et al.[25], the results of which suggested that a set of economic variables including industrial production, interest rate, inflation and oil price, are important in explaining expected stock returns. Campbell and Ammer[5], Binder and Merges[22], and Kim[26] are other studies of this sort. Besides economic factors, investor sentiment becomes a nonneglectable factor in recent studies. Baker and Wurgler[27] provided a general method to measure the investor sentiment, with which they have demonstrated how the sentiment affected the stock market as a whole. In addition, the reactions of the stock market to noise trader risk[28], aggregate earning news[29], and even terrorism activity[30] were investigated. These works are all based on different sentiment indicators, in which Chicago Board Options Exchange Volatility Index (VIX) is a common one for measuring the fear of investors.

We consider both of these two aspects in our model. Five leading indicators are selected as the major input of the DJIA based on the empirical analysis of several economic and sentiment indicators. We study the psychological effects of investors not only by using Chicago Board Options Exchange DJIA Volatility Index (VXD) as a sentiment indicator but also in terms of constructing the interest rate indicator. Time-varying causality test shows that all these five indicators Granger cause the external force, statistically evidencing the rationality of this input selection. With verified input factors, we introduce the estimation algorithms and results regarding to both internal model and adaptive filter, followed by the corresponding result analysis. The results also suggest that the determinant factors accounting for the stock market fluctuations change over time. With the correct inputs, our framework can provide very good results in modeling the market behavior.

3.1 Data Description

The stock market during recent global financial crisis provides a valuable sample for both investors and researchers. Investigations of the market behavior in this period can help us understand what the major exogenous influential factors are, how the market reacts to them, and most importantly, how the market is likely to behave if it confronts with similar situations in the future. Therefore, the DJIA which is often referred to as the snapshot of the health of the US market is used as an example in identifying the model associated with our adaptation framework. We use daily closing prices of the DJIA from January 2008 to December 2010, the period from the beginning of recent financial crisis to very recently. All the stock market data are obtained from the database of Yahoo Finance.

Since the major factors that affect the stock market vary at different stages and in order to pay special attention to the period in the recent financial crisis, we separate the whole period into four subperiods, from January to August 2008 (Subperiod 1), from September to December 2008 (Subperiod 2), from January 2009 to April 2010 (Subperiod 3) and from May to December 2010 (Subperiod 4). The variance of external force is another element that is considered in the partition of subperiods. Subperiod 2 starts from September 2008, the month when the financial crisis hit its most critical stage. This subperiod has the largest variance of external forces. After getting relatively small, the variance of external forces in May 2010 which is the beginning of Subperiod 4 starts to become bigger again.

3.2 Input Selection

The input is determined by a double selection method. First, some leading influential factors are selected based on an empirical analysis. After that, some statistical tests like causality and correlation tests are adopted to examine all these factors. Influential factors that can pass all these tests are selected as the input of our framework.

3.2.1 Influential Factor Selection

General economic situation and consensus of investors toward the market are two most important factors that affect the movement of the stock market. For this reason, we focus on two groups of indicators: economic indicators and sentiment indicators. Economic indicators are used to judge the health of the economy, whilst sentiment indicators can be a measurement of the situation of demand and supply. Sun[31] provided an empirical analysis on a wide range of indicators based on which we select five leading indicators within the scope of these two groups to investigate their contributions in generating the market force: three economic ones and two sentiment ones.

1) Interest Rate Indicator (IRI)

In economic indicators, the interest rate is one of the acknowledged determinant factors that have a direct impact on the general trend of the stock market. The size of federal funds rate target (FFRT) changes is often applied to investigate the role that the interest rate plays in the US stock market[32]. Recently, since the federal funds future rate (FFFR) embodies people's expectation on the monthly average of the daily effective funds rate, it is frequently used in measuring market reactions to interest rate changes[33−34]. We construct a similar indicator that takes people's expectation on a daily basis into consideration. Since the daily effective federal funds rate (DEFFR)* is a volume-weighted average of rates on trades arranged by a group of federal funds brokers who report to the Federal Reserve Bank of New York each day, we use the differences between DEFFR and FFRT, which is named interest rate indicator (IRI), as a gauge of the influence of interest rates on the US stock market:

$$\text{IRI} = \text{DEFFR} - \text{FFRT}. \tag{21}$$

2) Oil Price (OP)

The second one is from the commodity area, in which we select changes in the oil price (OP) as another key economic indicator. The oil price is based on the Cushing West Texas Intermediate (WTI) spot price, obtained from Energy Information Administration. Its change $c(n)$ is calculated by

$$c(n) = 100 \times \frac{o(n) - o(n-1)}{o(n-1)}, \tag{22}$$

where o is the value of oil price.

3) Baltic Dry Index (BDI)

The third economic indicator is the Baltic Dry Index (BDI), a daily average of global shipping prices for dry bulk cargoes. Functioning as an assessment of global trade and free of manipulation and speculation make it an excellent leading indicator of economic activity. We also use its changes calculated in the same way as (22) to measure its power in affecting the stock market.

4) Chicago Board Options Exchange DJIA Volatility Index (VXD)

Two sentiment indicators we choose are the VXD and the currency of Euro to Japanese Yuen (EUR/JPY). VXD is a futures contract based on the prices of options on the DJIA traded at the Chicago Board Options Exchange (CBOE). It is a kind of stock fear index, reflecting the market's expectation of the DJIA volatility over the next 30 days. The changes of these two variables are calculated in a similar way as that in (22).

*The definition from the Federal Reserve Bank of New York.

5) Currency of Euro to Japanese Yuen (EUR/JPY)

Financial crisis would inevitably disseminate fear, making investors tend to be risk-aversion. The currency pair EUR/JPY is a good one in such type of indicators. In addition, the empirical comparison of the DJIA and the EUR/JPY by Sun[31] suggested that they feed off of one another. That is the reason why we choose a currency pair that does not include US dollar.

Before testing, all these input influential factors will be adjusted to have the same date as the stock prices. In other words, if stock price exists on some day but some factors do not, we will fill them up with their data on the previous day; if stock prices do not exist while some factors do, we will remove input data on these days. After that adjustment, data normalization is performed to all the series but IRI by

$$\tilde{v}(n) = \frac{v(n) - \bar{v}}{s_v}, \tag{23}$$

where $v(n)$ denotes the original series, $\tilde{v}(n)$ denotes the data after normalization, $\bar{v}$ is the mean of series v in a specific period, and s_v is the standard deviation of v in the same period.

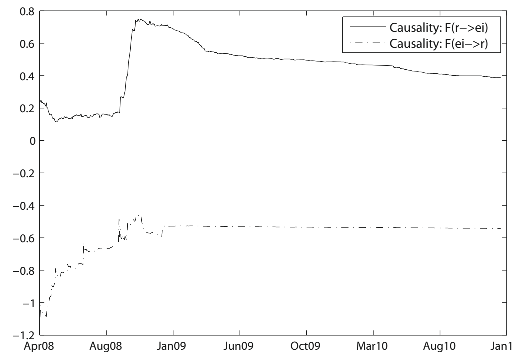
3.2.2 Causality Test

In order to establish a correct relationship of the DJIA and these indicators before introducing them to the framework, we need to investigate the causality relationship between the external force of the DJIA and each influential factor. A time-varying causality test is then conducted. By extending the idea of conventional Granger causality test for linear regression model, Geweke[35] quantified the causality relationship in the form of linear dependence between signals based on VAR models. The definition of Geweke can be easily extended in dynamic models. In fact, some applications have been reported in the area of neuroscience[36−37]. The similar method is used in this paper that a time-varying causality test is implemented based on the adaptive filter.

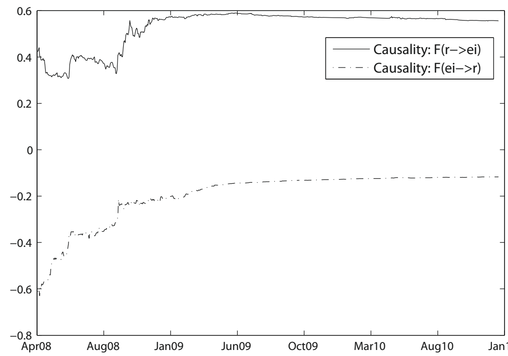
Considering the input and output time series r and e_i, the time-varying strength of causality from r to e_i and from e_i to r are defined as $F_{r\to e_i}(n)$ and $F_{e_i\to r}(n)$. If $F_{r\to e_i}(n) > F_{e_i\to r}(n)$, we can say that r Granger causes e_i in the time n, and vice versa. Under these definitions, time-varying causality testing results between the external force and the selected five influential factors are presented in Figure 5. $F_{r\to e_i}(n)$ is clearly lager than $F_{e_i\to r}(n)$ in all situations, providing an evidence for the validity of the market input selection in the statistical sense.

3.2.3 Correlation Test

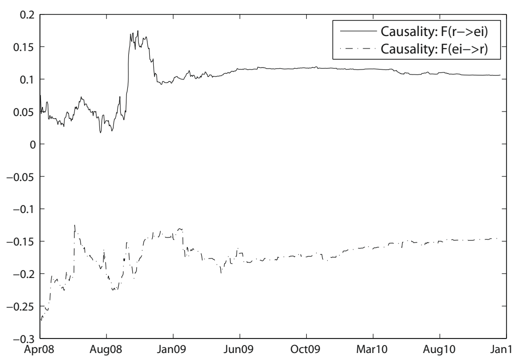
Besides the causal relationship, before combining these influential factors as the input to the system, we still need to figure out the correlation relationship between them. To do so, the Spearman rho correlation coefficient is calculated between these five factors to exclude some redundant ones. As given in Table 1, the EUR/JPY shows significant correlations with the OP and VXD, and the OP also significantly correlates to the VXD. Since the EUR/JPY and VXD are all belong to sentiment indicators but the OP is not, we finally choose IRI, OP, BDI, and VXD as the input to the framework.



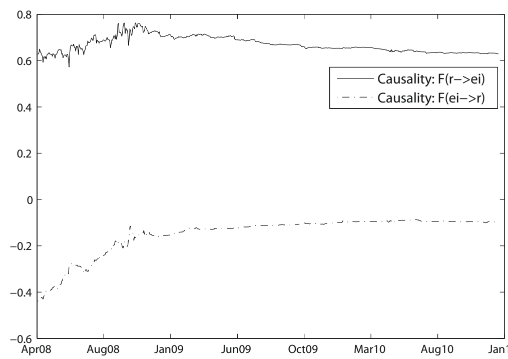
(a) Causality between external force and IRI



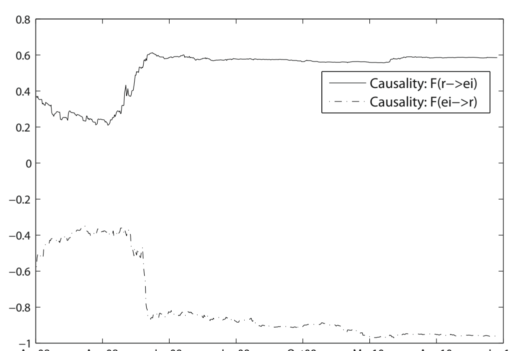
(b) Causality between external force and OP



(c) Causality between external force and BDI



(d) Causality between external force and VXD



(e) Causality between external force and EUR/JPY

Figure 5 Time-varying causality tests for the input influential factors

Table 1 Spearman rho correlation coefficients of the selected input influential factors

		IRI	OP	BDI	VXD	EUR/JPY
IRI	Correlation	1.0000				
	Sig. (2-tailed)	—				
OP	Correlation	0.028	1.0000			
	Sig. (2-tailed)	0.450	—			
BDI	Correlation	0.065	0.040	1.0000		
	Sig. (2-tailed)	0.076	0.279	—		
VXD	Correlation	0.020	−0.258*	−0.011	1.0000	
	Sig. (2-tailed)	0.583	0.000	0.760	—	
EUR/JPY	Correlation	0.014	0.418*	0.000	−0.398	1.0000
	Sig. (2-tailed)	0.710	0.000	0.993	0.000	—

*Correlation is significant at the 0.01 level (2-tailed).

3.3 Internal Model Estimation

Due to the daily data used, the sampling time is 1 day. Therefore, the reciprocal of the frequency is the number of days. According to the popular choice in creating Moving Average Convergence/Divergence (MACD), 12-days EMA is used, i.e., $N = 12$ in (3). The input number k of the OE model is selected to be 3, i.e., the EMA prices of previous two days and today constitute the input. Five-years daily closing prices previous to the starting point of the testing data are used as the training set to tune the parameters of the OE model, i.e., the DJIA daily closing prices from January 2003 to December 2007 in this case. The model parameters are estimated by the commonly used prediction error method in [38].

Let the estimation error of OE model be denoted by $e_{\mathrm{oe}}(n) = p_{\mathrm{ema}}(n) - \widehat{p}_{\mathrm{ema}}(n)$. Through (5) to (10), we have

$$\begin{aligned} e_{\mathrm{oe}}(n) &= p_{\mathrm{ema}}(n) - H_{\mathrm{oe}}(z)u_{\mathrm{oe}}(n-1) \\ &= p_{\mathrm{ema}}(n) - \sum_{j=1}^{k} H_j(z)p_{\mathrm{ema}}(n-j). \end{aligned} \tag{24}$$

Define parameter vector θ as

$$\theta = \left[\, \theta_1 \ \theta_2 \ \cdots \ \theta_k \,\right], \tag{25}$$

where

$$\theta_j = \left[\, \psi_{j,1} \ \psi_{j,2} \ \cdots \ \psi_{j,n_\Psi} \ \phi_{j,1} \ \phi_{j,2} \ \cdots \ \ \phi_{j,n_\Phi} \,\right]. \tag{26}$$

(24) can be written as

$$e_{\mathrm{oe}}(n) = p_{\mathrm{ema}}(n) - f\left(\theta, u_{\mathrm{oe}}(n-1)\right). \tag{27}$$

Here, $f(\theta, u_{\rm oe}(n-1))$ is the function of θ and $u_{\rm oe}(n-1)$ †. Under previous definitions, the model parameter vector θ is estimated to minimize the cost function which is defined as

$$V_K(\theta) = \sum_{n=1}^{K} e_{\rm oe}^2(n) = \sum_{n=1}^{K} \left[p_{\rm ema}(n) - f(\theta, u_{\rm oe}(n-1))\right]^2, \tag{28}$$

where K is the sample size. Obviously, $V_K(\theta)$ is nonlinear with respect to θ. The solution is to use Newton-Raphson method to iteratively minimize this cost function $V_K(\theta)$. The estimation process is then given by

$$\widehat{\theta}_K^{(i+1)} = \widehat{\theta}_K^{(i)} + \mu_K \left[\frac{\partial^2 V_K(\widehat{\theta}_K^{(i)})}{\partial \theta^2}\right]^{-1} \frac{\partial V_K(\widehat{\theta}_K^{(i)})}{\partial \theta} \tag{29}$$

where μ_K is the step size and $\frac{\partial^2 V_K(\widehat{\theta}_K^{(i)})}{\partial \theta^2}$ is the Hessian matrix that gives the search direction. The detailed algorithm of prediction error method please refer to [38].

The estimation process starts with zero initial states and stops when the improvement less than 10^{-4}. We then have the OE model for the DJIA during the period of interest

$$H_{\rm oe}(z) = \begin{bmatrix} \dfrac{0.9574z^{-1} - 0.5034z^{-2} - 0.321z^{-3} + 0.565z^{-4}}{1 - 1.329z^{-1} + 0.7312z^{-2}} \\ \dfrac{-5.127z^{-1} + 2.086z^{-2} + 0.3914z^{-3} + 0.8506z^{-4}}{1 - 0.62z^{-1} - 0.1239z^{-2}} \\ \dfrac{2.417z^{-1} + 3.049z^{-2} - 2.298z^{-3} - 1.82z^{-4}}{1 - 0.3215z^{-1} - 0.4643z^{-2}} \end{bmatrix}^{\rm T}, \tag{30}$$

and obtain the related external force. Figure 6 shows the external force where the shaded areax denote Subperiod 2 and Subperiod 4.

As depicted in the first shaded area, the variance of external forces from September 2008 to the end of 2008 is extremely large. That is also the reason for the partition of the investigation subperiods. We would like to point out that this is a general rule we adopt for estimating the OE model in our framework and it is also used in identifying market turning periods to be presented later.

3.4 Adaptive Filter Estimation

From (15), it is clear that the unknown parameters, or hyperparameters, in the covariance matrix Q determine the variations of the state variables. Therefore, the hyperparameter optimization is carried out first, and then the Kalman filter is used to perform the recursive estimation and prediction. We adopt the advanced maximum likelihood (ML) method[23,40] to

†In multiple-input and single-output OE model, let $\gamma_j(k) : k = 0, 1, \cdots$ be the impulse response parameters for $H_j(z)$, and the output can be written as

$$p_{\rm ema} = \sum_{j=1}^{pr} c_j x_j(n) + d_i(n),$$

where

$$c_{ir+j} = \gamma_j(i+1), \qquad j = 1, 2, \cdots, r; i = 0, 1, \cdots$$
$$x_{ir+j}(n) = u_{{\rm oe},j}(n-i-1), \qquad j = 1, 2, \cdots, r; i = 0, 1, \cdots$$

Therefore, $f(\theta, u_{\rm oe}(n-1))$ can be finally transformed to the linear regression form like the single-input and single-output OE model. Detailed definitions and identification processes are given in [39].

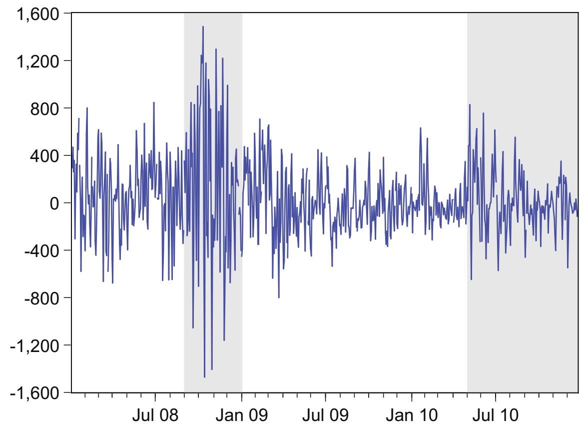


Figure 6 External force of the DJIA

estimate the hyperparameters based on a set of historical data. The major innovation of this method is introducing the Noise Variance Ratio (NVR) matrix Q_r and the $\widehat{P}$ to exclude the disturbance variance σ^2, which are respectively defined as

$$Q_r = \frac{Q}{\sigma^2} \quad \text{and} \quad \widehat{P} = \frac{P}{\sigma^2}, \tag{31}$$

where P is the prediction error covariance matrix associated with the estimated state vector $\widehat{X}$. In what follows, we adopt the notation $(n|n-1)$ to denote the estimation of a specific variable in step n conditional on information up to step $n-1$. In this way,

$$\widehat{P}(n|n-1) = \frac{P(n|n-1)}{\sigma^2} = \frac{1}{\sigma^2} E[(\widehat{X}(n) - \widehat{X}(n|n-1))(\widehat{X}(n) - \widehat{X}(n|n-1))^{\mathrm{T}}]. \tag{32}$$

Given a set of historical data $e_i(1), e_i(2), \cdots, e_i(\tau)$, the Log-likelihood function of $e_i(\tau+1), e_i(\tau+2), \cdots, e_i(K)$ conditional on these previous information is calculated via commonly used prediction error decomposition by

$$\begin{aligned} &\log L\big(e_i(\tau+1), \cdots, e_i(K) l | e_i(1), \cdots, e_i(\tau)\big) \\ = &\frac{-(K-\tau)}{2} \log 2\pi - \frac{1}{2} \sum_{n=\tau+1}^{K} \log \mathrm{var}(\widehat{e}(n)) - \frac{1}{2} \sum_{n=\tau+1}^{K} \frac{\widehat{e}(n)^2}{\mathrm{var}} (\widehat{e}(n)), \end{aligned} \tag{33}$$

where $\widehat{e}(n) = e_i(n) - \widehat{H}(n)\widehat{X}(n|n-1)$ is the one-step-ahead prediction error with the instrumental variables and $\mathrm{var}(\widehat{e}(n))$ is its variance which is obtained by

$$\mathrm{var}(\widehat{e}(n)) = \sigma^2 \left[1 + \widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n) \right]. \tag{34}$$

The Log-likelihood function needed to be maximized is

$$\log L(\cdot) = \frac{-(K-\tau)}{2}\log 2\pi - \frac{1}{2}\log\sigma^2 - \frac{1}{2}\sum_{n=\tau+1}^{K}\log\left[1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\right]$$
$$-\frac{1}{2\sigma^2}\sum_{n=\tau+1}^{K}\frac{\widehat{e}(n)^2}{1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)}. \tag{35}$$

Partially differentiating (35) with respect to σ^2 conditional on given information, we have the estimation of σ^2 as $\widehat{\sigma}^2$:

$$\widehat{\sigma}^2 = \frac{1}{K-\tau}\sum_{n=\tau+1}^{K}\frac{\widehat{e}(n)^2}{1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)}. \tag{36}$$

Substituting (36) into (35) and removing the constant term as well as the negative sign, it is transformed to a compact form as

$$\log L(\cdot) = \sum_{n=\tau+1}^{K}\log\left[1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\right]$$
$$+(K-\tau)\log\left[\frac{1}{K-\tau}\sum_{n=\tau+1}^{K}\frac{\widehat{e}(n)^2}{1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)}\right]. \tag{37}$$

Hence, hyperparameters are estimated by minimizing (37). The optimization algorithm we used is the Nelder-Mead simplex direct search method[41]. With the estimated hyperparameters, the Kalman filter provides recursive prediction and updating processes, in which the identification error $e(n)$ is fed back to tune the states. The algorithm is stated as follow:

1) Prediction:

$$\widehat{X}(n|n-1) = \widehat{X}(n-1|n-1), \tag{38}$$
$$\widehat{P}(n|n-1) = \widehat{P}(n-1|n-1) + Q_r, \tag{39}$$

2) Updating:

$$\widehat{X}(n|n) = \widehat{X}(n|n-1) + \widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\left[1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\right]^{-1}e(n), \tag{40}$$
$$\widehat{P}(n|n) = \widehat{P}(n|n-1) - \widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\left[1+\widehat{H}(n)\widehat{P}(n|n-1)\widehat{H}^{\mathrm{T}}(n)\right]^{-1}\widehat{H}(n)\widehat{P}(n|n-1). \tag{41}$$

First, an AR model is used to determine the lag order of $e_i(n)$ in model (11). For the DJIA data under test, when the order reaches 4, the Durbin-Watson statistic is 1.98 and the p value of Breusch-Godfrey test is 0.43, both showing no autocorrelation in the residuals. Thus, we set n_{a} to be 4. Considering to include about half-month impact, the lag orders of all the indicators are set to be 10. An interesting result is that when the time delay of IRI starts from third order, the model yields the best results. More specifically, the lag order of the IRI in our test is from 3 to 12.

We first investigate the influence of each influential factor by using only this factor with AR component at each time, and then combine some of them together as the input to test the predictive ability of the whole framework. When performing the hyperparameter optimization, to be practical, only the data from January to August 2008 is used and the hyperparameters of the AR part are fixed. In this way, we have 14 state variables and the same number of hyperparameters in the Q_r matrix. Initial hyperparameters are all set as 10^{-4} and initial covariance matrix $\widehat{P}$ is set as a diagonal matrix with all its diagonal entries equal to 10^5.

With the estimated hyperparameters, one-step-ahead prediction results are obtained through (38) to (41). Table 2 summarizes the associated adjusted R^2, the coefficient of determination[‡], which is defined as

$$\bar{R}^2 = 1 - \frac{K-1}{K-w-1}\frac{\sum(y-y_{\mathrm{h}})^2}{(y-\bar{y})^2}, \tag{42}$$

where K is the sample size, w is the total number of regressors in the model, y is the actual output, y_{h} is the predicted output, and $\bar{y}$ is the mean of y. This statistic increases only if the new term improves the model more than would be expected by chance; therefore, we use it to measure the contribution of each indicator to the external force.

Table 2 Prediction results (Adjusted R^2)

	Subperiod 1	Subperiod 2	Subperiod 3	Subperiod 4
AR	20.05%	13.40%	13.41%	14.77%
AR and IRI	31.35%	70.63%	17.67%	17.02%
AR and OP	44.23%	55.10%	51.30%	46.67%
AR and BDI	21.58%	25.80%	25.93%	19.86%
AR and VXD	59.97%	57.56%	47.56%	49.91%

3.5 Preliminary Analysis

Generally, in the US stock market, the force variation which is also considered as the market variance in Subperiod 2 is better explained than it in Subperiod 3. Among these indicators, the IRI is the most dominant factor in Subperiod 2 while the OP behaves best in Subperiod 3. The VXD is another significant factor by contributing nearly equally to the external forces in all subperiods compared with the OP. It even becomes most important in Subperiod 4. This situation is also consistent with our assumption that the determinant factors of the stock market change over time. From the beginning of Subperiod 2, the US financial crisis entered a dangerous phase. It is characterized by a series of collapses of financial institutions, such as the takeover of both Fannie Mae and Freddie Mac, the bankruptcy of Lehman Brothers and the merge of Bank of America and Merrill Lynch. The devastating effects they triggered began to spread to all other economic sectors from then, destroying the confidence of investors. Although an unprecedented \$700 billion rescue plan enacted by the US government finally got passed in October 2008, the original plan was rejected by the House of Representatives on September 29, 2008. All these events resulted in an extreme instability in the US stock market, which is shown in the market external force from September to December 2008 (Figure 6). It is also represented in the expectation of investors on the market, which we quantify by the IRI. Since the Federal Reserve decided to remain the FFRT at 2% on September 16, 2008, it deviated the major expectation on the open market. Although the FFRT was consecutively cut for two times in October 2008, the IRI still underwent a dramatic fluctuation. However, the IRI is not

‡The coefficient of determination of each step, R^2, is to provide the proportion of variability in a data set that is accounted for by a statistical model[42]. Adjusted R^2, i.e., $\bar{R}^2$, is a slight modification of R^2 by adjusting for the sample size and degree of freedom.

that important in Subperiod 3. A possible reason for that lies in the fact that Federal Reserve slashed the FFRT to 0.25% in December 2008 and kept it at this record low till today. That makes the IRI stay in a rational range.

In Subperiod 3, the OP seriously affects the US stock market, explaining more than 50% of the market variance. Unlike the interest rate, the relationship between the OP and the stock market is complicated and debatable in both academic and investment areas. However, as supported by our testing results, the OP leads the stock market and its influence exceeds much more than the IRI from 2009 to 2010. From the beginning of 2009, the OP began to rebound but was still rational. The DJIA moved in the same directions with it as a follower. The strengthening of oil demand can be considered as a symptom of the consolidation of economic recovery in USA. It is not surprising that McKay made the headline of the Wall Street Journal "Oil Lifts Dow to 10062.94 In Late Rally" on October 16, 2009 to analyze this situation in [43].

4 Market Forecasting

With the establishment of the market modeling framework, we are now ready to present the forecasting or prediction capability of the proposed approach. To be more specific, we are going to use the DJIA as an example to identify all necessary model parameters and utilize the obtained model to predict (or forecast) the one-day-ahead closing price of the market. As usual, the prediction performance is to be respectively measured by the mean absolute error (MAE) and root mean squared error (RMSE):

$$\mathrm{MAE} = \frac{1}{K}\sum_{n=1}^{K} |e(n)|\,, \tag{43}$$

$$\mathrm{RMSE} = \left[\frac{1}{K}\sum_{n=1}^{K} e^2(n)\right]^{\frac{1}{2}}, \tag{44}$$

where K is the number of samples interested.

4.1 Identified Model Parameters

Given the performances of four selected influential factors, while combining them together, the hyperparameters are estimated again with the complete input and output data from January to August 2008, the same training period in estimating the adaptive filter. The initial hyperparameters and covariances are 0.002 and 10^5, respectively. We obtain the estimation of the hyperparameters of the adaptive filter with a diagonal Q_r, whose main diagonal entries are respectively given as

$$0.0025,\ \ 0.0032,\ \ 1.1226\times 10^{-4},\ \ 2.3812\times 10^{-4}, \tag{45}$$

which are corresponding to the estimation of $\widehat{a}_1, \widehat{a}_2, \widehat{a}_3, \widehat{a}_4$;

$$0.0100,\ \ 0.0211,\ \ 0.0079,\ \ 0.0075,\ \ 0.0380, \tag{46}$$

$$0.7546,\ \ 6.7498\times 10^{-4},\ \ 0.3482,\ \ 0.1456,\ \ 0.1433, \tag{47}$$

which are corresponding to the estimation of $\widehat{b}_{1,1}, \widehat{b}_{1,2}, \cdots, \widehat{b}_{1,10}$, the coefficients associated with the Interest Rate Indicator input;

$$1.7388\times 10^{-4},\ \ 2.6124\times 10^{-4},\ \ 0.0056,\ \ 0.0656,\ \ 3.0712\times 10^{-4}, \tag{48}$$

$$1.9431\times10^{-4},\ \ 0.0011,\ \ 7.1140\times10^{-4},\ \ 0.0584,\ \ 0.0141, \tag{49}$$

which are corresponding to the estimation of $\widehat{b}_{2,1}, \widehat{b}_{2,2}, \cdots, \widehat{b}_{2,10}$, the coefficients associated with the Oil Price input;

$$1.7249\times10^{-4},\ 0.0153,\ \ 0.0015,\ \ 0.2589,\ \ 0.0023, \tag{50}$$

$$5.3271\times10^{-4},\ 7.9624\times10^{-4},\ 2.9077\times10^{-4},\ 4.3319\times10^{-4},\ 0.0267, \tag{51}$$

which are corresponding to the estimation of $\widehat{b}_{3,1}, \widehat{b}_{3,2}, \cdots, \widehat{b}_{3,10}$, the coefficients associated with the Baltic Dry Index input; and finally,

$$1.9714\times10^{-4},\ \ 5.9100\times10^{-4},\ \ 7.3694\times10^{-4},\ \ 3.5652\times10^{-4},\ \ 1.9268\times10^{-4}, \tag{52}$$

$$0.0012,\ \ 2.5038\times10^{-8},\ \ 5.3246\times10^{-4},\ \ 3.8426\times10^{-4},\ \ 1.1683\times10^{-4}, \tag{53}$$

which are corresponding to the estimation of $\widehat{b}_{4,1}, \widehat{b}_{4,2}, \cdots, \widehat{b}_{4,10}$, the coefficients associated with the Chicago Board Options Exchange DJIA Volatility Index input.

With the above given Q_r and initial conditions: $\widehat{X}(0) = 0$, $\widehat{P}(0)$ being a diagonal matrix with all its diagonal entries equal to 10^5, and $\widehat{H}(0)$ being set to its corresponding input data, the one-day-ahead prediction results can be obtained through the iterative process as given in (38) to (41). We compare the prediction ability of this framework over the traditional ARMAX model. The lag orders for ARMAX model are set similarly as used in our framework, 4 for AR and MA (moving average) terms and 10 for all the exogenous inputs.

4.2 Performance Analysis

Summarized in Table 3 is the prediction error results of our proposed framework and the traditional ARMAX method. It is clear that the system adaptation framework has outperformed the ARMAX a great deal. Figures 7 and 8 are the detailed prediction results of two time frames from September 2008 to January 2009 and from May 2010 to December 2010, respectively. The effectiveness of the framework structure, the ability of the dynamic design of the adaptive filter, and the inevitable function of the internal model are comprehensively testified.

Table 3 Comparison of the prediction accuracies between the ARMAX approach and the proposed framework

	Subperiod 1		Subperiod 2		Subperiod 3		Subperiod 4	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ARMAX	120.48	152.80	215.89	268.86	90.18	117.96	86.13	118.73
New Approach	31.49	40.86	29.45	38.29	29.10	38.62	36.14	48.76
Improvement	73.86%	73.26%	86.36%	85.76%	67.73%	67.26%	58.04%	58.93%

The prediction ability of our framework is excellent especially when it is in a complicated economic environment. The average daily changes of closing prices from subperiod 1 to 4 are 116.73, 263.82, 85.61, and 80.72, respectively. Compared with the MAE generated by ARMAX which are slightly smaller or even larger than this average daily price changes, our framework provides much smaller and much reasonable MAE in all subperiods. In this sense, our framework is much more meaningful. In Subperiod 2, while the traditional model fails to measure the dynamics of the market, our framework gives very precise predictions. All these results show the ability of this framework in understanding the dynamics of complex systems. In addition, this framework is developed in a systematic way that makes it flexible and easily to be expanded. Within its structure, the internal model and the adaptive filter could take

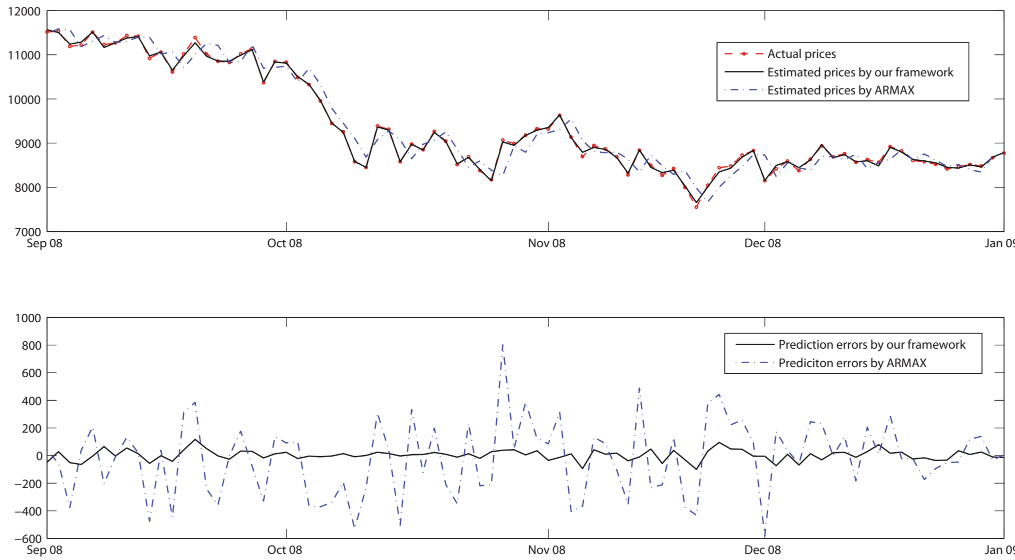


Figure 7 Prediction results of the proposed framework and the ARMAX approach in Subperiod 2

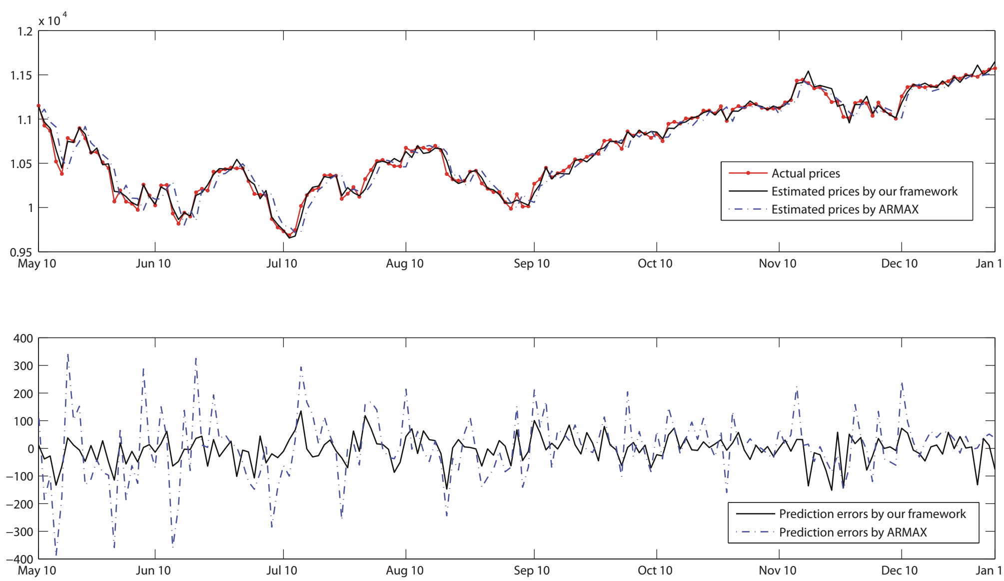


Figure 8 Prediction results of the proposed framework and the ARMAX approach in Subperiod 4

other models according to different modeling requirements or different markets. However, we also notice that both MAE and MSE keep increasing in Subperiods 3 and 4. The possible reason lies in the weakening of the determinant effect of input. Since the input plays a very important role on this framework, and in Subperiod 3 and 4 there is no such determinant factor as the IRI in Subperiod 2, the prediction ability of this framework is consequently weakened. Moreover, the variance of external force is smaller in Subperiod 3 than in Subperiod 2, but our framework works better in Subperiod 2, also proving the importance of having appropriate influential factors as the input. This result signals that from Subperiod 4 the input factors should be reconsidered.

5 Forecasting of the Market Trend Turning Periods

The crucial turning in the market trend can also be forecasted through the frequency contents of the external force. In this section, we first introduce the forecasting methods and process, and then present related results in New York, Shanghai and Singapore Stock Exchanges. After that, the Bai-Perron test is performed to find the structure breaks in some macroeconomic variables, supporting the forecasting results.

5.1 Turning Period Forecasting

When a long-term market trend alters its direction, there are some essential changes in the internal dynamics of the market. In our framework, such internal dynamics are measured by the internal model I, while the differences between the output of the internal model and the actual prices present in the external force. Thus, for the purpose of revealing such important market properties, the external force is analyzed in the frequency domain by the Fourier transform. Analyzing financial time series in the frequency domain is a comparatively underexplored area in the literature. Chen[44−45] and Turhan-Sayan and Sayan[46] applied the so-called time-frequency representation to present the information of stock market in both time and frequency domains simultaneously. Chen[44−45] proved the existence of persistent chaotic cycles in the US stock market which are around 3 to 5 years. In this sense, five-years data is enough for estimating the internal model. The general rules for choosing training data set are that the closing prices of previous five years are chosen to estimate the internal model when we start forecasting a new series. After two turning periods (a pair of peak and bottom) are found, the internal model will be reestimated by using the previous testing data set. By doing so, the internal model can better capture the market dynamics with a cycle of data.

We find that the frequency contents of the external force vary a lot in different time intervals. Generally, the significant components in the power spectrum can be clearly read at some specific time periods. Tests show that this sudden change always signifies the major turning periods in the stock market. More specifically, we select many sampling time spans sharing the same starting point, but with increasing length, such as half a month, one month, and two months. Due to the complicated wrestle of bullish and bearish forces in the stock market, reversals always happen in a period rather than on a single day. It is in good consistence with our tests. When a clear frequency pattern with significant peaks appears, it stays for several months before disappearing again. Therefore, the turning period defined in this paper is the period when the frequency contents have a clear pattern. To find out the next turning period, we use the same method but with a new starting point that is chosen from the beginning of a new trend. In order to avoid the big fluctuation during the transition of two trends, we tentatively select four to five months after the previous turning period as the new starting point. When the trend is confirmed, this starting point will be adjusted accordingly, forwarding or backwarding. Admittedly, this rule is empirical, but due to the long term trend we investigated, it is practical and proved to work well in the real stock markets. We would like to emphasize that this forecasting method only uses the historical data till the testing date. With testing results, we could instantly know that whether this testing date is in a turning period or not. No future date is needed in this method.

5.1.1 Frequency Pattern Identification Rules

In order to perform an automatic and objective judgement, we develop a set of rules to recognize the frequency patterns during turning periods. Our rules concern three aspects: magnitude, distance between frequency components, and range of one cluster of frequency components. Figure 9 demonstrates these three key elements, of which details are given below.

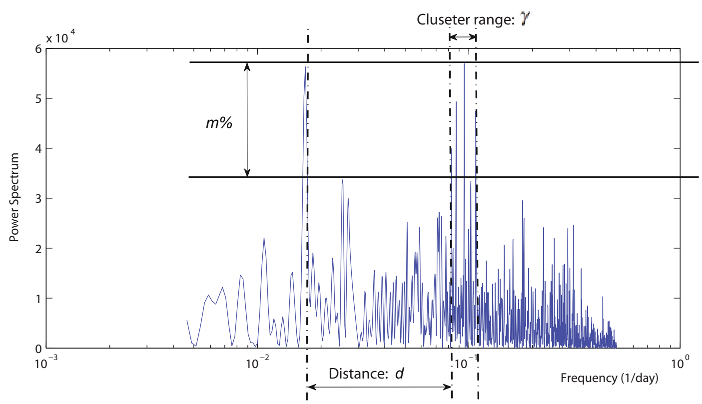


Figure 9 Elements in frequency pattern identification

With frequency contents of the external force in a specific time, our identification is composed of two procedures. First of all, based on the above three elements, some features of current frequency pattern are calculated which is called the feature extraction procedure. This procedure is used to identify and cluster dominant frequency components. The judgement procedure is performed after that by comparing these features with corresponding preset threshold values. This judgement considers the distribution of the dominant frequency components and the changing process of frequency contents.

1) Feature extraction procedure

Step 1 Only frequency components whose magnitudes are larger than $m\%$ of the maximum magnitude are considered while others are removed.

Step 2 Distance d between two adjacent frequency component is defined as their frequency differences (in 1/day) in a logarithmic scale. If the distance is larger than ι, we categorize these two frequency components to different clusters. The number of clusters is calculated in this step.

Step 3 The range of each cluster γ is defined as the logarithmic distance between the first and last frequency components in one cluster. The location of each cluster is defined as the frequency of the component with largest magnitude in this cluster. Both range and location of each cluster are calculated in this step.

2) Judgement procedure

Rule 1 The number of clusters should be less than κ.

Rule 2 If there is only one cluster, its range should be less than ϱ_1. If there are more than one clusters, each range should be less than ϱ_2.

Rule 3 If the frequency pattern fulfills Rules 1 and 2, the number of clusters and their locations are compared with them in the last identification. If the number is the same and the location change of each cluster is less than ς, the ending time of this sampling period is identified as in a turning period.

Sometimes noises may cause changes in the number of clusters during a turning period, such as suddenly increasing or decreasing one cluster. If current number of clusters is one less than previous and the location change of each cluster is still within ς, it is still considered as in a turning period. If one more cluster suddenly appeares, we remove the originally existed clusters of dominant components and then use Rules 1 and 2 to rejudge whether this new cluster has dominant components. If Rules 1 and 2 are fulfilled, it means there are some newly appearing dominant frequency components, signifying a situation changing. Otherwise, this new cluster could be ignored; thus, if it was in a turning period, the situation continues. In this rejudgement step, κ is set as 1 and the other corresponding threshold values are $\widetilde{m}\%$ and $\widetilde{\varrho}_1$.

5.1.2 Market Tuning Periods

We apply the proposed method to US, China and Singapore stock markets by investigating the DJIA, SSE and STI. For these three markets, the threshold values in the above identification rules are selected as

$$\iota = 0.2, \quad \kappa = 3, \quad \varrho_1 = 0.23, \quad \varrho_2 = 0.13, \quad \varsigma = 0.001, \quad \widetilde{m} = 75\%, \quad \widetilde{\varrho}_1 = 0.15. \tag{54}$$

Take the DJIA as an example, to begin with, daily closing prices of the DJIA from 1990 to 1994 are used to estimate the internal model, which gives

$$H_{\mathrm{oe,DJIA},1}(z) = \begin{bmatrix} \dfrac{-2.566z^{-1} + 2.120z^{-2} - 0.061z^{-3} - 0.371z^{-4}}{1 - 0.197z^{-1} - 0.110z^{-2}} \\ \dfrac{1.866z^{-1} - 1.581z^{-2} - 0.058z^{-3} + 0.537^{-4}}{1 - 0.563z^{-1} + 0.085z^{-2}} \\ \dfrac{1.707z^{-1} - 2.033z^{-2} + 2.263z^{-3} - 1.338z^{-4}}{1 - 0.231z^{-1} - 0.023z^{-2}} \end{bmatrix}^{\mathrm{T}}. \tag{55}$$

Transforming the external force to the frequency domain, the forecasting process for the first turning period is shown in Figure 10. All these nine subpictures have the same starting point which is January 1, 1995, while the sampling time span is increasing from 24 months in the first subpicture to 48 months in the last subpicture. The frequency contents have no clear pattern until one significant peak shows up in the fourth subpicture where the sampling ends at the end of July 1999. According to the judgement of identification rules, this date is the beginning of current turning period. After that, although the sampling time span keeps increasing, the frequency contents of the external force in the corresponding periods have a similar clear pattern. This clear pattern disappears from the last subpicture, signifying the ending of the turning period. The first forecasting process stops at this ending time. Based on the previous definitions and two-procedure identification rules, the first turning period of the DJIA is from August 1 to November 30, 1999. The next forecasting process begins at May 1, 2000, five months after the ending of the first turning period. The forecasting process is shown in Figure 11, where the period from August 1 to December 30, 2002 is the second turning period. In this time frame,

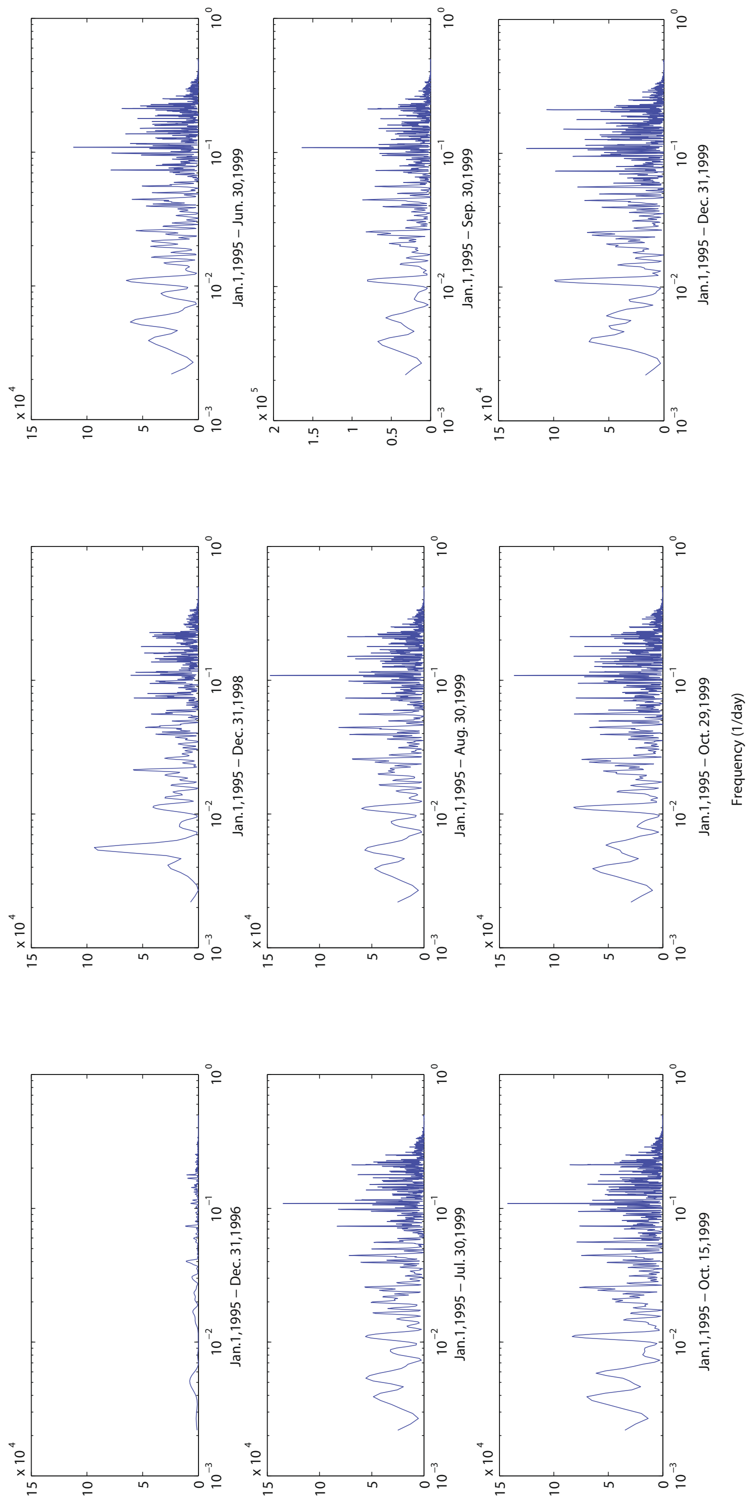


Figure 10 Frequency responses of the external force of the DJIA from January 1995 to December 1999

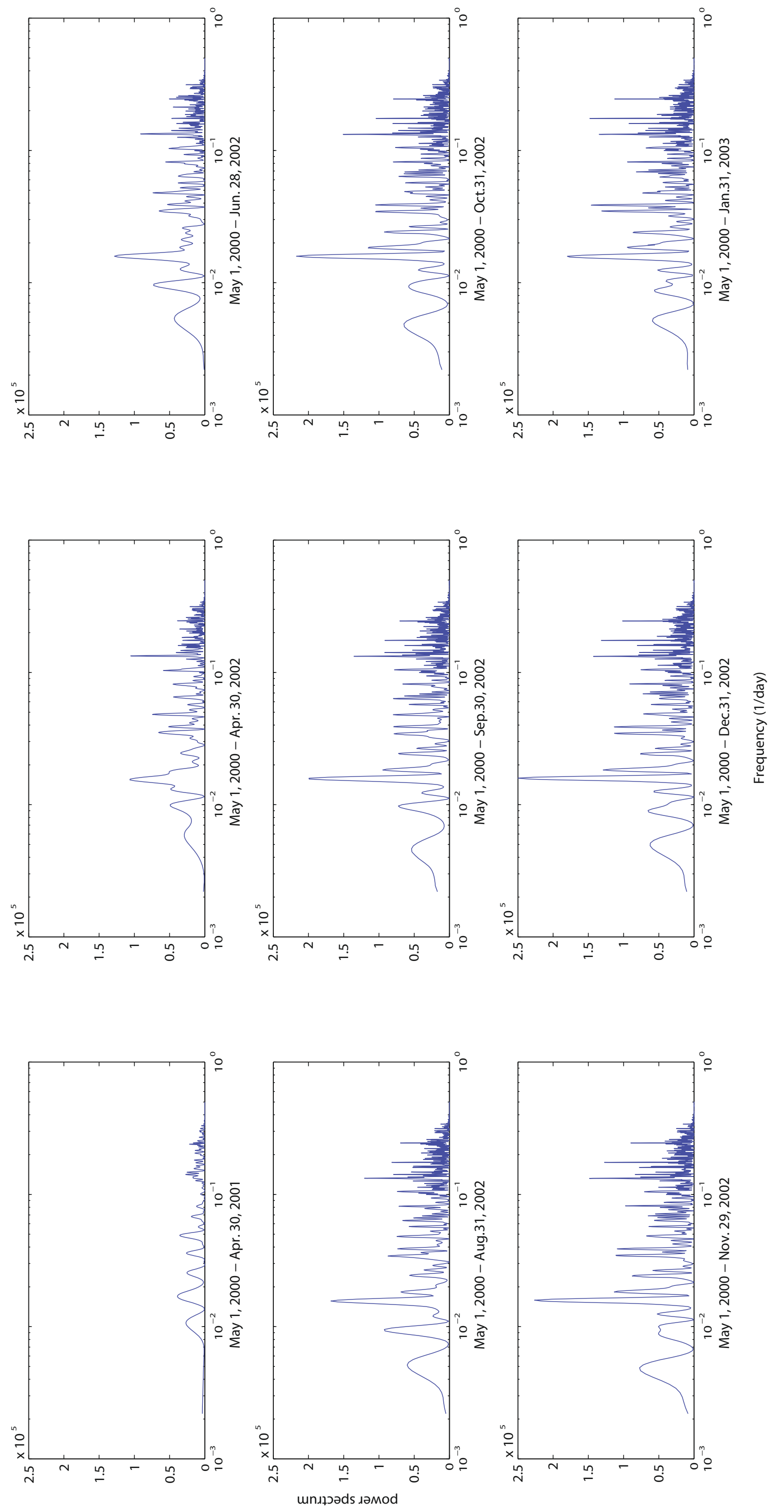


Figure 11 Frequency responses of the external force of the DJIA from May 2000 to January 2003

our rules give two more turning periods which are shown in red shaded areas in Figure 12. They are from February 1 to March 31, 2001 and from September 4 to December 5, 2001. Although they do not correspond to the major market trend turning, a big falling down happened in both periods. This might be the reason for this misjudgement.

After the second turning period has been forecasted, the internal model is reestimated based on the closing prices from January 1, 1995 to January 31, 2003, the previous testing period. The newly estimated internal model is

$$H_{\text{oe,DJIA},2}(z) = \begin{bmatrix} \dfrac{66.59z^{-1} + 58.66z^{-2} - 25.88z^{-3} - 0.942z^{-4}}{1 - 0.445z^{-1} + 0.034z^{-2}} \\ \dfrac{52.18z^{-1} - 24.66z^{-2} - 151.1z^{-3} - 76.65^{-4}}{1 + 0.825z^{-1} - 0.120z^{-2}} \\ \dfrac{2.515z^{-1} - 57.2z^{-2} + 76.16z^{-3} - 25.01z^{-4}}{1 - 1.373z^{-1} + 0.446^{-2}} \end{bmatrix}^{\mathrm{T}}. \tag{56}$$

Using the same rules, two turning periods are found one at a time from July 1, 2003 to April 30, 2010. The first one is from April 25 to December 30, 2007. Although it is composed of two periods, one from April 25 to May 30, 2007 and the other one from October 1 to December 30, 2007, they are considered in the same turning period as they are very close to each other. The second turning period is from December 1, 2008 to March 31, 2009.

By contrast with the stock price evolution of the DJIA in the corresponding time domain (as shown in Figure 12), it is found that these forecasted turning periods are all located in the transition periods of bull and bear market, matching well with the real ones. The DJIA first closed above 11,000 on May 3, 1999 and reached a record high close on January 14, 2000. This date is close to the ending time of the first turning period we forecasted and after that it began to decline. Although there was a rally in the second half of 2000, the DJIA bottomed out at its lowest close since October 1997 on October 9, 2002. That is considered as the ending of 3-year bear market. From March 2003, it turned to recover slightly, and then bulls got overpowered again. Our results also found this period. After reaching the record high of 14,164.53 on October 9, 2007, it started to crash. That is when the latest and the most serious financial crisis since the Great Depression started. It corresponds to our third turning period. In the wake of the global financial crisis and ensuing stock market collapse in 2008, a series of financial support measures were launched to stimulate the economy. The fourth turning period we forecasted matches the recovery of the DJIA from the beginning of 2009. Similarly, after this turning period, the internal model is reestimated as (57) by previous testing data. With this new model, a new turning period is identified which is from February 26 to April 30, 2010. It corresponds to the big slip from late April 2010. Although the DJIA got a rally after July 2010, it gives a warning signal for the fluctuation and a possible trend change afterward. This period is also highlighted in Figure 12.

$$H_{\text{oe,DJIA},3}(z) = \begin{bmatrix} \dfrac{-0.1398z^{-1} + 0.08447z^{-2} + 0.1189z^{-3} - 0.06742z^{-4}}{1 - 1.449z^{-1} + 0.4666z^{-2}} \\ \dfrac{-0.0292z^{-1} - 0.08049z^{-2} + 0.1867z^{-3} - 0.07857^{-4}}{1 - 1.363z^{-1} + 0.3784z^{-2}} \\ \dfrac{2.241z^{-1} - 0.1719z^{-2} - 0.2423z^{-3} - 0.027z^{-4}}{1 + 0.3715z^{-1} - 0.01451^{-2}} \end{bmatrix}^{\mathrm{T}}. \tag{57}$$

In order to illustrate the effectiveness of our forecasting method, the SSE in China and the STI in Singapore are also investigated using the same method and rules. The testing period

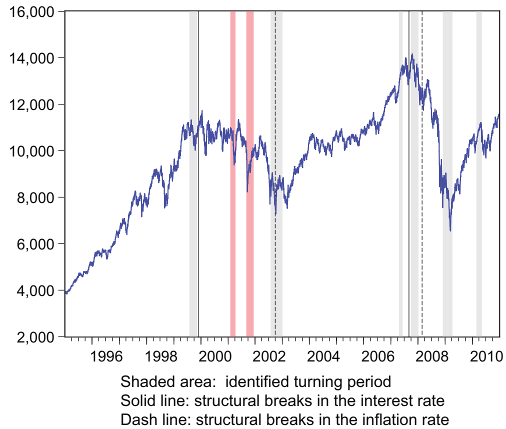


Figure 12 Turning periods of the DJIA

of the SSE is from January 2005 to December 2010, whereas the period from January 2003 to December 2010 is tested for the STI. Therefore, the training data for them are prices from 2000 to 2004 and from 1998 to 2002, respectively, based on which the following $H_{\mathrm{oe,SSE},1}(z)$ and $H_{\mathrm{oe,STI}}(z)$ are obtained to be respectively associated with their internal models:

$$H_{\mathrm{oe,SSE},1}(z) = \begin{bmatrix} \dfrac{2.242z^{-1} - 1.774z^{-2} + 0.734z^{-3} + 0.130z^{-4}}{1 - 0.843z^{-1} + 0.365z^{-2}} \\ \dfrac{-0.132z^{-1} - 0.450z^{-2} + 0.463z^{-3} + 0.245^{z}-4}{1 - 0.209z^{-1} - 0.090z^{-2}} \\ \dfrac{2.710z^{-1} - 3.672z^{-2} - 0.035z^{-3} + 0.341z^{-4}}{1 - 0.739z^{-1} + 0.119^{-2}} \end{bmatrix}^{\mathrm{T}} \tag{58}$$

and

$$H_{\mathrm{oe,STI}}(z) = \begin{bmatrix} \dfrac{5.701z^{-1} + 3.797z^{-2} + 1.257z^{-3} + 0.142z^{-4}}{1 + 0.503z^{-1} + 0.237z^{-2}} \\ \dfrac{0.228z^{-1} - 5.508z^{-2} + 0.609z^{-3} + 1.072^{-4}}{1 - 0.276z^{-1} - 0.106z^{-2}} \\ \dfrac{2.757z^{-1} - 6.686z^{-2} + 5.473z^{-3} - 1.512z^{-4}}{1 - 1.633z^{-1} + 0.690^{-2}} \end{bmatrix}^{\mathrm{T}}. \tag{59}$$

In the forecasting results, two turning periods are found in each market index. For the SSE, they are periods from August 15 to October 5, 2007 and from November 10, 2008 to May 13, 2009. The results are highlighted by shaded areas in Figure 13. The first turning period matches well with the reality but the second one lasts a little longer. In terms of the STI, periods from March 23 to October 25, 2007, from June 2 to September 8, 2008 and from December 1, 2008 to March 2, 2009 are the forecasting results, in which the second one is a misjudgement. The possible reason for this misjudgement may be related to the structural instability in macroeconomy and will be discussed in Subsection 5.2. These turning periods are highlighted in Figure 14.

Similarly, we reestimate the internal models for both the SSE and the STI to track recent market changes. A new turning period in the SSE which is from December 1, 2009 to February 10, 2010 is identified while there is no turning information found in recent STI series. The new internal model for the SSE is given in (60). Our test results suggest that the SSE have entered into a downturn trend. According to the recent trend of the SSE, it dropped a lot in the beginning of April 2010 and kept in fluctuating till now.

$$H_{\mathrm{oe,SSE},2}(z) = \begin{bmatrix} \dfrac{-36.18z^{-1}+42.26z^{-2}-12.51z^{-3}-4.804z^{-4}}{1-1.376z^{-1}+0.6349z^{-2}} \\ \dfrac{-40.64z^{-1}-83.02z^{-2}+2.381z^{-3}+19.39^{z}-4}{1+0.0975z^{-1}-0.2531z^{-2}} \\ \dfrac{2.805z^{-1}+36.38z^{-2}+86.03z^{-3}-80.92z^{-4}}{1-0.7763z^{-1}+0.04463z^{-2}} \end{bmatrix}^{\mathrm{T}}. \tag{60}$$

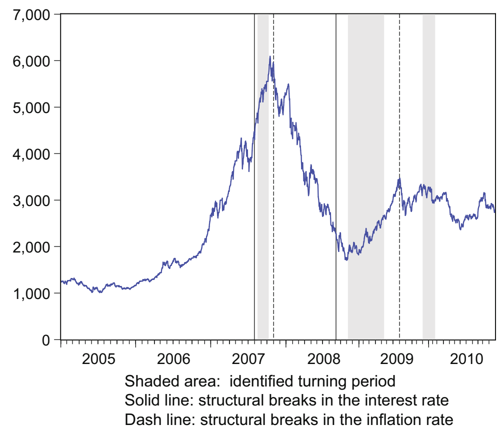


Figure 13 Market turning periods of the SSE

Supported by these empirical results, our method is capable of providing timely information on whether the long-term trend of the stock market is to change. We believe that the internal model of our framework plays an important role in forecasting the turning periods. As mentioned in Section 2, the internal model mainly measures the dynamics of the actual prices. The effect of time-invariant parameters associated with the internal model is to catch the consistent property in the market dynamics. When the internal model can well mimic the market dynamics in a bull or bear market even with some perturbations, no particular property can be found in the differences between actual and internal prices. However, the change of market fundamentals always alters the nature of the series itself. In consequence, the original model of the internal dynamics is not longer valid, reflecting this break by the presence of remarkable peaks in the frequency responses. In order to verify this explanation, we investigate the macroeconomic structural stability in the corresponding periods.

5.2 Structural Changes in the Macroeconomic Situation

As stated in the previous section, the stock market interacts with the macroeconomy. It is believed that the stock price is influenced by the macroeconomic factors in the long term. In

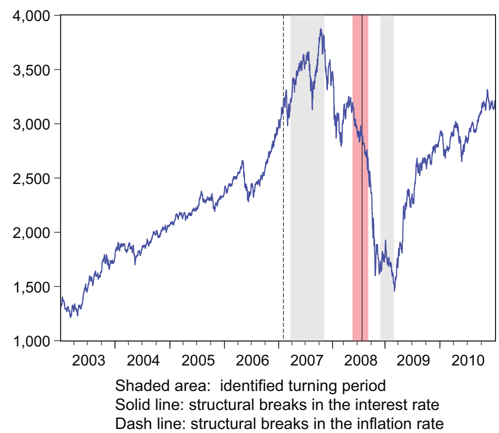


Figure 14 Market turning periods of the STI

this way, some macroeconomic factors are commonly used to judge whether the stock market is undergoing a correction or a major turning. Except for the interest rate which is the acknowledged most important macroeconomic factor to the stock market, a substantial body of the literature considers the inflation rate to be another macroeconomic factor that seriously affects the stock market. For example, Niemira[47] confirmed the relationship of the inflation rate and stock market cycle and suggested using it as an indicator to predict the turning points in stock market. Kim and In[48] used wavelet analysis to evidence this relationship in different time scale. When structural breaks of the interest rate and the inflation rate exhibit, they usually coincide with important economic events such as market structure changes and financial crisis. Reflecting to the stock market, a sharp crash or a rally presents. In this way, we choose the interest rate as the major indicator and the inflation rate as the complementary indicator to represent the economic situation. Structural stabilities of these two series in US, China, and Singapore of the same periods§ are investigated. Monthly money market interest rate and Consumer Price Index (CPI) (from EIU country database) are used in this study. Based on CPI data $C(n)$, the inflation rate $\pi(n)$ is defined as

$$\pi(n) = 100 \times [\ln(C(n)) - \ln(C(n-1))] \, . \tag{61}$$

By now, most of the existing methods for detecting structural breaks are statistical tests. The classical ones are Chow test for a single known break[49], Andrew-Ploberger test for a single unknown break[50], and Bai-Perron test for multiple unknown breaks[51−52]. They have shown their power in providing many evidences of structural instability in economic or financial time series. We employ the well developed Bai-Perron test to detect multiple structural breaks in these series. In our case, AR model is adopted as the basis for the linear model regression. The optimal lag order is set based on Akaike Information Criterion (AIC) with maximum ten lags. Durbin-Watson statistic and the Breusch-Godfrey serial correlation LM test (B-G LM Test) are used to judge whether the lag order is appropriate. The recommended two-steps testing strategy is adopted. The first step is to use a double maximum test which is considered to

§Since recent macroeconomic data are still unreachable, we only investigate the macroeconomic series before October 2010.

be more powerful than SupF test. If its statistics are significantly important, a sequential test will be carried out to determine the number of the breaks. Otherwise, structural stability is considered. Using this strategy, we have the test results given in Table 4.

For all the series, Durbin-Watson statistics are close to 2 and the p values of LM test are all larger than 0.05, denoting that there is no autocorrelation in the residuals with their optimal lags. The results of the double maximum test to these six series are all significant at least at 5% level, revealing their structural instability. Therefore, a sequential test with upper bound $M = 5$ is conducted, i.e., we consider at most 5 structural breaks in each series. It provides the number and locations of the breaks. Regarding to the test results, there is no structural break found in recent turning periods which may attribute to the trimming effect in the test. Due to the trimming factor, a proportion λ of the observations is trimmed at each end. It is selected according to the length of the series in order to make sure that the observations before or after the break point are sufficient to estimate the regression relationship. Moreover, the available data is before October 2010. In this way, to get a reliable test result in recent turning periods, we need to wait for more data. In order to compare these results, we illustrate the locations of these structural breaks together with the forecasted turning periods (see Figures 12–14).

In Figure 12, the US interest rate and inflation rate exhibit two structure breaks each, showing that the first three turning periods of the DJIA undergo a structural change in the economy. Due to the introduction of dot-com era, moderate monetary policy and other economic stimulus events, the DJIA increased rapidly from the end of 1980s, especially in 1990s. The US government adopted contractionary monetary policy to curb inflation in 1999, which partially explains the market structural change at that time and the consequent downturn in the stock market. Although US interest rate does not signal a structural break in 2003, a series of devastating events, such as September 11 attacks, Enron and WorldCom scandals and the collapse of IT industry gave big shocks to the stock market. The first trading day after September 11 attacks, the DJIA fell 7.1%. These shocks make stock market fluctuate and turn down a lot during that period. Considering the structural change in the US inflation rate, it is within that turning period. Both interest rate and inflation rate exhibit a break in 2007, which is caused by the global financial crisis originated from the US subprime mortgage crisis. It has led to plunging property prices, a slowdown in US economy, and billions in losses by stock investors. After that, the Federal Reserve kept the interest rates at record low levels till now and the CPI stayed flat. That is the reason for no structural break presenting in these two series during that period. However, the US economy began to show encouraging signs of recovery from the second half of 2009. Continuous historically low interest rates combined with other stimulus in the economy has shown their influence on the stock market that the DJIA started its rally from March 2009. Our model successfully catches this change in the market dynamics.

When it comes to the market in China, as shown in Figure 13, the first break point in the interest rate and the inflation rate matches the first turning period we forecasted. The second break point of the interest rate is located slightly before the second turning period, while the second break point of the inflation rate is close to the most recent turning period. In Singapore, the only break in the inflation rate is very close to the first turning period. However, the interest rate shows its only structure break in the middle of the misjudged turning period which is four month before the real one. Therefore, the misjudgement of turning period may attribute to the structural change in the macroeconomy and that inversely supports our explanation to the phenomenon that dominant frequency components always appear in turning periods. The comparison is presented in Figure 14.

Table 4 Bai-Perron test of structure breaks in macroeconomic indicators for the US, China and Singapore markets

Country	Optimal Lag	Durbin-Watson stat	B-G LM Test (p value)	UDmax	WDmax	$SupLR_T(l+1\|l)$		Break points
						$SupLR_T(2\|1)$	$SupLR_T(3\|2)$	
US interest rate	4	2.02	0.17	26.17***	36.27***	21.94**	10.70	Oct., 1999 Aug., 2007
US inflation rate	3	1.99	0.37	16.63**	27.96***	17.38**	6.73	Sep., 2002 Feb., 2008
China interest rate	2	1.99	0.98	25.95***	35.08***	11.75*	2.43	Jul., 2007 Aug., 2008
China inflation rate	4	1.98	0.79	29.72***	47.65***	26.08***	9.11	Oct., 2007 Jul., 2009
Singapore interest rate	2	1.96	0.89	90.61***	90.61***	5.02	–	Jul., 2008
Singapore inflation rate	3	1.99	0.98	19.86**	19.86**	7.57	–	Jan., 2007

Note: * 10% level, ** 5% level, *** 1% level.

6 Conclusion

We have proposed in this paper a system adaptation framework to identify the dynamics of the stock market, based on the well established systems theory. Treating the stock market as a highly complex system, the proposed framework consists of an internal model and an adaptive filter with focuses on the feedback and force in the market. In the internal model, the output-error linear model is adopted to generate the same trend as the actual price. The external force, which drives the estimated price to approach to the actual price, is obtained from the differences between the actual and internal prices. A time-varying state space model with instrumental variables is proposed to link the external influential factors to this market external force.

We have investigated the DJIA from January 2008 to December 2010, the period right after the global financial crisis happened. The whole period is separated into four subperiods according to the economic situation. This system adaptation framework outperformed the existing methods, such as the commonly used ARMAX approach, reported in the literature in terms of one-step-ahead prediction of stock prices. That shows a great ability of our framework in understanding the dynamics of the stock market. We have applied the framework to forecast the turning periods in the market trend. By analyzing the external force in the frequency domain, major turning periods in the market can be forecasted through the appearing of clear patterns in the related frequency contents. This method has successfully applied in US, China and Singapore markets. All these results reveal that the system based framework is a promising method for modeling financial markets. We are currently investigating interaction behavior between more input influential factors especially the behavior of investors, and building a trading system based on our proposed framework.

References

[1] F. J. Fabozzi, F. G. Modigliani, and F. J. Jones, *Foundations of Financial Markets and Institutions*, 4th edition, Prentice Hall, New Jersey, 2009.

[2] W. F. Sharpe, G. J. Alexander, and J. V. Bailey, *Investments*, Prentice Hall, New Jersey, 1998.

[3] D. M. Hanssens, L. J. Parsons, and R. L. Schultz, *Market Response Models: Econometric and Time Series Analysis*, Kluwer Academic Publishers, Boston, 1990.

[4] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.

[5] J. Y. Campbell and J. Ammer, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance*, 1993, **48**(1): 3–37.

[6] J. Friedman and Y. Shachmurove, Using Vector Autoregression Models to Analyze the Behavior of the European Community Stock Markets, CARESS Working Paper, 1997.

[7] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation, *Econometrica*, 1982, **50**(4): 987–1008.

[8] E. M. J. H. Hol, *Empirical Studies on Volatility in International Stock Markets*, Kluwer Academic Publishers Boston, 2003.

[9] A. N. Refenes, A. Zapranis, and G. Francis, Stock performance modeling using neural networks: A comparative study with regression models, *Neural Networks*, 1994, **7**: 375–388.

[10] H. Q. Yang, Margin Variations in Support Vector Regression for the Stock Market Prediction, Master's thesis, Chinese University of Hong Kong, 2003.

[11] D. I. Michael, *Interaction between Economics and Systems Theory, Optimization Techniques: Lecture Notes in Control and Information Sciences*, Springer, Berlin/Heidelberg, 1980.

[12] T. Poggio, A. W. Lo, B. D. LeBaron, and N. T. Chan, Agent-based models of financial markets: A comparison with experimental markets, MIT Sloan Working Paper, 2001, 4195-01.

[13] S. H. Chen and C. H. Yeh, On the emergent properties of artificial stock markets: The efficient market hypothesis and the rational expectations hypothesis, *Journal of Economic Behavior & Organization*, 2001, **49**(2): 217–239.
[14] S. H. Chen and C. C. Liao, Agent-based computational modeling of the stock price-volume relation, *Information Sciences*, 2005, **170**(1): 75–100.
[15] D. Orrell and P. E. McSharry, System economics: Overcoming the pitfalls of forecasting models via a multidisciplinary approach, *International Journal of Forecasting*, 2009, **25**(4): 734–743.
[16] C. Chiarella and S. Gao, Modelling the value of the S&P 500 – A system dynamics perspective, University of Technology Working Paper, 2002, 115.
[17] X.R. Cao and D. X. Wang, How good is technique analysis? Keynote Speech of the 6th International Conference on Computer Science and Education, 2011.
[18] L. Gerencsér and Z. Mátyás, A behavioral stock market model, *Mathematical Methods of Operations Research*, 2008, **67**: 43–63.
[19] X. Zheng and B. M. Chen, Modeling and analysis of financial markets using system adaptation and frequency domain approach, *Proceedings of the 7th IEEE International Conference on Control and Automation*, Christchurch, New Zealand, 2009.
[20] X. Zheng and B. M. Chen, Identification of market forces in the financial system adaptation framework, *Proceedings of the 8th IEEE International Conference on Control and Automation*, Xiamen, China, 2010.
[21] P. C. Young, Time-variable parameter and trend estimation in non-stationary economic time series, *Journal of Forecasting*, 1994, **13**(2): 179–210.
[22] J. J. Binder and M. J. Merges, Stock market volatility and economic factors, *Review of Quantitative Finance and Accounting*, 2001, **17**: 5–26.
[23] P. C. Young, Stochastic, dynamic modeling and signal processing, time variable and state dependent parameter estimation, in *Nonlinear and Nonstationary Signal Processing* (ed. by W. J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young), Cambridge University Press, Cambridge, 2000.
[24] C. J. Taylor, D. J. Pedregal, P. C. Young, and W. Tych, Environmental time series analysis and forecasting with the Captain toolbox, *Environmental Modelling and Software*, 2007, **22**(6): 797–814.
[25] N. Chen, R. Roll, and S. Ross, Economic forces and the stock market, *Journal of Business*, 1986, **59**: 383–403.
[26] K.H. Kim, Dollar exchange rate and the stock price: Evidence from multivariate cointegration and error correction model, *Review of Financial Economics*, 2003, **12**: 301–313.
[27] M. Baker and J. Wurgler, Investor sentiment in the stock market, *Journal of Economic Perspectives*, 2007, **21**: 129–151.
[28] W. Y. Lee, C. X. Jiang, and D. C. Indro, Stock market volatility, excess returns, and the role of investor sentiment, *Journal of Banking & Finance*, 2002, **26**(12): 2277–2299.
[29] S. P. Kothari, J. Lewellen, and J. B. Warner, Stock returns, aggregate earnings surprises, and behavioral finance, *Journal of Financial Economics*, 2006, **79**(3): 537–568.
[30] G. Kaplanski and H. Levy, Sentiment and stock prices: The case of aviation disasters, *Journal of Financial Economics*, 2010, **95**(2): 174–201.
[31] W. Sun, A comprehensive approach to predicting market bottoms, Bachelor of Engineering Thesis, Department of Electrical & Computer Engineering, National University of Singapore, 2009.
[32] D. L. Thornton, Tests of the market's reaction to federal funds rate target changes, *Federal Reserve Bank of St. Louis Review*, 1998, **1998**: 25–36.
[33] K. N. Kuttner, Monetary policy surprises and interest rates: Evidence from the Fed funds futures market, *Journal of Monetary Economics*, 2001, **47**(3): 523–544.
[34] B. S. Bernanke and K. N. Kuttner, What explains the stock market's reaction to federal reserve policy? *Journal of Finance*, 2005, **60**(3): 1221–1257.
[35] J. F. Geweke, Measurement of linear dependence and feedback between multiple time series, *Journal of the American Statistical Association*, 1982, **77**: 304–324.
[36] W. Hesse, E. Moller, M. Arnold, and B. Schack, The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies, *Journal of Neuroscience Methods*,

2003, **124**: 27–44.

[37] A. Roebroeck, E. Formisano, and R. Goebel, Mapping directed influence over the brain using Granger causality and fMRI, *Neuroimage*, 2005, **25**: 230–242.

[38] L. Ljung, *System Identification: Theory for the User*, Prentice Hall, New Jersey, 1999.

[39] J. Ding, F. Ding, and S. Zhang, Parameter identification of multi-input, single-output systems based on FIR models and least squares principle, *Applied Mathematics and Computation*, 2008, **197**(1): 297–305.

[40] A. C. Harvey, *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, 1989.

[41] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, Convergence properties of the nelder-Mead simplex method in low dimensions, *IAM Journal of Optimization*, 1998, **9**(1): 112–147.

[42] R. G. D. Steel and J. H. Torrie, *Principles and Procedures of Statistics: A Biometrical Approach*, McGraw-Hill, New York, 1960.

[43] P. A. McKay, Oil lifts Dow to 10062.94 in late rally, *Wall Street Journal*, 16 October, 2009, 1.

[44] P. Chen, Random walk or color chaos on the stock market? — Time-Frequency analysis of S&P indexes, *Studies in Nonlinear Dynamics & Econometrics*, 1996, **1**(2): 87–103.

[45] P. Chen, Trends, shocks, persistent cycles in evolving economy: business cycle measurement in Time-Frequency representation, *Nonlinear Dynamics and Economics* (ed. by W. A. Barnett, A. P. Kirman and M. Salmon), Cambridge University Press, Cambridge, 1996.

[46] G. Turhan-Sayan and S. Sayan, Use of Time-Frequency representations in the analysis of stock market data, *Computational Methods in Decision-making, Economics and Finance/Applied Optimization Series* (ed. by E. J. Kontoghiorghes, B. Rustem and S. Siokos), Kluwer Academic Publishers, Dordrecht, 2002.

[47] M. P. Niemira, Forecasting turning points in the stock market cycle and asset allocation implications, *Analyzing Modern Business Cycles: Essays Honoring Geoffrey H. Moore* (ed. by P. A. Klein), M. E. Sharpe, New York, 1990.

[48] S. Kim and F. In, The relationship between stock returns and inflation: New evidence from wavelet analysis, *Journal of Empirical Finance*, 2005, **12**(3): 435–444.

[49] G. Chow, Tests of equality between sets of coefficients in two linear regressions, *Econometrica*, 1960, **28**(3): 591–605.

[50] D. W. K. Andrew and W. Ploberger, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 1994, **62**(6): 1383–1414.

[51] J. Bai and P. Perron, Estimating and testing linear models with multiple structural changes, *Econometrica*, 1998, **66**(1): 47–78.

[52] J. Bai and P. Perron, Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, 2003, **18**(1): 1–22.