IDENTIFICATION OF COUPLED SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS IN FINANCE INCLUDING INVESTOR SENTIMENT BY MULTIVARIATE ADAPTIVE REGRESSION SPLINES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BETÜL KALAYCI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

JULY 2017

Approval of the thesis:

IDENTIFICATION OF COUPLED SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS IN FINANCE INCLUDING INVESTOR SENTIMENT BY MULTIVARIATE ADAPTIVE REGRESSION SPLINES

submitted by **BETÜL KALAYCI** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Financial Mathematics, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics** ______________

Assoc. Prof. Dr. Yeliz Yolcu Okur
Head of Department, **Financial Mathematics** ______________

Prof. Dr. Gerhard-Wilhelm Weber
Supervisor, **Scientific Computing, METU** ______________

Examining Committee Members:

Assoc. Prof. Dr. Ali Devin Sezer
Financial Mathematics, METU ______________

Prof. Dr. Gerhard Wilhelm Weber
Scientific Computing, METU ______________

Assist. Prof. Dr. Özge Sezgin Alp
Accounting and Financial Management, Baskent University ______________

Date: ______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: BETÜL KALAYCI

Signature :

ABSTRACT

IDENTIFICATION OF COUPLED SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS IN FINANCE INCLUDING INVESTOR SENTIMENT BY MULTIVARIATE ADAPTIVE REGRESSION SPLINES

Kalaycı, Betül
M.S., Department of Financial Mathematics
Supervisor : Prof. Dr. Gerhard-Wilhelm Weber

JULY 2017, 66 pages

Stochastic Differential Equations (SDEs) rapidly become the most well-known format in which to express such diverse mathematical models under uncertainty such as financial models, neural systems, micro-economic systems, and human behaviour. They are one of the main methods to describe randomness of a dynamical model today. In a financial system, different kinds of SDEs have been elaborated to model various financial assets. On the other hand, economists have conducted research on several empirical phenomena regarding the behaviour of individual investors, such as how their emotions and opinions influence their decisions. Emotions can affect the way of thinking. A negative state leads to be risk-averse, while a positive state leads to be ambitious, and to act in a risky way even. All those emotions and opinions are described by the word Sentiment. In finance, stochastic changes might occur according to investors' sentiment levels.

In our study, we aim to represent the mutual effects between some financial process and investors' sentiment with constructing a coupled system of non-autonomous SDEs, evolving in time. These equations are hard to assess and solve. Therefore, we express them in a simplified manner of an approximation by discretization and Multivariate Adaptive Regression Splines (MARS) model. MARS is a strong method for flexible regression and classification with interactive variables, based on high-dimensional and big data. Hereby, we treat time as another spatial variable. Afterwards, we will present a modern application with real-world data. The thesis finishes with a conclusion and an outlook towards future studies.

Keywords: SDEs, parameter estimation, economics, investor sentiment, Multivariate Adaptive Regression Splines

ÖZ

FİNANSTA YATIRIMCI DUYARLILIĞINI İÇEREN BAĞLANTILI STOKASTİK DİFERENSİYEL DENKLEM SİSTEMLERİNİN ÇOK DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ TARAFINDAN TANIMLANMASI

Kalaycı, Betül
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Prof. Dr. Gerhard-Wilhelm Weber

Temmuz 2017, 66 sayfa

Stokastik diferansiyel denklemler, finansal modeller, sinir sistemleri, mikro-ekonomik sistemler ve insan davranışı gibi belirsizlik altındaki çeşitli matematiksel modelleri ifade etmede hızlı bir şekilde en iyi bilinen format oldular. Onlar bugün bir dinamik modelin rastgeleliğini tanımlamada temel yöntemlerden birisi. Bir finansal sistemde çeşitli finansal varlıkları modellemek için farklı türlerde stokastik diferansiyel denklemler geliştirilmiştir. Diğer yandan iktisatçılar, bireysel yatırımcıların duygularının ve görüşlerinin kararlarını nasıl etkilediği gibi davranışlarıyla ilgili birçok ampirik olgunun üzerinde araştırma yürütmüşlerdir. Duygular, düşünce biçimini etkileyebilir. Olumsuz bir durum riskten kaçınmaya neden olurken, olumlu bir durum hırslı olmaya ve hatta riskli bir şekilde davranmaya neden olur. Tüm bu duygu ve düşünceler, duyarlılık kelimesi tarafından tanımlanır. Finansta stokastik değişiklikler yatırımcıların duyarlılık seviyelerine göre ortaya çıkabilir.

Çalışmamızda bazı finansal süreçler ile yatırımcıların duyarlılığı arasındaki karşılıklı etkileri, zamanla gelişen ve otonom olmayan bağlantılı bir stokastik denklem sistemi kurarak temsil etmeyi amaçlıyoruz. Bunlar değerlendirilmesi ve çözümü zor denklemler. Bu nedenle bunları ayrıklaştırma ve Çok Değişkenli Uyarlanabilir Regresyon Eğriler (MARS) modeli sayesinde basitleştirilmiş bir yaklaşım şekliyle ifade edeceğiz. MARS, yüksek boyutlu ve büyük verilere dayanan etkileşimli değişkenlerle sınıflandırma ve esnek regresyon için güçlü bir yöntemdir. Burada zamanı bir başka

uzaysal değişken olarak değerlendiriyoruz. Sonrasında gerçek dünya verileriyle modern bir uygulama sunacağız. Tez, gelecek çalışmalara yönelik sonuç ve görüşler ile tamamlanmaktadır.

Anahtar Kelimeler: Stokastik diferansiyel denklemler, parametre tahmini, ekonomi, yatırımcı duyarlılığı, Çok Değişkenli Uyarlanabilir Regresyon Eğrileri

To My Family

ACKNOWLEDGMENTS

I would like to sincerely thank my thesis supervisor, Prof. Dr. Gerhard-Wilhelm Weber. It is a great honor for me to be his MSc. student and also his friend. Prof. Dr. Weber encouraged, guided and supported me during the development and preparation of this thesis from beginning to end. I am very grateful for all of his contributions and motivations during the whole period of my MSc. studies. Without his help, this thesis would not even exist.

I also owe my deepest thanks to Dr. Ayşe Özmen for her share of experiences and expertise. Thanks to her, I have learned many invaluable pieces of information about my studies which I would not be familiar without her at all. I have benefited greatly from many productive discussions with her. I am also thankful for her patience and enlightening visions in all steps of our study. She has become my role model from our first day of studying to the end of my life.

I would like to sincerely thank Dr. Azar Karimov for his help and contributions. When we were stuck in a complicated point during our thesis, he was the saver with his brilliant ideas and support. I am very grateful to him.

I would like to thank my dear jury members; Assoc. Prof. Dr. Ali Devin Sezer and Assist. Prof. Dr. Özge Sezgin Alp for their support and motivation.

I deeply thank all members of the Institute of Applied Mathematics (IAM) at Middle East Technical University for their friendship, understanding and help during my studies.

I am also grateful to my Professors; Assoc. Prof. Dr. Fatma Karakoç and Prof. Dr. Cihan Orhan from Ankara University, Department of Mathematics. Without their encouragement and motivation, I would never have imagined studying in my dream university one day.

Finally, and most significantly, I would like to thank each member of my family for their patience and love:

I owe my deepest thanks to my mother Aynur Kalaycı who has always supported and believed in me at my both best and worst days, from the first day of my life to now.

I owe my grateful thanks to my father Sami Kalaycı who has enlightened my vision, guided me with his invaluable experiences and thoughts about every area of life since I was a little girl.

I also would like to thank my brother Haluk Kalaycı who always shares his suggestions with me and always wishes my best.

This thesis would not have been written without my family's perpetual support, patient and endless love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AAE	Average Absolute Error
AAII	American Association of Individual Investors
AD	Alzheimar's Disease
BF	Basis Function
CART	Classification and Regression Trees
CEED	Closed End Equity Discount
CEFD	Closed End Fund Discount
CFA	Chartered Financial Analyst
CGPLM	Conic Generalized Partial Linear Model
CMARS	Conic Multivariate Adaptive Regression Splines
CPI	Consumer Price Index
DDE	Delay Differential Equations
P^{D-ND}	Dividend Premium
fMRI	funtional Magnetic Resonance Imaging
GCV	Generalized Cross Validation
GMM	Generalized Method of Moments
IDE	Impulsive Differential Equation
IPO	Initial Public Offering
LM	Levenberg-Marquardt
LSE	Least-Squares Estimation
MARS	Multivariate Adaptive Regression Splines
MCI	Mild Cognitive Impairment
MCCI	Michigan's Consumer Confidence Index
MLE	Maximum Likelihood Estimation
NAV	Net Asset Value
NIPO	Number of Initial Public Offering
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
PFC	Pre Frontal Cortex
R^2_{adj}	Adjusted R^2

RCMARS	Robust Conic Multivariate Adaptive Regression Spline
RCGPLM	Robust Conic Conic Generalized Partial Linear Model
RIPO	Returns of Initial Public Offering
RMSE	Root Mean Square Error
RSS	Residual Sum of Squares
SDE	Stochastic Differential Equation
SS-ANOVA	Smoothing Spline of Analysis of Variance
VST	Ventral Striatum
YSMI	Yale School of Management's Stock Market Index

CHAPTER 1

INTRODUCTION

1.1 Introduction

Dynamical processes in nature, economy and, especially, in finance are exposed to random effects and noise. These time-dependent processes are generally characterized by their large number and by a high frequency. Since large dataset from financial observations have discretely discontinuous piecewise constant structures, expressing these high-frequency observations is challenge [31, 80]. Therefore, to cope with these oscillations, a mathematical model has to be established. While composing such a model, high sensivity with respect to oscillations and nonsmoothness of the data should be taken into consideration [75, 80].

Economists barely talk about the “mind”, but they do view mind leading to reasoning. They take into account several empirical facts regarding the behavior of indiviual investors, such as how their emotions, opinions and views influence their decisions; this constitutes the new subject of Behavioral Finance [16, 23]. The general name of all these emotional states is called *Sentiment*. Behavioral Finance investigates the effect of Investor Sentiment in stock markets, in the economy. This is reflected in treating beliefs about the world as expectations of variables’ conditions upon current data or current information [16].

In our study, it has been dealt with real-world dataset which belong to the sentiment of investors and financial processes. Before showing the results and effects of these datasets, we have represented those factors with Financial Mathematics of Stochastic Differential Equations (SDEs). These equations can be about processes of prices, volatility, interest rates, etc., about underlyings and derivatives of varied types as well. As in reality many phenomena are influenced by random noise, behavior of the noise should be reflected in the differential equations. For that reason, stochastic differential equations have to be used with care. They increasingly became a popular representation of the natural processes [61, 75]. These equations play an important role in applications of different kinds of fields; but, in our study, applications in finance and human behavior are focused on.

While representing financial processes and investor sentiment with SDEs, there is an issue which needs to be known. In fact, SDEs are hard to represent and solve because of the unknown parameters which they involve. Therefore, those parameters have to be estimated from observations of the process [80]. However, these observations are hard to apply because of the assumptions that need to be provided. Since we work on real-world datasets from different kinds of sectors as financial processes and developments of human factor, these sectors are mostly described by their large number and by a high fluctuation. Therefore, it should be dealt with those fluctuation and high variations for such observations carefully, by using a parameter estimation method which will decline this high diversification and will provide a smoother approximation of the data [61, 75]. For this reason, we will use Multivariate Adaptive Regression Splines (MARS) method to estimate the components of SDEs on our stochastic processes. As a result, we have identified a system of SDEs in terms of financial processes and Investor Sentiment for MARS by working with real-world data.

1.2 State of the Research

When beginning to prepare this project, our main purpose was to construct a coupled systems with the representation of SDEs in the sectors of finance and neuroscience. By saying "neuro", we want to identify how financial decisions are affected according to human's neurologic acitivities, what is happening in investors' brains while they are making financial decisions; e.g., selling, buying and which hormones are activating when they are in a risky or risk-averse situation. The studies on these topics are explained in a field called "neurofinance" which examines the neurological basis of the mental state in financial decisions [51].

While making financial decisions, some neuroscientists measure people's responses via functional Magnetic Resounance Imaging (fMRI). Previously, Lo and Repin (2002) presented psychophysiological evidence of traders who exhibit an important emotional response, as gauged by elevated levels of "skin conductance and cardiovascular variables" via fMRI, during market events such as price volatiliy or intra-day breaks in trend [39]. Frydman (2015) measured subjects's brain activity via fMRI while they traded stocks in an experimental market [21].

Celik, Weber, Eyuboglu and Oguz (2017) remarked that "Neuroscience is of emerging importance along with the contributions of Operational Research to the practices of diagnosing neurodegenerative diseases with computer-aided systems based on brain image analysis". More detailed information about the studies in Neuroscience is given in Chapter 3.

After researching for several months and communicating with scientists and practitioners who work in this area, we have noticed that to find suitable neurologic data was nearly impossible in general, which means that we cannot show our investigations of "neuro" on the part without the data. Afterwards, we decided to examine Sentiment which is included in Behavioral Finance. In this manner, our study converted into identification and representation of coupled systems of SDEs, but with both finance and Investor Sentiment. Instead of conducting research about people's brain and hormonal activities according to their behaviors, we use data which show how people act while they are making financial decisions, and we assess them based on psychological procedures coupled with financial data [69].

Pioneering and popular studies about sentiment and "aggregate stock returns" appeared in the 1980s. The state of sentiment was not explicit, and evidence from the statistical point of view was not so strong in these studies [8]. To deal with Investor Sentiment, two branches of literature showed up: the first is given by articles using market variables as proxies for Investor Sentiment, and the second one consists of publications using Investor Sentiment surveys [12].

As an example of the first kind of articles, Fisher and Klimek (1999) studied the effect of sentiment in different groups of people. They stated that the sentiment of Wall Street strategic experts is not related to the sentiment of "individual investors", or to the sentiment of "newsletter writers", even though the sentiments of some groups of participants are nearly related. Furthermore, they stated that the sentiment of each of the groups are negatively associated with the future stock returns [18].

Brown and Cliff (2004) constructed a sentiment index based on twelve measures and tested its correlation with stock returns. It is shown by them that "the sentiment levels and shifts are strongly correlated with contemporal market returns". But sentiment cannot predict future stock returns for short-term horizons [77].

Baker and Wurgler (2006) constructed a sentiment index based on six proxies. These proxies are "trading volume as measured by NYSE turnover, the dividend premium, the closed-end fund discount; the number and first-day returns on IPOs, and the equity share in new issues". Baker and Wurgler (2007) suggested that main factors for some stocks which are more prone to wide changes are due to the subjectivity of their valuations. They say that "a wave of investor sentiment has larger effects on securities whose values are highly subjective and difficult to arbitrage" [8].

Zouaoui, Nouyrigat and Beer (2011) expressed that "we examine the influence of investor sentiment on the probability of stock market crises. We find that investor sentiment increases the probability of occurrence of stock market crises within a one-year horizon" [82].

Bathia and Bredin examined if investor sentiment has an important effect on both value and growth of stock returns and aggregate market returns, by using "a range of investor sentiment proxies, including investor survey, equity fund flow, closed-end equity fund (CEEF) discount and equity put-call ratio" [10].

Uygur and Tas (2014) found an important evidence on the fact that a change in Investor Sentiment has a higher impact on conditional volatility of industry, banking, and beverages and food sector indexes when compared to sectors such as retail or telecommunication. After checking for macroeconomic shocks, the weekly trading volume of Istanbul Stock Exchange 100 is employed as investor sentiment proxy [70].

Huang, Jiang, Tu and Zhou (2014) proposed a "new investor sentiment index that is aligned with the purpose of predicting the aggregate stock market. By eliminating a common noise component in sentiment proxies, the new index has much greater predictive power than existing sentiment indices have both in and out of sample, and the predictability becomes both statistically and economically significant" [27].

Gao, Ren and Zhang (2016) studied how the investor sentiment influences stock markets worldwide. According to Google search behavior of households', the authors built up a weekly search based on the sentiment for 40 countries during the 2004–2014 period and, at the end, they formed an significant role of global sentiment in driving sentiment and anticipating returns across countries [24].

On the other hand, Investor Sentiment surveys being also used in articles as a measure of investor sentiment are: CFA (Chartered Financial Analyst) Institute's Global Market Sentiment Survey, Barron's Investor Sentiment Readings, Franklin Templeton Global Investor Sentiment Survey, and the American Association of Individual Investors (AAII) weekly poll of Investor Sentiment.

Chang, Yu, Reinstein and Churyk (2016) recommended a decision making process, consisting of four steps: First, an individual investor should choose an investor sentiment index with the highest predictability from an available resource such as AAII, CFA or Barron's sentiment survey conducted. Second, he should describe and establish the highest level based on bullish and bearish sentiment. Third, he foresees the market direction and forms expectations for the future. In the fourth step, he can make an investment decision [77].

Hengelbreck, Theissen and Westheide (2010) analyzed whether data can identify immediate market reaction, and whether "the sign of such a reaction corresponds to the sign of the intermediate and long-term predictive ability". To examine these, the authors used Investor Sentiment indicators from the Sentix sentiment index and AAII index [26]. Bormann (2013) clarified "market sentiment" by using psychological definitions, and used data from the German sentiment index "Sentix" which indicated the usefulness of these definitions [12].

From these studies, we see that there has been a lot of research done about investor sentiment via using different resources. In our thesis, we employ a sentiment index consisting of proxies, which is constructed by Baker and Wurger in 2006 [6]. It will be addressed in the following chapters.

Since we denote a coupled system of SDEs with financial processes and investor sentiment, the parameters of these SDEs need to be estimated. For the parameter estimation of SDEs, some approaches have been developed. These different kinds of approaches commonly have been studied in several papers.

Gonzales (2012) explained that he aims to estimate the parameters, because the process cannot be observed exactly. For different datasets there will be different values of the parameters, with these resulting parameters addressing different markets, but always employing the basically same kind of structure of SDEs. As long as the author observed the process, he acquired the estimates more authentically [25].

Rahman, Bahar, Bazli, Rosli, Saleh and Weber (2009) provided a review of literature in estimating parameters of SDEs. They estimated the parameters of the logistic equation by applying the optimization method of nonlinear least-squares Levenberg-Marguardt Method (LM method) [47].

Teka (2013) studied parameter estimation of Black-Scholes-Merton model which models asset prices for European options and discussed an estimation procedures. To estimate their parameters, drift and diffusion, the author used the method of Maximum Likelihood Estimation (MLE) [64].

Scheider, Craigmile and Herbei (2014) obtained a sequential sample of parameter values by approximating with MLE method in SDE models by a kriging-based optimization strategy [55].

Cicha (2010) estimated parameters of an SDE with cubic drift and constant diffusion by minimizing the difference between parametric marginal distribution and a nonparametric estimate of the marginal distribution [14].

As it is seen from those various studies, there are several methods that can be used to estimate parameters, some of them are: Exact Maximum Likelihood, Pseudo-likelihood methods (discretising the SDE), Approximated Likelihood methods (approximating the likelihood function, e.g., with Hermite polynomials), Generalized Method of Moments (GMM) (matching theoretical and sample moments), Martingale Estimating functions, and (Adapted) Kalman Filter method [25]. Classical methods such as Maximum Likelihood Estimation (MLE), Methods of Moment and Least-Squares Estimation (LSE) have been widely used in the identification of Ordinary Differential Equation (ODE) parameters.

Necessities of these classical methods, such as MLE applied for an SDE, are that the approximation includes an intense computation, inadequate accuracy and difficulty for applying them on a multivariate SDE. Timmer (2008) noted that for the superior statistical properties, MLE seems to be the most preferred method. However, he argued it is extremely expensive and not feasible. Then, the author discussed the relation between MLE and quasi MLE, and suggested quasi MLE [68]. Furthermore, some drawbacks of classical methods such as MLE consist of the decline of the estimators' sufficiency because of the computational problem which emerges from the extensive needed investigation of local optima and the high computation time (Brunel 2008) [48]. Thus, Rahman, Bahar, Rosli and Salleh (2012) stated that Varziri et al. (2008) proposed a "new version of a two-step method via the minimization of the negative of natural logarithm of approximate probability density function to estimate the drift and the spline parameters of the SDE". They purposed to estimate the parameters of SDEs, with a Bayesian regression spline [48].

As it is mentioned above, some serious problems may occur in classical methods for parameter estimation. Besides, these methods are hard to apply because of the assumptions that have to be fulfilled. Therefore, in order to deal with this issue, Multivariate Adaptive Regression Splines (MARS) method has been preferred in order to identify SDEs with their parameters in a simplified manner of approximation. Similarly to the previously explained research, there is a number of studies which also use, regression splines but with the difference of not applying Bayesian regression. MARS has been used for various applications; it includes a forward stepwise regression. MARS is an important tool in classification and regression. It is very helpful in many areas of finance, neuroscience, technology and nature.

The MARS algorithm consists of two sub-algorithms. Taylan, Weber and Yerlikaya (2008) proposed not to use the second one (backward stepwise algorithm) in their study, but they constructed a penalized residual sum of squares for receiving a Tikhonov regularization problem, for which they imply continuous optimization [63]. In another study, Taylan and Weber (2008) approximated SDEs by additive models which are based on splines and a discretization. Their parameter estimation method addresses linearly contained spline coefficients as prepared in their previous study (Taylan and Weber, 2007) and partially nonlinearly contained probabilistic parameters. They constructed a "penalized residual sum of squares model and face parametric nonlinearities by Gauss-Newton's and Levenberg-Marquardt's method for determining the iteration steps" [74].

As we mentioned previously, Celik, Weber, Eyuboglu and Oguz (2017) introduced a voxel-based procedure, by "Voxel-MARS" in order to detect the Alzheimer's disease (AD) and mild cognitive impairment (MCI) via the use of MARS. This method was employed as a classifier in the area of brain MRI analysis for the first time [13].

In fact, MARS is one of the nonparametric approaches that makes no specific assumptions to estimate the components of SDEs on some given stochastic process.

In addition to the various theories, hypotheses and results of neural, behavioral and sentiment-based investigations, which we have presented above and also in Chapter 3, in our research project, we focus on describing human behavior with SDEs. We express the effect of human behavior and any financial processes onto each other according to time. As a result, we obtain a coupled system of SDEs.

The main objective of our study is to propose MARS as an identification method for a system of SDEs and to elaborate this rather new methodology on quite different kinds of coupled stochastic dynamics. These are financial processes and human factor processes, without any further assumptions. Herewith a special contribution of the thesis is a mathematical representation of the "human factor". Of course, we mean that with all the needed care. This thesis is a pioneering work on a quite "hard" subject, and it invites the scientific community to future research.

1.3 Scope of the Thesis

This thesis consists of six chapters. The structure of the thesis is as follows:

- Chapter 1: It is the introduction of the thesis. The content of the study is surveyed in this chapter.
- Chapter 2: Background information is provided here about stochastic SDEs (one-dimensional and multi-dimensional), regression, linear model, nonlinear model, nonparametric model and Multivariate Adaptive Regression Splines (MARS).

- Chapter 3: Explanations of Neurofinance and Investor Sentiment, which is a field of Behavioral Finance, is provided here.
- Chapter 4: Representation of coupled systems of SDEs in terms of finance and human behavior, and estimation of these differential equations with MARS method are detailed in this part of the thesis.
- Chapter 5: Here we present our application on the coupled financial process and human factor process with two real-world datasets by MARS.
- Chapter 6: Conclusion and outlook of the thesis are elaborated in this last chapter.

CHAPTER 2

MATHEMATICAL AND STATISTICAL BACKGROUND

2.1 Introduction

In this chapter, we begin by mentioning about the construction of an SDE, how and by whom it was first introduced. Subsequently, the general form and parts of the SDE are identified. Afterwards, the properties of random noise part are explained which plays a major role in constituting the SDE. Then, our MARS (Multivariate Adaptive Regression Splines) method is introduced. However, in order to comprehend MARS sufficiently, priorly, it is explained what regression is and, in contrast to MARS, what simple regression models are.

2.2 Stochastic Differential Equations

Many phenomena in natural, technological and economical processes can be better described and analyzed by the help of mathematical modeling, statistical learning, machine learning or data mining. A wide range of literature on mathematical modeling of dynamical processes can be found for deterministic differential equations, e.g., ordinary differential equations (ODEs), partial differential equations (PDEs), impulsive differential equations (IDEs), or delay differential equations (DDEs), whereby the element of noise is not considered. Since, however, in reality, many phenomena are influenced by random fluctuation, the behavior of noise in differential equations should be explained. For that reason, *Stochastic Differential Equations* (SDEs) are required [47]. There is a broad field of applications where SDEs are given, for example, in finance, economics, physics, population dynamics, hydrology, biometry, engineering and social sciences.

SDEs have become an increasingly common tool for the representation of natural, technological and societal processes in many fields [61]. These equations are useful to model most financial assets, such as stock price or interest rate processes [36]. Methods for the computational solution of SDEs are based on similar procedures as for ODEs, but extended to provide support for implying stochastic dynamics [53]. More generally, these equations are acquired by allowing randomness in the coefficients of a differential equation [45].

In various fields, dynamical processes are modelled by means of a deterministic differential equation with initial value $x_0 \in \mathbb{R}$:

$$\dot{x} := \frac{dx}{dt} = a(x,t), \qquad x(0) = x_0. \tag{2.1}$$

Without loss of generality, we assume the time domain of the solution to be $[0, \infty)$, where $x(\cdot) : [0, \infty) \to \mathbb{R}$ is a solution trajectory satisfying the initial condition $x(0) = x_0$, and $a : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ is a smooth function. In general, x is not known explicitly, but it can be approximated in terms of its time derivative by a.

Likewise for numerous applications in physics, biology, social sciences, engineering, in finance, processes are exposed to some random environmental effects. However, this type of modelling as in Equation (2.1) omits stochastic oscillations and is not suitable for financial assets, e.g., stock prices. In fact, random effects cannot be represented only with an ODE. By allowing such a randomness in the coefficients of differential equations, we become able to describe the behaviour of noise [61].

For the one-dimensional case, the noise term can be represented as ξ_t, so that, with initial value $X(0) = x_0$,

$$\dot{X} = a(X,t) + b(X,t)\xi_t, \qquad t \in [0, \infty). \tag{2.2}$$

Here, a is sometimes named as the deterministic part, $b\xi_t$ is the stochastic part, and the noise ξ_t indicates a generalized stochastic process, also called as *white noise*. There is an example which is commonly used in models from chemistry, physics, economy and finance, that is the *Wiener process W*. This process represents random influences and models the motion of stock prices, etc., which immediately answers to the numerous upcoming kinds of information. Otherwise, in the absence of random fluctuation, we have a deterministic system, that is not sufficient to model such fluctuations for a variety of reasons [53, 61].

A Wiener process, named after Norbert Wiener, is a mathematical construct that forms random behavior observed by the botanist Robert Brown in 1987, widely called a *Brownian motion* [53]. He studied the tiny and apparently random motion of pollen grains suspended on the surface of water [40]. Brownian motion is significant in the modeling of stochastic processes, because it describes the integral of white noise, which is idealized noise that is independent of frequency [53]. Indeed, a French mathematician, Louis Bachelier, discovered the Brownian motion in his PhD. thesis "The Theory of Speculation", which was published in 1900. This study introduced *Mathematical Finance* to the world and provided a source of problems and applications for probability theory and stochastic analysis [5].

The first mathematical characterization of a Brownian motion is due to Albert Einstein [40]. His purpose was to prove the existence of atoms in theory as a tiny particle which collides with many individual atoms or molecules and should be regarded as exposed to a Brownian Motion. After a short while, he discovered the experimental work of Robert Brown. Therefore, the history of SDEs can be considered to have begun from the classic paper of Einstein in 1905, which is named as *"On the Motion of Small Particles Suspended in a Stationary Liquid, as Required by the Molecular-Kinetic Theory of Heat"*. Here, he presented a mathematical connection between a microscopic random motion of particles and a macroscopic diffusion equation [40, 52].

Following Einstein's explanation of an observed Brownian motion, Paul Langevin and others attempted to formulate the dynamics of such a motion in terms of differential equations. The resultant equations were represented in the form of Equation (2.2) [33]. In addition to these studies, Japanese mathematician Kiyoshi Itô also introduced a differential calculus based on a stochastic integral which is called the *Itô calculus* [36].

A one-dimensional Wiener process (or a Brownian motion) is a time-continuous process performing the subsequent features:

- For each t, the random variable W_t is normally distributed with mean 0 and variance t.
- It has independent increments: giving any finitely many times $0 < t_1 < t_2 < \ldots < t_N$, the random variables (increments) $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_N} - W_{t_{N-1}}$ are independent.
- It has stationary increments: $W_t - W_s \sim N(0, t-s)$ for all $0 \leq s < t \leq T$.

Resulting from the first and third properties, for all $0 \leq t \leq T$ the random variable W_t is normally distributed with mean $E(W_t) = 0$ and variance $\text{Var}(W_t) = E(W_t^2) = t$.

A main problem is that a Wiener process (W_t) is not differentiable nearly anywhere. But (W_t) is treated as if it was differentiable, for approximating and smoothing the model. Then, by smoothing the white noise, ξ_t is defined as $\xi_t = \dot{W}_t = dW_t/dt$; herewith, we get Wiener process. Finally, if $\xi_t d_t$ is replaced by dW_t in Equation (2.2), the formula that we obtained is a *Stochastic Differential Equation* [53, 61].
We can rewrite the equation as:

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t; \tag{2.3}$$

here, the "deterministic" or drift term is $a(X_t, t)$ and the "stochastic" or diffusion term is $b(X_t, t)$, respectively. Furthermore, (X_t) is a solution which we will try to determine based on emprical data. Equation (2.3) is called an Itô SDE.

2.3 Multi-dimensional Stochastic Differential Equations

If a finite number of SDEs is given in a model, then the multi-dimensional case should be considered. This situation can arise for the modeling of, e.g., the price evolution of multiple stocks, interest rates, volatilities or variances, wealths, etc. [36, 45].

For the case of random fluctuation in higher dimensions, let $\boldsymbol{W}_t = (W_t^1, W_t^2, ..., W_t^m)^T$ denote an m-dimensional Brownian motion (or a Wiener process) at time t.

Here, the drift part given by $\boldsymbol{a} : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$, namely, $\boldsymbol{a}(X_t, t)$, is a measurable vector process, and the diffusion part expressed through $\boldsymbol{b} : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^{d\times m}$, namely, $\boldsymbol{b}(X_t, t)$, is a measurable matrix-valued process.

For the underlying probability space $(\Omega, \mathcal{F}, P)$, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, a d-dimensional stochastic process $\boldsymbol{X} = (\boldsymbol{X}_t : t \in [0, \infty))$, also denoted as $(\boldsymbol{X}_t)_{t\geq 0}$, is represented by d differential equations, a d-dimensional initial vector and suitable conditions [75].

If we write a coupled system of SDEs with d processes of states and m-dimensional Brownian motions, we obtain [80, 75]:

$$\begin{aligned}
dX_t^1 &= a^1(\boldsymbol{X}_t, t)dt + b^{11}(\boldsymbol{X}_t, t)dW_t^1 + ... + b^{1m}(\boldsymbol{X}_t, t)dW_t^m, \\
dX_t^2 &= a^2(\boldsymbol{X}_t, t)dt + b^{21}(\boldsymbol{X}_t, t)dW_t^1 + ... + b^{2m}(\boldsymbol{X}_t, t)dW_t^m, \\
&\vdots \\
dX_t^d &= a^d(\boldsymbol{X}_t, t)dt + b^{d1}(\boldsymbol{X}_t, t)dW_t^1 + ... + b^{dm}(\boldsymbol{X}_t, t)dW_t^m.
\end{aligned} \tag{2.4}$$

Here, we refer to

$$\boldsymbol{X}_t = \begin{bmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^d \end{bmatrix}, \quad \boldsymbol{a}(\boldsymbol{X}_t, t) = \begin{bmatrix} a^1(\boldsymbol{X}_t, t) \\ a^2(\boldsymbol{X}_t, t) \\ \vdots \\ a^d(\boldsymbol{X}_t, t) \end{bmatrix}, \quad d\boldsymbol{W}_t = \begin{bmatrix} dW_t^1 \\ dW_t^2 \\ \vdots \\ dW_t^m \end{bmatrix},$$

$$\boldsymbol{b}(\boldsymbol{X}_t, t) = \begin{bmatrix} b^{11}(\boldsymbol{X}_t, t) & b^{12}(\boldsymbol{X}_t, t) & \cdots & b^{1m}(\boldsymbol{X}_t, t) \\ b^{21}(\boldsymbol{X}_t, t) & b^{22}(\boldsymbol{X}_t, t) & \cdots & b^{2m}(\boldsymbol{X}_t, t) \\ \vdots & \vdots & \ddots & \vdots \\ b^{d1}(\boldsymbol{X}_t, t) & b^{d2}(\boldsymbol{X}_t, t) & \cdots & b^{dm}(\boldsymbol{X}_t, t) \end{bmatrix}.$$

2.4 Regression

Regression analysis is a mathematical and statistical tool for the investigation of relationships between a response variable and one or more predictor variables. It is largely used for prediction and estimation and, most generally, it estimates the conditional expectation of a response variable, given some predictor variables.

There are numerous applications of regression. These applications occur in almost every area, including the physical and chemical sciences, engineering, economics, management, earth, biological and ecological sciences, and the social sciences. In fact, regression analysis has offered a possibility of usage that has become the most broadly used statistical technique [42].

We will mention about some regression models briefly in this chapter, namely: linear regression model, nonlinear regression model, non-parametric model, additive model and MARS.

2.5 Linear Regression Models

The case of simple linear regression considers a single independent X and a dependent variable Y. The correct relationship between Y and X is a straight line and that the observation Y at each level of X is a random variable [43]. Then, the equation of a straight line regarding these two variables is

$$Y = \beta_0 + \beta_1 X, \tag{2.5}$$

where β_0 is the intercept and β_1 is the slope. The difference between the observed value of Y and the straight line $(\beta_0 + \beta_1 X)$ be *noise* ϵ. The term ϵ can be considered as a statistical error, which is a random variable that explains the defect of the model to fit the data completely. This error may be consisting of the impacts of other variables such as measurement errors. Therefore, a more conceivable model is

$$Y = \beta_0 + \beta_1 X + \epsilon. \tag{2.6}$$

Equation (2.6) is called a *linear regression model*. Here, as usual, X is called the independent (predictor or regressor) variable and Y is called the dependent (response) variable [41]. Since Equation (2.6) includes just one predictor variable, it is called a *simple linear regression model* [41]. For an extra comprehension into the linear regression model, we can fix the value of the predictor variable X and observe the related value of the response Y. If X is fixed by a value x, the random component ϵ on the right-hand side of Equation (2.6) designates the features of Y.

The linear model presumes that the regression function $E(Y|X)$ is linear. It has been also assumed by the Hastie et al. that the aberrations of Y around its expectation are Gaussian and additive [20]. Thus

$$E(Y|X = x) = E(\beta_0 + \beta_1 x + \epsilon) = \beta_0 + \beta_1 x. \tag{2.7}$$

The variance of Y given any value of x of X is

$$Var(Y|X = x) = Var(\beta_0 + \beta_1 x + \epsilon) = \sigma^2, \tag{2.8}$$

where the ϵ is a Gaussian random variable with mean 0 and variance σ^2, written $\epsilon \sim N(0, \sigma^2)$ [20].

Although the mean of the Y is a linear function of x, the variance of Y does not depend on the value of X. Besides, since the errors are not correlated with each other, the responses are also not correlated [20, 41].

The parameters β_0 and β_1 which are called *regression coefficients* are unknown and must be estimated. *Maximum-Likelihood Estimation* and *Least-Squares Estimation* can be used to estimate unknown parameters. The easiest of them is the Least-Squares estimation method (LS). That is, the coefficients β_0 and β_1 are estimated in order that the sum of the squares of the differences between N observations y_i and the straight line along the positions x_i is at a minimum [41, 81], where

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad i = 1, 2, ..., N. \tag{2.9}$$

In fact, this has been written in terms of the N pairs of data (x_i, y_i) $(i = 1, 2, \ldots, N)$. The LS estimation method which addresses the sum of the squares of the aberrations of the observations from the model regression line is given by the objective function

$$S(\beta_0, \beta_1) = \sum_{i=1}^{N} \epsilon^2 = \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_i)^2. \tag{2.10}$$

Equation (2.10), makes no assumptions about the feasibility of model (2.9) [20].

The *LS estimators* of the intercept β_0 and the slope β_1 which minimize $S(\beta_0, \beta_1)$, are denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$, and the prediction equation or the fitted simple regression model is stated as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x. \tag{2.11}$$

These LS estimators β_0 and β_1 must satisfy that the partial derivatives $\partial S/\partial\beta_0$ and $\partial S/\partial\beta_1$ both are equal to zero.

The difference between the observed value y_i and the related fitted value $\hat{y}_i$ is a *residual*. The ith residual is

$$e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i), \quad i = 1, 2, ..., N. \tag{2.12}$$

Residuals play an significant role for searching the adequancy of the model and constitute the error in the fit of the model in comparison to the ith observation y_i.

There are many applications in regression analysis where the dependent variable Y may be related to k predictors, $X_1, X_2, \ldots, X_k$, i.e., the model ought to include more than one regressor, as represented in following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \epsilon. \tag{2.13}$$

Equation (2.13) is called a *multiple linear regression model*. Again, the method of LS estimation may be used to estimate the unknown parameters in this regression model of Equation (2.13).

In general form, the linear regression model can be written as

$$\begin{aligned} Y &= \boldsymbol{X}^T\boldsymbol{\beta} + \epsilon, \\ &= f(\boldsymbol{X}, \boldsymbol{\beta}) + \epsilon, \end{aligned} \tag{2.14}$$

where $\boldsymbol{X} = (1, X_1, X_2, \ldots, X_k)^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^T$ and our linear model function f as defined in Equation (2.14). Since the expected value of the model errors is zero, the expected value of the dependent variable is

$$\begin{aligned} E(Y|\boldsymbol{X} = \boldsymbol{x}) &= E[f(\boldsymbol{x}, \boldsymbol{\beta}) + \epsilon], \\ &= f(\boldsymbol{x}, \boldsymbol{\beta}). \end{aligned} \tag{2.15}$$

Here, $f(\boldsymbol{x}, \boldsymbol{\beta})$ is the *expectation function*, which is a linear mapping of the unknown parameters [41].

2.6 Nonlinear Regression Models

Although linear regression models sometimes satisfy a rich and flexible structure that fits the needs of many analysts, they are not coherent for all situations. Various problems occur in applications of engineering and the sciences where the dependent variable and the independent variables are associated with a *nonlinear* function, e.g., exponential function, logarithmic function, trigonometric function, Gaussian function, or combination of such functions. This brings about *nonlinear regression models* [41].

Any model that contains at least one nonlinear parameter is a *nonlinear regression model*, which means that in a nonlinear model at least one derivative according to a parameter must contain that parameter [81]. We may write some examples for a nonlinear regression model:

$$\begin{aligned} &Y = \theta_1 e^{\theta_2 X} + \epsilon, \\ &Y = \theta_1 X + e^{-\theta_2 X} + \epsilon. \end{aligned} \tag{2.16}$$

The symbol θ is used to stand for a parameter in a nonlinear model to put emphasis on the difference between the linear and the nonlinear case. Then, the general form of a multivariate nonlinear regression model is:

$$Y = f(\boldsymbol{X}, \boldsymbol{\theta}) + \epsilon, \tag{2.17}$$

Similarly to the linear regression case, $f(\boldsymbol{x}, \boldsymbol{\theta})$ represents the *expectation function* for the nonlinear regression model [3, 41, 81]:

$$\begin{aligned} E(Y|\boldsymbol{X} = \boldsymbol{x}) &= E[f(\boldsymbol{x}, \boldsymbol{\theta}) + \epsilon] \\ &= f(\boldsymbol{x}, \boldsymbol{\theta}). \end{aligned} \tag{2.18}$$

For N observations, the model has the following form:

$$y_i = f(\boldsymbol{x}_i, \boldsymbol{\theta}) + \epsilon_i \quad (i = 1, 2, \ldots, N). \tag{2.19}$$

Here, y_i is the ith observation of the response variable, $\boldsymbol{x}_i$ is the ith input data; $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N)^T$ are parameters [81].

The residual sum of squares is given by [81]:

$$S(\boldsymbol{\theta}) = \sum_{i=1}^{N} {\epsilon_i}^2 = \sum_{i=1}^{N} (y_i - f(\boldsymbol{x}_i, \boldsymbol{\theta}))^2, \tag{2.20}$$

where, in terms of the underlying random variables (measurements), the noisy error terms ϵ_i are independent and follow a $N(0, \sigma^2)$ distribution [1], and $\boldsymbol{\theta}$ is the vector of parameters to be determined [3, 41, 81].

2.7 Nonparametric Regression Models

Linear correlation between the response variable and the covariates is a strong assumption. However, it may not be surely valid for each model in real life, that is, there are models where the linearity assumption is not satisfied. If it is insisted on fitting models with a linear method, then a model will not be specified accurately and the prediction will be deceptive. Therefore, a nonparametric model would be suitable with respect to these situations. In general, a *nonparametric model* takes the form [72]

$$Y_i = g(X_{i1}, ..., X_{ip}) + \epsilon_i, \qquad i = 1, ..., N, \tag{2.21}$$

where Y_i is the ith observations, N is the number of observations, and $g(\cdot)$ is a non-specified function. Generally, it is assumed that $g(\cdot)$ is smooth and, in particular, continuous. If $g(X_{i1}, ..., X_{ip}) = \beta_1 X_{i1} + ... + \beta_p X_{ip}$, then a nonparametric model becomes a linear regression model. While in nonparametric regression the function of g is to be determined, in parametric regression, the model parameters are searched [79, 72].

In nonparametric modeling, how to estimate the function $g(\cdot)$ plays an important role. Kernel estimations, regression splines and smoothing splines are widely known estimation methods [72].

Kernel estimation methods use linear estimators to predict the value at a particular point $\boldsymbol{x}$. One of the well-known linear estimator is *Nadaraya-Watson estimator*, which can be expressed as

$$\hat{g}_h(\boldsymbol{x}) = \frac{\sum_{i=1}^{N} K_h(\boldsymbol{x}_i - \boldsymbol{x}) y_i}{\sum_{i=1}^{N} K_h(\boldsymbol{x}_i - \boldsymbol{x})}, \tag{2.22}$$

[1] As we assumed for linear regression, Aster, Borchers and Thurber (2005) also for nonlinear regression permit possibly different Normal distributed measurement errors ϵ_i, with mean 0 and variances σ_i^2. In a setting of Maximum Likelihood Estimation, the authors arrive at minimization problems of least squares (in short: L_2-regression), including a weighting by the standard deviations σ_i. We note that similar arguments apply for Exponential distribution and L_1- rather than L_2-regression.

where K is a kernel function and h is an interval length which is a smoothing parameter that checks the size of the local neighborhood. Gaussian kernel and symmetric Beta family kernel are some well-known kernel functions [72, 76].

Regression splines and smoothing splines, which are other ways to estimate a nonparametric function, are drived from a different point of view than kernels. Both with regression splines and with smoothing splines, one constructs an estimation entirely from a family of selected basis functions. These basis functions are called *base splines* [67, 72].

As a second nonparametric model, *regression spline* estimation with a set of basis functions is represented as follows:

$$\hat{g}(x) = a_0 + a_1 x + a_2 x^2 + ... + a_r x^r + \sum_{j=1}^{m} A_j (x - k_j)_+^r, \qquad (2.23)$$

where r is the order of the regression spline, k_j is the jth knot, $a_0, ..., a_r, A_1, ..., A_m$, are the unknown coefficients, and $(x - k_j)_+$ is a piecewise linear extension and equals to $x - k_j$ if $x > k_j$, and 0 otherwise. The representation of regression splines such as in Equation (2.23) is a linear combination of two parts, namely, $(x^j)_{j=0,...,r}$ and $((x - k_j)_+^r)_{j=1,...,m}$ [72, 76].

Some other known examples for basis functions are B-spline basis, natural splines, and radial basis functions [72].

The third kind of nonparametric model is the smoothing spline analysis of variance (SS-ANOVA) which was suggested by Wahba (1990). The regression function in SS-ANOVA can be written as

$$g(\boldsymbol{x}) = a + \sum_{j=1}^{p} g_j(x_j) + \sum_{1 \leq j < k \leq p} g_{jk}(x_j, x_k) + ... + g_{1...p}(x_1, ..., x_p), \qquad (2.24)$$

where a is a constant, g_j are main effects, and g_{jk} are two-way interactions.

There are also some other nonparametric regression models such as Classification and Regression Trees (CART) and projection-pursuit regression [81]. In our study, we will concentrate on another regression-spline-based method offered by Friedman (1991), which is also nonparametric and called Multivariate Adaptive Regression Splines (MARS).

2.8 Classical Additive Models

In real life, traditional linear models are often not successful enough, because many influences are generally nonlinear. In order to describe these influences, responsive statistical methods such as *nonparametric regression* can be used [19, 62]. Even though nonparametric regression is helpful, we can experience some difficulties, such as, if the number of regressor variables is large in the models, many forms of nonparametric regression will not be fitting sufficiently. There is also difficulty about assessing nonparametric regression which relies on multivariate smooth spline estimates. In order to cope with these difficulties, one of the very important regression models is used, called *Additive Models*. They are used to estimate an additive approximation of the multivariate regression function [62].

For N observations on a dependent variable Y, recorded via the vector $\boldsymbol{y} = (y_1, y_2, ..., y_N)^T$, which is measured at N design or input data vectors $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{im})^T$ $(i = 1, 2, ..., N)$, the *additive model* is defined as follows [74]:

$$Y = \alpha_0 + \sum_{j=1}^{m} f_j(X_j) + \epsilon, \tag{2.25}$$

the error ϵ has mean zero, being independent of the factors X_j [20, 61]. Generally, the functions f_j are considered to be splines, which are auxiliary unknown and univariate functions, whereas, in general, they can be multivariate as, e.g., in MARS. Their estimates are denoted by $\hat{f}_j$. A special assumption on X_j says that $E(f(X_j)) = 0$, since otherwise there will be a free constant in any of the functions. In fact, the intercept (bias) α_0 summarizes all those constants, and each function is estimated by an algorithm offered by Friedman and Stuetzle, which is called *backfitting (or Gauss-Seidel) algorithm*. The estimation function $\hat{f}_j$ then reveals possible nonlinearities in the impact of X_j. All the functions $\hat{f}_j$ do not have to be nonlinear, but they can be mixed in linear and other parametric forms which include nonlinear terms. These terms are defined in two or more variables, or separately for each level of the X_k variables [20, 62].

The mean of the response variable is used for estimating the coefficient α_0: $\hat{\alpha}_0 = E(Y)$. The method can be represented as follows, depending on the partial residual against X_j [74]:

$$r_j = Y - \hat{\alpha}_0 - \sum_{k \neq j} \hat{f}_k(X_k), \tag{2.26}$$

and comprises of estimating each smooth function f_j with respect to the rest of the functions; then, $E(r_j|X_j) = \hat{f}_j(X_j)$, which minimizes $E(Y - \hat{\alpha}_0 - \sum_{j=1}^{m} \hat{f}_j(X_j))^2$ [74].

2.9 Multivariate Adaptive Regression Splines (MARS)

MARS can be seen as a form of stepwise linear regression, according to its introduction by Jerome Friedman (1991). MARS is an adaptive method for regression [20]. Differently from additive models, MARS does not treat the input variables separately, and MARS takes into account their interactions.

MARS is described as an approach of using smoothing splines to fit the relationship between a finite number of regressors and a response variable. In a piecewise sense, it yields a very smooth line, or surface, that can capture "shifts" in the relationship between these variables. These shifts occur at positions assigned as "knots" and ensure a smooth transition between "regimes". MARS algorithm not only investigates over all possible knot positions, but also across all variables and all mutually affected positions among them. This is carried out via the use of integrations of variables named as *basis functions*, called again as splines. When the optimal numbers of basis functions and knot positions are designated by MARS, estimates of the fitted value via the chosen basis functions are provided with a final least-squares regression [56]. As a consequence, the resulting multivariate additive model is determined with a two-stage process, which are *forward stage* and *backward stage*.

At the *forward stage*, MARS finds which basis functions (BFs) are attached to the model by using a fast searching algorithm, and builds a likely large model that usually overfits the dataset. The process continues until the model reaches the maximum number of basis function, which is a specific value set by the users. In fact, BFs contribute both most and least to the overall performance together in this model. That is why the model is more complicated and contains many inaccurate terms in the forward stage. At the *backward stage*, the overfit model is clipped to diminish the complexity of the model. However, the model supports the overall performance with taking into consideration the fit to the data. At this backward stage, the BFs which contribute to the smallest raise in the *residual sum of squares (RSS)* are eliminated from the model at each iteration. Consequently, an optimally estimated model is created [20, 37, 46, 81].

MARS uses extensions in these piecewise linear one-dimensional basis functions of the form $(x-t)_+$ and $(x-t)_-$:

$$(x-t)_+ = \begin{cases} x-t, & \text{if} \quad x > t, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad (x-t)_- = \begin{cases} t-x, & \text{if} \quad x < t, \\ 0, & \text{otherwise.} \end{cases}$$

Each such a map is a *truncated linear function*, with a univariate knot at value t, determined using the dataset. Figure 2.1 shows BF pairs for $t = 0.5$ as an example:

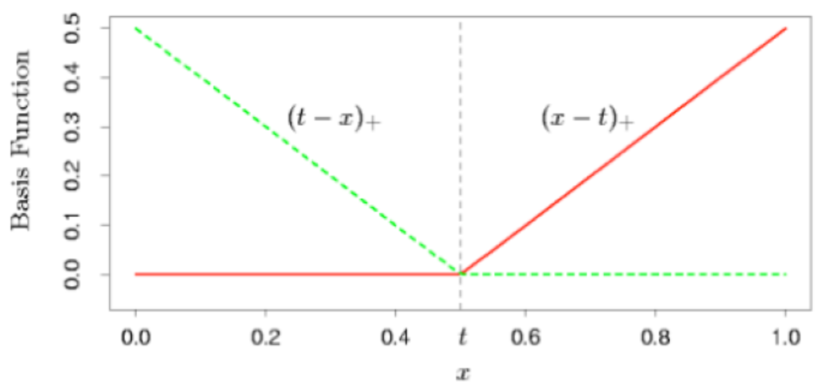

Figure 2.1: The BFs used by MARS for t=0.5 [81].

These two functions are called a *reflected pair*. The aim is to present a reflected pair for each input X_j with knots at each observed values x_{ij} of that input. Thus, the aggregation of the basis functions is

$$C = \{(x_j - t)_+, (x_j - t)_- : \quad t \in \{x_{1j}, x_{2j}, ..., x_{Nj}\}, j = 1, 2, ..., p\}, \tag{2.27}$$

where N is the number of the observations, p represents the dimension of the input space so that, in case that all of the input values are distinct, there are $2Np$ one-dimensional basis functions. Although each of the aforementioned basis function depends only on a single X_j, such as by $h(\boldsymbol{x}) = (x_j - t)_+$, it is regarded as a function over the whole input space $\mathbb{R}^p$.

As it was stated earlier, the "model-building strategy" begins like a "forward stepwise linear regression", but unlike using the original inputs, it is allowed to involve functions from the set C and their products. As a result, the model has the first following form:

$$f(\boldsymbol{x}) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(\boldsymbol{x}). \tag{2.28}$$

This was at the forward stage of the MARS, where $\mathcal{M}$ is the set of basis functions in the present model, and $h_m(\boldsymbol{x})$ are multivariate basis functions from C or products of two or more such functions. Furthermore, β_m stands for the unknown coefficients at the constant 1 $(m = 0)$ or the mth basis function, and $\boldsymbol{x} = (x_1, x_2, ..., x_p)^T$. The form of the mth basis function is represented as follows:

$$h_m(\boldsymbol{x}) = \prod_{k=1}^{K_m} \left(s_{km} \cdot (x_{v(k,m)} - t_{km})\right)_+, \tag{2.29}$$

where K_m is the number of "truncated linear functions" multiplied in the mth basis function, $x_{v(k,m)}$ is the input variable which refers to the kth truncated linear function in the mth basis function, t_{km} is the knot value suitable to the variable $x_{v(k,m)}$, and $s_{km} = \pm 1$.

A *lack-of-fit criterion* is performed to assess the probable basis functions.

To form the model, MARS forward stepwise algorithm begins with the constant function $h_m(\boldsymbol{x}) = 1$ to estimate β_0, and the whole rest of the elements in the set C are candidate functions. There are some obtainable forms of the basis functions $h_m(\boldsymbol{x})$:

- 1,
- x_k,
- $(x_k - t_i)$,
- $x_k x_l$,

- $(x_k - t_i)_+ x_l$,
- $(x_k - t_i)_+(x_l - t_j)_+$.

For each of these basis functions, input variables cannot be identical in the MARS algorithm. For that reason, the basis functions which are stated above use different input variables, x_k and x_l, and their knots t_i and t_j. We take into account as a new basis function pair all products of a function h_m from the model set $\mathcal{M}$ with one of the reflected pairs in C, at each stage. To the model set the term of the form is added which generates the biggest decline in the training error [20, 46, 81]:

$$\hat{\beta}_{M+1} h_l(x) \cdot (x_j - t)_+ + \hat{\beta}_{M+2} h_l(x) \cdot (t - x_j)_+, \qquad h_l \in \mathcal{M}. \tag{2.30}$$

Here, $\hat{\beta}_{M+1}$ and $\hat{\beta}_{M+2}$ are coefficients determined by LS estimation, together with all the rest of the $M+1$ coefficients in the model. Then the resulting products are supplemented the model, and the process ends when the model set $\mathcal{M}$ includes a maximum number of predetermined terms. To give an instance, the subsequent basis functions can be possible candidates of being added to the model [20, 81]:

- 1,
- x_k,
- $(x_k - t_i)$, if x_k is already in the model,
- $x_k x_l$, if x_k and x_l are already in the model,
- $(x_k - t_i)_+ x_l$, if $x_k x_l$ and $(x_k - t_i)$ are already in the model,
- $(x_k - t_i)_+(x_l - t_j)_+$ if $(x_k - t_i)_+ x_l$ and $(x_k - t_i)_+ x_k$ are already in the model.

In the final of this forward process at each step, we have obtained a large model which is frequently overfitting the data. Therefore, a backward deletion is applied, too. At each step, we remove from the model the term which leads to a smallest rise in the residual squared error, by using a backward algorithm. This procedure stops when an optimal number of effective terms is obtained in the last model. Then, this has produced an estimated best model $\hat{f}_\lambda$ of any size (= number of terms) λ. In the MARS model, the so-called *Generalized Cross-Validation (GCV)* is employed to estimate the optimal value of λ; this criterion is identified as

$$GCV(\lambda) = \frac{\sum_{i=1}^{N} (y_i - \hat{f}_\lambda(x_i))^2}{(1 - M(\lambda)/N)^2}. \tag{2.31}$$

Here, the value $M(\lambda)$ is the influential number of parameters in the model, and N is the number of sample observations.

Some classical statistical methods such as from regression techniques can be good at conducting interaction terms. Therefore, they need to try many combinations of the variables in the dataset; hence, these methods cannot provide computations in a feasible way. On the other hand, MARS automatically searches for appropriate interactions among independent variables, which can be particularly preferred whenever there is a large number of interactive variables. MARS identifies interactions, and a comparatively small number of regressor variables which are complex transformations of initial variables. In addition to these advantages, it also provides an opportunity to discover nonlinearities which can appear in the relationship between dependent and independent variables, while also producing graphs that help visualizing and understanding the interactions [20, 81].

For all of the aforementioned reasons, in our study, we prefer to use MARS for identifying systems of SDEs, which will be explained in the subsequent chapters.

CHAPTER 3

THE MUTUAL EFFECT OF FINANCE TO NEURONAL STATE AND BEHAVIOR VIA SENTIMENT

3.1 Introduction

For many centuries, theoreticians have attempted to construct certain models of how people make judgments, decisions, and choices. In this chapter, these decisions and choices are mentioned in terms of finance and explained with their relations to both the neural and the behavioral sides. In addition to the behavioral finance part, the term of sentiment is introduced. Measurement techniques of investor sentiments will be described, too.

3.2 A Relation between Neuro and Finance Sectors

The term *Neuroscience* was introduced in the mid 1960s, to comprehend the structure and function of the normal and the abnormal brain. Today, Neuroscience encompasses a wide variety of research attempts from moleculer biology of nerve cells to disordered behavior, cognition and emotion [59].

In our study, we wish to focus on the relation between neuroscience and the financial processes. We aim to learn answers to questions such as how the neural world can affect financial decisions of investors, in which situations it affects more or less, etc. From this point of view, we have researched on connection between neuroscience and financial processes.

Neuroscience makes a connection with finance to bring about the interdisciplinary area of *Neurofinance*, which comprises all the phenomena and relations, explaining feelings and thoughts that escape from our comprehension. Investors' financial performances which tend to grieve when they are emotionally reactive and lead to poor impulse control, can gradually be calculated by using numerous neuroscience techniques now. These techniques provide a determination of the regions in the brain that are quite active in the course of decision making processes and, especially, decision making in finance [57, 69].

Many financial decisions highly affect people's lives, and these decisions are made at various stages in the economy [22]. There was found an observed neural activity in investment decisions which is consistent with a preference for "social status" and "relative wealth" concern. This neural activity is measured with *functional magnetic resonance imaging* (fMRI) [21]. Most of the new research in the area of neurosciences of social and financial decision-making is obtained from studies on functional magnetic resonance imaging (fMRI). By using an MRI scanner, neuroscientists can watch the changes in the cerebral blood flow as subjects play conversational social games. These games are generally computerized games and based on selecting choices by using buttons. In comparison with other neuroimaging techniques, fMRI is less costly, less disrupted, and it allows for a better analysis [49]. For example, Kuhnen and Knutson (2006) examined how anticipatory neural activity predicts optimal or suboptimal choices of nineteen healthy volunteers, who are divided as "experts" as PhD. students in finance and economics, "non-experts" as PhD. students in humanities, respectively, in a financial decision-making test employing fMRI [34].

While making decisions, to better comprehend the relation between neuroscience and financial processes, regions of the human brain are investigated. Kahneman and Frederick (2002) explained that the processes of the human brain can be subdivided into two components: *control processes* and *automatic processes*. *Control processes* are those that consists of a range of logic and are aforethought. The latter ones, *automatic processes*, occur as a result of unconsious behaviours. Therefore, decisions made by automatic processes have a tendency to occur without conscious effort, and most of the time, humans do not quite figure out why they made such decisions. Controlled processes generaly occur in orbital regions of the brain, e.g., the *pre-frontal cortex* and the *forebrain*. These regions of the brain differentiate human brains from animals' brains. In fact, these regions are the last enhanced and evolved parts in human beings. On the other hand, generally, the automatic processes occur in the "occipital, parietal and temporal" parts of the brain. The *nucleus accumbens* and the *amygdala* are the regions which are involved in these parts. Since there are constraints in resources, automatic processes are far more broadly used than controlled processes. On the controlled processes, the brain becomes activated when an unexpected situation occurs. Referring to these situations, the processes of both risk and return which are the main components of finance can be examined under this aspect [57].

In an evolutionary manner, risk is treated in new functional *mesial prefrontal cortex*, which is actuated in absolute cases, particularly, when "certain stimuli" are induced, such as pain. In one respect, when human beings feel a risk, then the brain system reacts adversely to such risks. Kuhnen and Knutson (2005) conducted a research to understand the mistakes which traders seemingly make in their investment strategies. They considered that these mistakes can take part in two categories: *risk seeking* and *risk-aversion*. It is indicated by an fMRI scanner that the *nucleus accumbens* is activated before risk seeking mistakes and before choices which were risky, while *anterior insula* is activated, whenever there are risk aversion mistakes or choices with almost no risk. These different neural cases lead people to make different financial choices [34].

Similarly to risk, the expectancy of rewards highly influences human behaviour. The

Ventral Striatum (vSt) encodes for a wide range of expected rewards, counting primary rewards like liquid, learned rewards like money and social rewards, and reward-prediction errors [22].

Looking for potentially new reward targets is something that human beings are continually doing. Furthermore, imagining about future rewards increases patience, probably it makes the future more conspicuous. This imagination effect is linked to a relation between the *hippocampus* and *cingulate cortex* [22]. Thus, the effects of these rewards are measured by the *pre-frontal cortex*, which is associated with decision making, personal statement, and social behavior. Besides, *pre-frontal cortex* adjusts the behavior according to these rewards [57].

Here, the main focus in not only behavioral, but also rational theories of financial decision making are of our interest, related with the tradeoff between risk and return. Previous researches in neurofinance started with examining the neural correlates of these two key variables. The former is the larger neural activity in the *ventral-striatum* that was connected with investing in safer assets. The latter result is coherent with an effect of the *anterior insula* in interception of bodily states and emotions, including comfortless sensations such as suspense and pain [22].

Recent works in neuroscience have investigated a wider array of topics such as asset price bubble formation [21]. Price bubbles occur when asset prices increase much more than intrinsic value. These bubbles can be examined in laboratory experiments by constructing a known intrinsic value that springs from monetary payments to subjects holding an asset which consists of 15 trading periods, decreasing in value at each step. It has been observed from the experiments that prices inclined to rise towards a peak beyond the intrinsic value, in contrast with bubbles which are smaller and shorter. In addition to these experiments, one of the fMRI studies employing this paradigm determined that in market sequences which displayed bubbles, activity in the *ventral* and *dorsomedial prefrontal cortex* (PFC) was more satisfying to a trader's portfolio value. Besides, these regions were more activated with each other throughout bubbles. These results suggest that a coordination of valuation and mentalizing in the *ventral* and *dorsal prefrontal* regions is associated with bubbles [22].

Aumann (1997) denoted that people generally do not act in a consciously rational side in their daily lives and activities. He proposed a work called "rules of thumb" which is an evolutionary process. This process is like genes or generally alleles. If they function inadequately, they become sparse and in time they disappear; if they function adequately, they become productive and they reproduce [4].

People nowadays are living longer in most places of the world; how financial decisions are affected by this aging becomes a new, important question. As people are getting old, they become more patient, and they also seem to make poorer financial decisions. In one line with this evidence, reduction in neural activity in the *vSt* is smaller in older adults [22].

People can also change their financial and social decisions according to television advertisings. There have been some neuroscience studies analyzing the relationship between brain activity when they see television advertisings, and ensuing changes in consumer choice. The time gap between advertising watching and consumer choice is sufficiently long to depend on *long-term memory* (LTM) [58].

The main aims of neurofinance are to obtain a better comprehension of financial markets by describing some physiological features influencing trading behavior and trading results, to combine these features with trading results, to improve methods and technologies, and to design a convenient training to enhance the trading performance. As neuroscience methods are becoming less costly and more useful, the interconnection between finance and neuroscience permits to ensure emerging developments into this interface [57, 69].

3.3 A Relation between Behavior and Finance: The Effects of Investor Sentiment

Much of academic finance is rest on the hypothesis that individuals behave rationally, while buying and selling stocks, etc. They try to reach all useful information in the decision-making process to compose "rational expectations" about the future in designating the value of companies and the general situation of the economy. As a result, stock prices represent fundamental values and will only move up and down if there are any unexpected positive or negative news observed. Hence, economists have made an inference that financial markets are durable and adequate, stock prices pursue a "random walk" and the comprehensive economy is tending to a "general equilibrium". In reality, however, investors may not always be as rational as we think [44, 71].

A recent area of finance, named *Behavioral Finance*, studies why and how some traders manage success, while others lose a fortune at gambling. Behavioral Finance models are generally based on a concept of individual investors who tend to make judgment and decision-making errors. That is, Behavioral Finance is a kind of finance from a wider science point of view, containing psychology and sociology. It provides a new approach to financial markets, sheds light on why people purchase or sell stocks and even why they do not purchase stocks at all [44, 71].

Behavioral Finance figures out and expounds how emotions influence investing decisions. Humans differ widely according to facts such as risk aversion, time preference and tastes. The two basic principles of Behavioral Finance are the limit to arbitrage and cognitive psychology. The *limit to arbitrage* explains that it can be challenging for rational traders to compensate the divisions which are done by less rational traders. *Cognitive psychology* addresses how people consider. It states the limit to arbitrage when the market is not sufficient. Besides, it lists the kinds of variations from full rationality that traders have an expectation to see [9, 44].

There is a wide psychology literature indicating reasons why people make systematic errors: they always prefer to make decisions easier (heuristics). They are overconfident, they care too much their current experiences (representativeness). Traders disconnect decisions which should be connected (mental accounting). They introduce incorrectly the individual issues (framing), and finally, they incline to be slow to keep up with the changes (conservatism). Their choices may also generate a deflection when they sell shares whose prices have increased, and they keep the asset whose price has decreased in value (disposition effect) [44]. Apart from these reasons, according to Kent et al. (2001), there are most widespread behaviors' of majority of the investors, while making investment decisions: Investors generally want to benefit from their past performance as a sympton of future performance to make decisions for buying stocks. Investors trade emotionally and ambitiously on stock markets, they often do not attend all asset and security sections, and they behave based on a current situation. Individual investors display a loss-averse behavior, they do not display efficient portfolios all the time. These investors act coherently to each other, and they are impressed by historical high or low trading stocks [44].

Investor behaviors which we mentioned above vary, especially, during the bubble period. Stock market bubbles occur when the prices of stocks rise much more than their fundamental value. This principal value can be considered as the price which investors are keen for paying if they are forced to hold the asset eternally. The investors' confidence levels are very high due to the fact that stock market bubbles are only noticed in course of the continuous bull market situations. They believe that stocks will always provide them profit and the demand for their stocks will continue permanently. This opinion about the stock market leads to irrational anticipations and increases the stock prices upward and also overcharges the size of the bubble in stock market. Hence, before prices decline, they start to sell their stocks, inevitably with the feeling of "panic". Generally, this process ends with a sharp decrease, which is named *bubble burst* [32, 35].

In fact, investors have drawn their attention to a particular point. Kahneman (1974) denoted that people are disposed to "cognitive illusions", such as becoming wealthy and well-known, or being able to find a reasonable way to leave the market before a bubble breaks. People dramatize the factor of skill which they think they have, and they disclaim the task of chance in their decision-making process. Kahneman considers as humans optimistic. Because of this property which comes from human nature, we can make an inference why the casino is full twenty-four hours a day with humans who are looking for a luck. This optimistic human nature encourages investors to purchase stocks and shares when their market prices have attained a historic high. In case of this euphoric market condition, investors should sell their stock and shares [44].

Kent, Hirshleifer and Siew (2002) found that studies about the psychology of investors was conducted by searching the effect of stock returns and variables to each others, on factors, such as weather and societal happiness. These differing investigations are activated by recent theories in psychological economics on visceral factors and the risk-as-feeling point of views. *Visceral factors* are the wide variety of emotions, moods and direct states experienced by people during decision making process. The *risk-as-feeling* view explained that these visceral factors could influence, and even prevail, rational considerations on decisions including risk and uncertainty. This provides anticipated patterns in stock returns, since people in good moods are prone to be more optimistic in their estimates and judgments than people in bad moods. Concerning stock pricing, optimistic or pessimistic approaches about the future anticipations from the business direction are prevalent, stock prices should be higher when most investors are in good moods than they are in neutral or bad moods [44].

Investors have tried to locate “trends and turning points” in stock prices, in equity markets. There is a model which used to describe and analyze trend changes at the begining and to sustain an investment status until we have sufficient evidence to show that the trend has reversed. This process is conducted by the field *Technical Analysis*. According to this analysis, the evidence suggests that subjective probability distributions of investors are too tight, especially, for difficult works as anticipating stock prices. Investors employ a wide range of rules and commitment mechanisms to adjust emotion. Numerous investors trade shares on random tips from people which they know, without planning in advance. A first problem here is that investors are too optimistic about almost everything which is related with their personal life. A second problem is that Sentiment of traders traces the market which leads investors to purchase shares in bull markets and sell these in bear markets [44]. In our study, this is the point that we will concentrate on regarding the Sentiment of investors.

There has been no singular broadly approved definition of *Investor Sentiment* so far. There are various studies about sentiment and aggregate stock returns. A pioneering and famous of them was first seen in the 1980s [8]. Investor Sentiment was characteristically assessed as a myth by general financial theories and has not much appealed by researchers previous to 1990 [78]. These studies were mostly atheoretical, a testing in many kinds of ways whether the entire stock market could be mispriced [8]. The Sentiment term is considered in various ways by financial analysts, academic researchers, and the media, many of them come to terms that Investor Sentiment is economically important. Some researchers address Investor Sentiment as a tendency to trade on noise instead of information. On the other hand, some of them used it in a colloquial manner, such as “investor optimism or pessimism”. The term Sentiment has an association with feelings; therefore, the media sometimes addresses it as *investor fear* or *risk-aversion* [78].

Some stocks are more susceptible to wide changes in the tendency to speculate, because of the subjectivity of their valuations. The lack of an acquisition history together with the existence of apparently unlimited growth opportunities, such as canonical young, uprofitabless or excessive growth stock, allows inexperienced investors to trade, with equal reasonableness, a broad spectrum of valuations. These are ranging from “much too low” to “much too high” with respect to their sentiment. Throughout a bubble period, when the tendency to speculate is high, this profile of features also allows investment bankers to assess further for the high result of valuations [6]. Most of these companies go bankrupt, when the bubble burst. As a result, this brings about a rise in the unemployment rates [35]. In contrast to this, the value of a firm with a long acquisition history, tangible assets and stable dividends is much less subjective and, therefore, its stock is possibly less influenced by variations in the tendency to speculate. That is, investors with a low tendency to speculate are able to demand profitable, dividend-paying stocks not just because profitability and dividends are associated with some unperceivable firm property which identifies safety to the investor. But for sure because the obvious characteristics “profitability” and “dividends” are mainly taken to define a safety channel [6].

A growing amount of theoretical and empirical works shows that arbitrage inclined to be particularly risky and expensive for new, small, profitless, excessive growth, or constrained stocks. The main point of this discussion is that the same stocks which are the most difficult to arbitrage prone to be the most difficult to value as well [6].

Tetlock (2007) stated that “Most theoretical models of the effect of investor sentiment on stock market pricing make two important assumptions (see, e.g., DeLong et al. (1990a))”. One of these assumptions is that models assume the presence of two kinds of traders, the ones called *noise traders* whose beliefs are random about future dividends. The other ones are named as, *rational arbitrageurs* whose beliefs are Bayesian.

Investor Sentiment models, as proposed by, e.g., DeLong et al. (1990a), anticipate that a low sentiment will constitute a downward price pressure, and extraordinarily increased or decreased values of sentiment will constitute a “high volume” [65].

The limits of arbitrage can be determined by utilizing the misperceptions of these noise traders. Arbitrageurs have a tendecy to be risk averse and to have reasonably short horizons. Consequently, they have limited willingness to take positions against noise traders. If noise traders are pessimistic and therefore, have decreased asset’s price today, an arbitrageur buying this asset should consider that noise traders might become even more pessimistic and decrease the price even beyond in the near future. When the arbitrageur has to liquidate before the price gets well, he suffers a loss. His original arbitrage position can be bounded because of the fear of this loss. In reverse, an arbitrageur selling an asset short, when bullish noise traders have increased its price, must consider that noise traders might become even more bullish tomorrow. Hence noise traders must make a decision which explains the risk of a beyond price rise when he has to purchase back the stock. The risk coming from stochastic shifts in noise traders’ beliefs and ideas increases the probability by which on-average bullish noise traders earn a higher expected return, than rational, advanced investors deal with arbitrage against noise trading. This result is attained because a noise trader’s risk makes assets less winning to risk-averse arbitrageurs, and thus it decreases prices. If noise traders on-average overvalue returns or undervalue risk, they invest more into the risky asset in an on-average way than advanced investors. Furthermore, they have a possibility to earn higher average returns [17]. Consequently, noise traders’ anticipations about asset returns are susceptible to movement in sentiment [78].

Effects of Investor Sentiment can differ according to different countries and cultures. Chui, Titman and Wei (2008) asserted that cultural differences might play a role for the behavioral biases between countries. For example, collectivistic countries have societies where people are in powerful groups. This construction causes a consensus opinion among society and, therefore, leads to *herd-like overreaction*. Herd-like overreaction described as related actions of noise traders based on extremely optimistic or pessimistic expectations. This reaction is exactly what is assumed to direct the sentiment-return relation in financial markets. In stock markets, it is expected that stock markets in pluralistic countries are more affected by Investor Sentiment compared to individualistic countries. These countries are where people are liable to discourse on their own information and opinions and, as a result, are less influenced by behavioral biases [54].

While examining these assumptions and descriptions about Sentiment, it may be possible to measure the Sentiment. Several Sentiment measures currently exist, which vary from measures derived for the purpose of an academic article to daily indices, used for opposite aims by options traders.

Some of the widely used Sentiment measures are listed by Zhang (2008) which are: “Closed-end fund discount consumer confidence indices, Investor intelligence surveys, Market liquidity, Implied volatility of index options, Ratio of odd-lot sales to purchases, and Net mutual fund redemptions”. These measures show the fact that there is no obvious, general explanation of Sentiment by both academic researchers and professional traders. The former ones adopt Investor Sentiment measures to make discussios in favor of market efficiency or to tell the reason and the result of the specific movements. The latter are likely to adopt Sentiment indicators as a probable trading tool. However, most researchers cannot even commonly share with each other a number of same ideas about a basic definition of sentiment. Thus, there are obviously conflicts about how to measure it. Zhang (2008) examined two approaches for measuring sentiment “indirect market-based proxies for sentiment” and “direct survey data”. The first one is the *market-based* approach, which aims to accumulate Sentiment from financial proxies indirectly. To illustrate this, the the put-call ratio and closed-end fund discount are widely mentioned as proxies for sentiment obtained from market reactions. The second approach is to *measure sentiment directly*, employing and benefitting from surveys and questionnaires given to investors. Yale School of Management’s Stock Market Confidence Index (YSMI), University of Michigan’s Consumer Confidence Index (MCCI), CFA Institute’s Global Market Sentiment Survey, and American Association of Individual Investors’ (AAII) weekly poll of Investor Sentiment can be given as examples [78].

Baker and Wurgler (2006) elaborated that “Candidate ways of measuring sentiment, ordered from origins in investor psychology to responses by corporate insiders, include: surveys, mood proxies, retail investor trades, mutual fund flows, trading volume, dividend premia, closed-end fund discounts, option implied volatility, first-day returns on initial public offerings, volume of initial public offerings, new equity issues, and insider trading”[8]. Let us briefly mention about the effects of some of these:

Investor Surveys: By learning investors’ optimism level from these surveys, we can get a view about the marginal investor behaviour [8].

Investor Mood: Stock prices are tied up with exogenous changes in human emotions. For example, it has been found that the market returns are on average lower in fall and winter which is the seasonal effect of the disorder. In fact, benefitting much less from hours of the daylight causes depressive disorder [8].

Retail Investor Trade: There is a higher possibility for the inexperienced retail or individual investor to be subject to sentiment when compared with a professional investor. As an example, Greenwood and Nagel (2006) found that during the peak of the internet bubble, younger investors were more hopeful than older investors to buy stocks [8].

Insider Trading: Since institutional managers have a better information about the real value of their firms or companies than outside traders, managers’ personal portfolio decisions may unearth their opinions about the mispricing of their firms either. If Sentiment brings about connected mispricings across firms, insider trading patterns possibly contain a systematic sentiment component [8].

Trading Volume: This also named by, liquidity, can also be seen as an investor sentiment index. Baker and Stein (2004) denoted that if short-selling is more expensive than opening and closing long positions, irrational investors are more prone to trade, and so they supplement liquidity. Those investors display this kind of behaviour when they are optimistic, and gambling on rising stocks in return for when they are pessimistic and gambling on falling stocks. If there is a rise in trading volume, then it will reflect the attendance of self-assertive investors in the market, and show an rise in Investor Sentiment [8, 38].

Dividend premium: This term was defined by Baker and Wurgler (2004) as "the difference between the average market-to-book-value ratios of dividend payers and nonpayers". The dividend premium points out sufficiently the main historical trends in firms' tendency to pay dividends, which means, when dividends are at a "premium", firms are more inclined to pay them. Hence, when dividends are "discounted", the firms pay less. They seem to provide succession sentiment for or against "safety" when deciding on whether to pay dividends [8].

Closed-end Fund discount: Closed-end funds are investment companies which put on the market a fixed number of shares, and then trade on stock exchanges. The closed-end fund "discount" (or, occasionally, "premium") is defined as the "difference between the net asset value of a fund's actual security holdings and the fund's market price". The fundamental value of a closed-end fund is completely its net asset value. Various authors have discussed that if closed-end funds are unequally kept by retail investors, the average discount on closed-end equity funds may be a sentiment index [8, 78].

IPO First-Day Returns: Numerous modelled facts about earning such remarkable returns on their first trading day of Initial Public Offerings, or IPOs, have been considered to be expressed by "investor enthusiasm". Average first-day returns exhibit summits and troughs which are highly correlated with IPO volume mentioned subsequently [8, 78].

IPO Volume: The IPOs is generally said to be highly fragile to investor sentiment. Investment bankers mention "windows of opportunity" for an IPO which flexibly opens and closes. This variability could be an explanation for the fact that IPO volume exhibits wild fluctuations [8, 78].

Equity Issues Over Total New Issues: Baker and Wurgler (2000) found that "high values of the equity share portend low stock market returns, and suggest that this pattern reflects firms shifting successfully between equity and debt to reduce the overall cost of capital". But it does not have to indicate that individual firms or managers of them can predict all prices on the market [8, 78].

To measure Sentiment, Baker and Wurgler (2006) preferred to construct a sentiment index which averages six commonly used proxies instead of addressing a single measure for Investment Sentiment. In the following chapters, we will explain the application of this *Theory of Sentiment Index* deeply.

CHAPTER 4

REPRESENTATION AND PARAMETER IDENTIFICATION OF SYSTEMS OF SDEs IN FINANCIAL AND HUMAN FACTOR: SENTIMENT PROCESSES

4.1 Introduction

In this chapter, we describe and combine SDE in different kinds of fields; one of them represents a financial process and the other one represents sentiment. These SDEs are explicitly related to each other through the state variables. Herewith, we will obtain coupled systems of SDEs indeed. Then, parameters of these SDEs will be estimated by MARS method, which provides a smoother approximation for the data.

4.2 Representation of SDEs in terms of Finance and Human factor: Sentiment

As mentioned previously, SDEs have rapidly become a most well-known format in which one can express mathematical models of processes under various forms of uncertainty, such as in finance, neuroscience, neural networks, micro-economic systems, and human factors [15]. They are a main tool to describe the randomness of a dynamical system today [73].

In a financial system, different kinds of SDEs have been gradually refined to model specific financial products or classes of products. Financial institutions benefit from SDEs to price their derivatives or to gauge the risks of their portfolios [73].

Similarly, economists often want to model problems in a way which integrates the aspects on how today's choices will affect future decisions (which is modeled by the drift part), but also allows for the fact that there will be small random effects shifting people's state in future periods (which is modeled by the diffusion part, representing the randomness) [28].

Feelings can influence thinking styles. Positive mood brings about more creative solutions; however, they might be risky, while negative mood brings about more prudent thinking. These kinds of optimistic or pessimistic expectations can persist and affect asset prices for significant periods of time, eventually leading to crises [12, 82].

All these emotions, opinions, views and ideas are based on a feeling about a situation described by the word *sentiment*, which we have already broadly explained in the previous chapter.

In order to represent the influence of any financial process and a process of sentiment by traders on each other, we describe a coupled system of two SDEs.

Our system of SDEs is the formulation of a two-dimensional stochastic process $\boldsymbol{X} = (\boldsymbol{X_t} : t \in [0,\infty))$, also denoted by $(\boldsymbol{X}_t)_{t\geq 0}$, driven in terms of two differential equations. From these two SDEs, the first one (X_t^1) describes a branch of an economy or financial process, which is affected by sentiment, while the second SDE (X_t^2) describes sentiment which is affected by financial process.

The formulas below involve random fluctuations, where each state variable has just one Brownian motion in its equation. Then, our coupled system of SDEs is on the following form:

$$\begin{aligned} dX_t^1 &= a^1(X_t^1, X_t^2, t)dt + b^{11}(X_t^1, X_t^2, t)dW_t^1, \\ dX_t^2 &= a^2(X_t^1, X_t^2, t)dt + b^{22}(X_t^1, X_t^2, t)dW_t^2. \end{aligned} \tag{4.1}$$

With Equation (4.1) we have an example of multi-dimensional SDE, which we introduced in its general structure in Chapter 2. Differently from the general structure, in our model, we decrease the model complexity in order to increase the stability of the model (i.e., to decrease its sensitivity) against various forms of noise and perturbation. For this purpose, our system does not include the mixed terms (b^{ij}) with $i \neq j$.

In our study case, we refer to $d = m = 2$, where $\boldsymbol{a}(\boldsymbol{X}_t, t)$ constitutes a drift part, $\boldsymbol{b}(\boldsymbol{X}_t, t)$ establishes the diffusion part, and $(\boldsymbol{W}_t)$ denotes a two-dimensional Brownian motion at time t:

$$\boldsymbol{X}_t = \begin{bmatrix} X_t^1 \\ X_t^2 \end{bmatrix}, \quad \boldsymbol{a}(\boldsymbol{X}_t, t) = \begin{bmatrix} a^1(X_t^1, X_t^2, t) \\ a^2(X_t^1, X_t^2, t) \end{bmatrix}, \quad d\boldsymbol{W}_t = \begin{bmatrix} dW_t^1 \\ dW_t^2 \end{bmatrix},$$

$$\boldsymbol{b}(\boldsymbol{X}_t, t) = \begin{bmatrix} b^{11}(X_t^1, X_t^2, t) & 0 \\ 0 & b^{22}(X_t^1, X_t^2, t) \end{bmatrix}.$$

4.3 Parameter Identification for Systems of Stochastic Differential Equations with Multivariate Adaptive Regression Splines

Estimation of the parameters of SDEs is significant in practice because in many applications, such as modelling of the behavior of stock prices, the deterministic components in the dynamics are not easy to represent [11, 80]. In our study, SDEs are considered to approximate the connection between our mental spheres and the real world of finance. Many of the SDEs do not have closed-form solutions. Therefore, in general it is necessary to numerically approximate them through a time-discretized version of the solutions. Our method for estimating the drift and diffusion terms of SDEs is based on both a discretization of the SDEs and an employment of MARS method [53, 75, 80].

4.4 Discretization of Stochastic Differential Equations

Simulation methods are generally based on discrete approximations for continuous solution of SDEs. There are various discretization schemes for SDEs such as Euler, Milstein and Runge-Kutta discretization. One of the most well-known and widely applied schemes of approximation is *Euler's method*. It is originally used to produce solutions to ODEs. There is also *Euler-Maruyama method*, which is an analogue of the Euler method for ODEs. Here, for the SDEs the solution is approximated by Euler-Maruyama method instead of Euler method [53, 80].

From the side on strong-order on convergence, the Euler method for ODEs has order 1.0, while the Euler-Maruyama method for SDEs has strong order 1/2. To construct a strong-order 1.0 method for SDEs, another term has to be added to the method. At this point, *Milstein scheme* should be considered by which a 1.0 scheme is obtained as a result of an application of Itô Formula. That is, it makes use of Itô Lemma to raise the accuracy of the approximation by including a second-order term. Then, we can say that the Milstein scheme is same with the Euler-Maruyama scheme if there is no function of x in the diffusion part of the equations. Mostly, Milstein method converges to the right stochastic solution process more rapidly than Euler-Maruyama if the step size goes to zero [29, 53, 80].

The distribution of the process is often unknown, so that the simulated discretized version of the SDE should be simulated. We represent the Milstein approximation (X_{t_i}), with respect to any increasing sequence of times $t_i \geq 0$, in short: $\boldsymbol{x}_i$ $(i = 1, 2, \ldots, N)$ [75]. Our case is two-dimensional, with the kth components of Milstein scheme, where $k = 1, 2$. A general representation of Milstein scheme for the multi-dimensional case is given by [75]

$$\hat{x}_{i+1}^k = \hat{x}_i^k + a^k(\hat{\boldsymbol{x}}_i, t_i)(t_{i+1} - t_i) + \sum_{j=1}^{2} b^{k,j}(\hat{\boldsymbol{x}}_i, t_i)(\hat{w}_{i+1}^j - \hat{w}_i^j) + \sum_{j_1, j_2=1}^{2} L^{k,j_1} b^{k,j_2}(\hat{\boldsymbol{x}}_i, t_i) I_{j_1,j_2}. \tag{4.2}$$

Here, for $j_1 = j_2$:

$$I_{j_1,j_2} := \tfrac{1}{2}((\hat{w}_{i+1}^{j_1} - \hat{w}_i^{j_1})^2 - (t_{i+1} - t_i)),$$

and, for $j_1 \neq j_2$:

$$I_{j_1,j_2} := \tfrac{1}{2}((\hat{w}_{i+1}^{j_1} - \hat{w}_i^{j_1}) - (\hat{w}_{i+1}^{j_2} - \hat{w}_i^{j_2}));$$

and we refer to the differential operator

$$L^{k,j_1} := \sum_{l=1}^{2} b^{k,j_1}(\hat{\boldsymbol{x}}_i, t_i)\tfrac{\partial}{\partial x^l}.$$

In our coupled system of SDEs, $b^{12} = b^{21} = 0$. Thus, we confine us to $j_1 = j_2$ and $k = j$.

Then, Equation (4.2) has the form:

$$\hat{x}_{i+1}^k = \hat{x}_i^k + a^k(\hat{\boldsymbol{x}}_i, t_i)(t_{i+1} - t_i) + b^{k,k}(\hat{\boldsymbol{x}}_i, t_i)(\hat{w}_{i+1}^k - \hat{w}_i^k) + L^{k,k}b^{k,k}(\hat{\boldsymbol{x}}_i, t_i)I_k. \quad (4.3)$$

Here, $\hat{w}_i$ stands for an estimation of W_{t_i}. When we refer to t_i, sometimes we denote W_{t_i} by W_i. We write $\bar{W}_i$ for real-valued denotations in contexts of random numbers drawn and data referred to. Let us particularly advert to finitely many given data points $(\bar{\boldsymbol{x}}_{\boldsymbol{i}}, \bar{t}_i)$, where $0 \leq \bar{t}_1 < \bar{t}_2 < \ldots < \bar{t}_i < \bar{t}_{i+1} < \ldots < \bar{t}_N$ are sampling times. By rearranging Equation (4.3) and inserting the values of both the data and simulated Brownian motions into the discrete dynamics, we obtain a more symbolic form [75, 80]:

$$\bar{y}_i^k = a^k(\bar{\boldsymbol{x}}_i, \bar{t}_i) + \frac{1}{\bar{h}_i} b^{k,k}(\bar{\boldsymbol{x}}_i, \bar{t}_i)(\Delta\bar{w}_i^k) + \frac{1}{\bar{h}_i} L^{k,k}b^{k,k}(\bar{\boldsymbol{x}}_i, \bar{t}_i)I_k. \quad (4.4)$$

Here, $\bar{y}_i^k$ represents the difference quotient which comes from the ith and $(i+1)$th data values $\bar{x}_i^k$ and $\bar{x}_{i+1}^k$, respectively, on step lengths $\Delta\bar{t}_i := \bar{h}_i = \bar{t}_{i+1} - \bar{t}_i$ between neighbouring sampling times [75, 80]:

$$\bar{y}_i^k = \begin{cases} \frac{1}{\bar{h}_i}(\bar{x}_{i+1}^k - \bar{x}_i^k), & \text{if} \quad i = 1, 2, \ldots, N-1, \\ \frac{1}{\bar{h}_N}(\bar{x}_N^k - \bar{x}_{N-1}^k), & \text{if} \quad i = N. \end{cases}$$

Since $W_{t_i}^k \sim N(0,t)$ $(k = 1, 2)$, the increments $\Delta W_{t_i}^k$ are independent on non-overlapping intervals. This means that $W_{t_2}^k - W_{t_1}^k, \ldots, W_{t_N}^k - W_{t_{N-1}}^k$ are independent, for any times $0 < t_2 < t_2 < \ldots < t_N$. Furthermore, Var$(\Delta W_{t_i}^k) = \Delta\bar{t}_i$, where we refer to increments $\Delta W_{t_i}^k := W_{t_{i+1}}^k - W_{t_i}^k$ and $\Delta\bar{t}_i := t_{i+1} - t_i$. That is why, the increments having normal distribution can be simulated with the help of standard normal distributed random numbers Z_{t_i}. To simplify notation, instead of writing $W_{t_i}^k$ and $Z_{t_i}^k$ we prefer to write W_i^k and Z_i^k, respectively. Consequently, we have a discrete model for a Wiener process [80]:

$$\Delta W_i^k = Z_i^k \sqrt{\Delta\bar{t}_i}, \qquad Z_i^k \sim N(0,1). \quad (4.5)$$

By inserting Equation (4.5) into our Milstein approximation and drawing random numbers $\bar{z}_i$ of Z_i, we obtain the following form:

$$\bar{y}_i^k = a^k(\bar{\boldsymbol{x}}_i, \bar{t}_i) + b^{k,k}(\bar{\boldsymbol{x}}, \bar{t}_i)\frac{\bar{z}_i^k}{\sqrt{\bar{h}_i}} + \frac{1}{\bar{h}_i}L^{k,k}b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)I_k. \tag{4.6}$$

If we abbreviate this in a more compact form, it can be written as follows:

$$\bar{y}_i^k = \bar{G}_i^k + \bar{H}_i^k + \bar{F}_i^k. \tag{4.7}$$

Here, $\bar{G}_i^k := a^k(\bar{\boldsymbol{x}}_i, t_i)$, $\quad \bar{H}_i^k := (1/\sqrt{\bar{h}_i})b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)\bar{z}_i^k$ and $\bar{F}_i^k := (1/\bar{h}_i)L^{k,k}b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)I_k$.

4.5 Construction of Minimization Problem for Parameter Estimation of Stochastic Differential Equations

For the purpose of finding the minimum of the entire squared residual differences between the right- and left-hand sides of Equation (4.7) (k=1,2), the following minimization problem is introduced:

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \sum_{i=1}^{N}\sum_{k=1}^{2}(\bar{y}_i^k - (\bar{G}_i^k + \bar{H}_i^k + \bar{F}_i^k))^2, \tag{4.8}$$

where the vector $\boldsymbol{\theta}$ will contain all the unknown coefficients in our upcoming parameter representation of these functions: $\bar{G}_i^k$, $\bar{H}_i^k$ and $\bar{F}_i^k$. We state that also larger vector-valued processes could be studied, addressing sums of squared Euclidean norms $\|\cdot\|_2^2$ [75, 80].

The data underlying our system of stochastic processes is situated both in the financial sector and on the human factor: sentiment, have a high variation. Therefore, classical linear Gaussian least-squares estimation can not be sufficient to solve this optimization problem given in Equation (4.8). We have to employ a parameter estimation method which will regard this high fluctuation and provide a smoother approximation of the data [75, 80].

4.6 Parameter Estimation of Stochastic Differential Equations using Multivariate Adaptive Regression Splines

Spline approximation is *adaptive*, so it is quite flexible. A large fluctuation might occur from a high-degree polynomial approximation, and it may be based on strongly diversifying data and outliers or “spikes” existing. Spline approximation can be organized to avoid this large fluctuation of the solutions. Splines are representable as linear combinations of basis splines which are generally low-dimensional. We will approximate the data $(\bar{\boldsymbol{x}}_i, \bar{t}_i)$ “smoothly”, but we have to come to terms with smoothness by the simplicity of certain discrete nondifferentiabilities [20, 75, 80].

In our study, we use MARS basis functions (BFs) as our splines. MARS’ 1-dimensional piecewise linear functions form a set of BFs, shown as reflected pairs in Equation (2.27). This collection $C = \{(x_j - t)_+, (x_j - t)_- : t \in \{x_{1j}, x_{2j}, ..., x_{Nj}\}, j = 1, 2, ..., p\}$ was already introduced in Chapter 2.

In our study, a special MARS application is conducted in the dimension $p = 2$. Therefore, we approximate each of the three functions underlying the terms $\bar{G}_i^k := a^k(\bar{\boldsymbol{x}}_i, t_i)$, $\bar{H}_i^k := (1/\sqrt{\bar{h}_i})b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)\bar{z}_i^k$ and $\bar{F}_i^k := (1/\bar{h}_i)L^{k,k}b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)I_k$ by using the basis functions of MARS and describing a “regularization” through GCV. Since MARS is appropriate for large and high-variational problems, this procedure will be very useful for the stability of our model, under the existence of numerous and highly diverging data [75].

Our purpose is to estimate systems of SDEs in a simplified manner by using MARS method. That is why we make an approximation, including with a control of complexity, by the help of MARS. For each function we use the multivariate-additive form of MARS BFs with respect to subsets of variables (coordinates), i.e., with subvectors. Addressing each of those subvectors, the basis functions are multiplicative, which we will state below. The approach to $\bar{G}_i^k$, $\bar{H}_i^k$ and $\bar{F}_i^k$ (k=1,2) with MARS’ basis functions is represented as follows [75, 80]:

$$
\begin{aligned}
\bar{G}_i^k &= a^k(\bar{\boldsymbol{x}}_, t_i) = \beta_0^k + \sum_{l=1}^{d^{g,k}} \beta_l^k b_l^k(\bar{\boldsymbol{u}}_{i,b}^{l,k}),\\
\bar{H}_i^k &= \frac{1}{\sqrt{\bar{h}_i}} b^{k,k}(\bar{\boldsymbol{x}}_i, t_i)\bar{z}_i^k = \gamma_0^k + \sum_{m=1}^{d^{h,k}} \gamma_m^k c_m^k(\bar{\boldsymbol{u}}_{i,c}^{m,k}),\\
\bar{F}_i^k &= \frac{1}{\bar{h}_i} L^{k,k} b^{k,k}(\bar{\boldsymbol{x}}_i, t_i) I_k = \delta_0^k + \sum_{n=1}^{d^{f,k}} \delta_n^k d_n^k(\bar{\boldsymbol{u}}_{i,d}^{n,k}),
\end{aligned}
\tag{4.9}
$$

where we use the unifyfing notation of 3-dimensional vectors $\bar{\boldsymbol{u}}_{i,b}^{l,k}, \bar{\boldsymbol{u}}_{i,c}^{m,k}, \bar{\boldsymbol{u}}_{i,d}^{n,k} = (\bar{\boldsymbol{x}}_i, \bar{t}_i)$ $(i = 1, 2, \ldots, N)$, and $d^{g,k}, d^{h,k}, d^{f,k} \in \mathbb{N}_0$ are suitable numbers $(k = 1, 2)$. The form of b_l^k, c_m^k, d_n^k in Equation (4.9) is given by MARS BFs built, e.g., as stated subsequently:

$$
\begin{aligned}
b_l^k(\bar{\boldsymbol{u}}_b^{l,k}) &= \prod_{p=1}^{3} (s_{\mathcal{K}_p^{l,k}} \cdot (\bar{u}_{\mathcal{K}_p^{l,k}} - \tau_{\mathcal{K}_p^{l,k}}))_+,\\
c_m^k(\bar{\boldsymbol{u}}_c^{m,k}) &= \prod_{p=1}^{3} (s_{\mathcal{K}_p^{m,k}} \cdot (\bar{u}_{\mathcal{K}_p^{m,k}} - \tau_{\mathcal{K}_p^{m,k}}))_+,\\
d_n^k(\bar{\boldsymbol{u}}_d^{n,k}) &= \prod_{p=1}^{3} (s_{\mathcal{K}_p^{n,k}} \cdot (\bar{u}_{\mathcal{K}_p^{n,k}} - \tau_{\mathcal{K}_p^{n,k}}))_+.
\end{aligned}
\tag{4.10}
$$

Here, $(q)_+ = \max\{0, q\}$ $(q \in \mathbb{R})$, $\bar{u}_{\mathcal{K}_p^{l,k}}$, $\bar{u}_{\mathcal{K}_p^{m,k}}$ and $\bar{u}_{\mathcal{K}_p^{n,k}}$ are the pth truncated linear variables involved in the lth, mth and nth BFs, $\tau_{\mathcal{K}_p^{l,k}}$, $\tau_{\mathcal{K}_p^{m,k}}$ and $\tau_{\mathcal{K}_p^{n,k}}$ are the knot values related to the variables $\bar{u}_{\mathcal{K}_p^{l,k}}$, $\bar{u}_{\mathcal{K}_p^{m,k}}$ and $\bar{u}_{\mathcal{K}_p^{n,k}}$, and $s_{\mathcal{K}_p^{l,k}}$, $s_{\mathcal{K}_p^{m,k}}$ and $s_{\mathcal{K}_p^{n,k}}$ are selected signs $+1$ or -1, respectively. For the sake of better reading, we have not displayed all the dependences on discrete parameters.

To closely represent the residual sum of squares (RSS) which is the summation term in Equation (4.8), first we need to describe the following expression for k=1,2:

$$\begin{aligned}\bar{G}_i^k + \bar{H}_i^k + \bar{F}_i^k &= \beta_0^k + \sum_{l=1}^{d^{g,k}} \beta_l^k b_l^k(\bar{\boldsymbol{u}}_{i,b}^{l,k}) + \gamma_0^k + \sum_{m=1}^{d^{h,k}} \gamma_m^k c_m^k(\bar{\boldsymbol{u}}_{i,c}^{m,k}) + \delta_0^k + \sum_{n=1}^{d^{f,k}} \delta_n^k d_n^k(\bar{\boldsymbol{u}}_{i,d}^{n,k}) \\ &= \bar{\boldsymbol{A}}_i^k \boldsymbol{\theta}^k,\end{aligned} \tag{4.11}$$

with $\bar{\boldsymbol{A}}_i^k := (\boldsymbol{b}_i^k, \boldsymbol{c}_i^k, \boldsymbol{d}_i^k)$, where $\boldsymbol{b}_i^k := (1, b_1^k(\bar{\boldsymbol{u}}_{i,b}^{l,k}), b_2^k(\bar{\boldsymbol{u}}_{i,b}^{l,k}), \ldots, b_{d^{g,k}}^k(\bar{\boldsymbol{u}}_{i,b}^{l,k}))$, $\boldsymbol{c}_i^k := (1, c_1^k(\bar{\boldsymbol{u}}_{i,c}^{m,k}), c_2^k(\bar{\boldsymbol{u}}_{i,c}^{m,k}), \ldots, c_{d^{h,k}}^k(\bar{\boldsymbol{u}}_{i,c}^{m,k}))$, $\boldsymbol{d}_i^k := (1, d_1^k(\bar{\boldsymbol{u}}_{i,b}^{n,k}), d_2^k(\bar{\boldsymbol{u}}_{i,b}^{n,k}), \ldots, d_{d^{f,k}}^k(\bar{\boldsymbol{u}}_{i,b}^{n,k}))$, and $\boldsymbol{\theta}^k := (\boldsymbol{\beta}^T, \boldsymbol{\gamma}^T, \boldsymbol{\delta}^T)^T$, where $\boldsymbol{\beta}^k := (\beta_0, \beta_1, \ldots, \beta_{d^h})^T$, $\boldsymbol{\gamma}^k := (\gamma_0, \gamma_1, \ldots, \beta_{d^f})^T$, $\boldsymbol{\delta}^k := (\delta_0, \delta_1, \ldots, \delta_{d^g})^T$.

Then, we obtain the RSS as the sum of squared lengths of the difference vectors between $\bar{\boldsymbol{y}}^k$ and $\bar{\boldsymbol{A}}^k \boldsymbol{\theta}^k$, where the matrix $\bar{\boldsymbol{A}}^k = (\bar{\boldsymbol{A}}_1^T, \bar{\boldsymbol{A}}_2^T, \ldots, \bar{\boldsymbol{A}}_N^T)^T$ involves the row vectors $\bar{\boldsymbol{A}}_i^k$, and the vector of difference quotients $\bar{\boldsymbol{y}}^k = (\bar{y}_1^k, \bar{y}_2^k, \ldots, \bar{y}_N^k)^T$ represents the change rates of the given data (difference quotients). Herewith, the objective function of Equation (4.8) takes the following form [75]:

$$\sum_{i=1}^{N} \sum_{k=1}^{2} (\bar{y}_i^k - \bar{\boldsymbol{A}}_i^k \boldsymbol{\theta}^k)^2 = \sum_{k=1}^{2} \left\| \bar{\boldsymbol{y}}^k - \bar{\boldsymbol{A}}^k \boldsymbol{\theta}^k \right\|_2^2 . \tag{4.12}$$

For a compact notation, the single matrix-vector form of term in Equation (4.12) looks as follows:

$$\left\| \bar{\boldsymbol{y}} - \bar{\boldsymbol{A}} \boldsymbol{\theta} \right\|_2,$$

where, in respective dimensions,

$$\bar{\boldsymbol{y}} = \begin{bmatrix} \bar{\boldsymbol{y}}_1^1 \\ \vdots \\ \bar{\boldsymbol{y}}_N^1 \\ \bar{\boldsymbol{y}}_1^2 \\ \vdots \\ \bar{\boldsymbol{y}}_N^2 \end{bmatrix}, \quad \bar{\boldsymbol{A}} = \begin{bmatrix} \bar{\boldsymbol{A}}^1 & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{A}}^2 \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}^1 \\ \boldsymbol{\theta}^2 \end{bmatrix}.$$

CHAPTER 5

A REAL-WORLD APPLICATION

5.1 Introduction

Until now, we have constructed a coupled system of SDEs with two differential equations, one of them has been represented by a financial process and the other one has been represented on the Sentiment of Investors. Afterwards, we have approximated the solutions of these SDEs under Milstein discretization scheme. Since our stochastic processes have a high variation, we have used MARS method as a parameter estimation which avoids this high variation and gives a smoother approximation of the underlying data. In this chapter, to test our approach and give a numerical example we use two different datasets which are taken from the real world; for application of the financial process we will address the Consumer Price Index (CPI) and for the Investor Sentiment we will refer to the Sentiment Index.

5.2 General Information and Description of the Consumer Price Index (CPI)

The CPI is employed as the broadest measure of inflation in the United States. It influences almost all Americans in various ways. It is used as an economic symptoms test, a regulator of US dollar value, and a deflator in regards to other economic events. CPI is based on the expenses of all classes of people living in urban and metropolis cities. This includes the self-employed, the unemployed, the indigent, retired, salary earners and clerical employees. The CPI reflects the expenditure of goods and services on a monthly basis for the entire country. Those goods and services include food, beverages, dwelling, attire, health benefit, transportation, education and communication, etc. [30, 60].

5.3 General Information and Description of the Investor Sentiment Index

As we have denoted earlier in Chapter 3, Zhang (2008) examined two approaches for measuring Sentiment, namely, “indirect market-based proxies” for sentiment and “direct survey data”. The aim of an approach which is based on the market, is to gather sentiment from financial proxies implicitly [78]. In our application, we use a dataset of Investor Sentiment which also consists of the financial proxies.

To measure Sentiment, instead of using a single measure of Investment Sentiment, Baker and Wurgler (2006) preferred to construct an index which averages six mutually employed proxies. They formed "a composite index of sentiment that is based on the common variation in six underlying proxies for Sentiment: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium" [6, 78].

We have already introduced these 6 proxies in Chapter 3 together with other candidate sentiment measures. Here, we explain where these proxies are taken from and how they are formalized in general sentiment indexes.

The *closed-end fund discount (CEFD)*, as indicate in Chapter 3, is defined by Baker and Wurgler as the "average difference between the net asset values (NAV) of closed-end stock fund shares and their market prices". They claim that CEFD is inversely related to sentiment [78]. They took the closed-end fund discount for 1934 to 1964 ("domestic stock funds") from Neal and Wheatley (1998); for 1965 to 1985 (general equity funds only) from Lakonishok, Shleifer, Vishny (1991); for 1986 from CDA/Wiesenberger; and for 1987-2010 from Herzfeld; Morningstar from 2011 (unleveled general equity only) [6].

NYSE share turnover (TURN) is based on the ratio of notified share volume to average shares itemized from the "NYSE Fact Book". Baker and Wurgler have defined the "TURN as the natural log of the raw turnover ratio, detrended by the 5-year moving average" [6].

Authors Baker and Wurgler take the "number of IPOs (NIPO) and the average first-day returns (RIPO) from Jay Ritter's website" [6].

The *equity share (S)* is obtained by utilizing data from the "Federal Reserve Bulletin" [6].

The sixth and final proxy is *Dividend Premium (P^{D-ND})* which is also defined by Baker and Wurgler (2004). According to their study, they have found a connection between the closed-end fund discounts and the dividend premium. The values differ slightly due to subsequent improvements in the CRSP/ Compustat merge procedure. [6].

The authors state that any single proxy will be defective and noisy for Sentiment. Thus, they proposed a useful approach which uses to average diversified proxies and employ the mutual component as a sentiment measure [78].

Each of these sentiment proxies involves a sentiment component together with non-sentiment-related components. Therefore, the main sentiment index is constituted using fundamental component analysis to insulate the common components. That is, this method separates the first fundamental component of the six sentiment proxies and removes the mutual properties into a centring index. By extracting the mutual components of the sentiment proxies, it represents a smoother measure of sentiment

[6, 8, 78].

There is also a problem with the "relative timing of the variables" in constituting a sentiment index. In other words, if some variables display "lead-lag relationships", they might reflect a specific change in sentiment before the other variables. Firm supply response proxies such as S and NIPO usually have delays, when compared to investors' demand and behavior driven proxies such as TURN, RIPO, P^{D-ND} and CEFD [6].

Then, they construct a compound index which catches the mutual component in those six proxies and comprises for the fact that several variables take more time to find out the same sentiment. For that reason, they begin by "estimating the first principal component of the six proxies and their lags". Then, they have attained a first-stage index with 12 variables, for both the present and lagged proxies. Afterwards, they have estimated the correlation between the first-stage index and both the present and lagged values of each of the proxies. Lastly, the authors defined *Sentiment* as the "first principal component of the correlation matrix of six variables-each respective proxy's lead or lag, whichever has higher correlation with the first-stage index". The index has unit variance, since the coefficients have been rescaled [6]:

$$\begin{aligned} Sentiment_t = \; & -0.241CEFD_t + 0.242TURN_{t-1} + 0.253NIPO_t \\ & +0.257RIPO_{t-1} + 0.112S_t - 0.283P_{t-1}^{D-ND}. \end{aligned} \tag{5.1}$$

The authors have standardized each of these index components. According to the first principal component analysis, the sample variance is 49$\%$; thus, they concluded that one component reflects most of the joint variation. The correlation is 0.95 between the 12-terms first-stage index and the Sentiment index [6].

In the model of the Sentiment Index given in Equation (5.1), each individual proxy's sign is as expected. Variables positively associated with equity share in new issues, share turnover, the IPO volume and first-day returns, and negatively related with the dividend premium and the closed-end fund discount. Furthermore, all proxies' timing is also as expected except of CEFD, since investor behavior and price variables might bring about firm supply variables [6, 8].

5.4 Application

We use the monthly CPI dataset from Bloomberg website [1]. For the Sentiment Index, we also include the monthly data and the latest updated version, which is on March 31, 2016. With this update, NYSE turnover has been removed as one of the six sentiment indicators. Therefore, from now on, the Sentiment Index is based on five proxies. This dataset is taken from the Baker and Wurgler's web page [6, 7] The applications and results are provided by *MARS Salford Systems* [2].

As we mentioned several times, we want to show the effects of financial process and Sentiment of traders onto each other in our thesis. We also include *time* as a component. Hereby, we consider the most affected time interval by the Sentiment factor so that we can see the relation between financial process and Sentiment explicitly. For this purpose, according to the economic history of the USA, we decide to take our time interval between 2001-2012. In this time interval, we see train and test results semiannually. To be more precise, the train period begins with the second half of 2001 and continues with the test period in the first half of 2002. Likewise, train period goes on with the second half of 2002 and test period with the first half of 2003. We continue to see the train period for the second half of 2003 and test period for the first half of 2004. In the same manner, train period goes on with the second half of 2004 and test period with the first half of 2005. This process continues on an ongoing basis until the end of 2012. Lastly, we see the train period in the second half of 2011 and test period in the first half of 2012.

We train for 6 months and test for 6 months. We indicate the train and test periods together for each 6 months on separate graphics. In other words, we obtain a graph which shows the train period for the second half of 2001 and the test period for the first half of 2002. Similarly, we attain separate graphs for each of the other train-test periods also. As a result, we acquire 11 time periods together with train and test values. Thus, we have 11 graphics for both CPI and Sentiment. By using these values of the time periods, we construct a MARS model with 60 data points.

For simplicity, we denote Investor Sentiment as *S*, CPI as *CPI* and time as *T*.

At first, to see the total effect of Investor Sentiment, CPI, and time to CPI, we obtain the following MARS basis functions:

$$
\begin{aligned}
&BF_1 = \max\{0, CPI_{t-1} - 0.735672\},\\
&BF_2 = \max\{0, T_t - 2001\},\\
&BF_5 = \max\{0, S_{t-1} - 0.0262489\} \cdot BF_2,\\
&BF_7 = \max\{0, S_{t-1} + 0.220402\} \cdot BF_1.
\end{aligned}
$$

Then, the model has the form:

$$
\begin{aligned}
&CPI_t = 0.731639 + 0.725255 \cdot BF_1 + 0.0207088 \cdot BF_2 - 0.0266736 \cdot BF_5\\
&+ 0.218117 \cdot BF_7.
\end{aligned}
$$

Here, BF_1 is a standard BF and the reflected BF for the input variable CPI_{t-1} (CPI), BF_2 is also a standard BF and the reflected BF for the input variable T_t (time). However, if we examine BF_5 and BF_7, we see that BF_5 employs BF_2 to state the interaction between the input variables S_{t-1} (Sentiment) and T_1 (time). Similarly, BF_7 employs BF_1 to state the interaction between the S_{t-1} (Sentiment) and CPI_{t-1} (CPI).

For the first time period, we start with 2001. Here, we train for the second half of 2001 and test for the first half of 2002. That is, we train for 6 months and test for 6 months. Figure 5.1 shows the total impact on CPI and the predicted impact by MARS for both train and test periods.

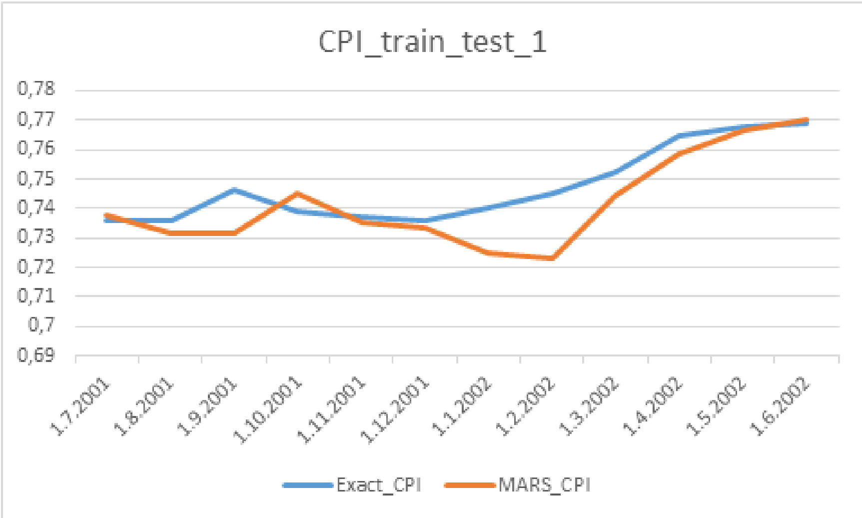

Figure 5.1: Modelling of CPI for the 2001-2002.

For the second graph, we train for the second half of 2002 and test for the first half of 2003. Figure 5.2 shows the total impact on CPI and the predicted impact through MARS for both train and test period. Our results show train period for 6 months and test period for 6 months.

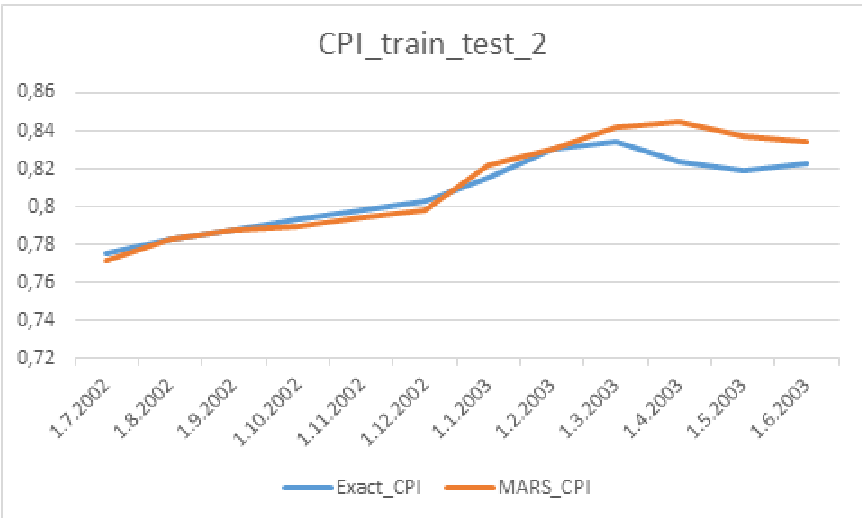

Figure 5.2: Modelling of CPI for the 2002-2003.

Figure 5.3 shows the total impact on CPI and the predicted impact through MARS for both train and test period. Here, the train period begins with the second half of 2003, lasts for 6 months and the test period begins with the first half of 2004, also lasts for 6 months.

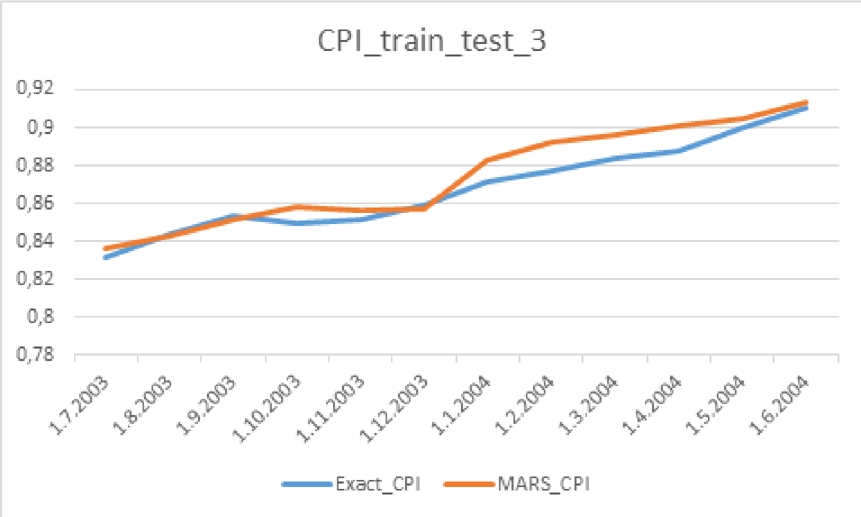

Figure 5.3: Modelling of CPI for the 2003-2004.

Figure 5.4 depicts the total impact on CPI and the predicted impact through MARS for the second half of 2004, which is the train period and shows for the first half of 2005, which is the test period.

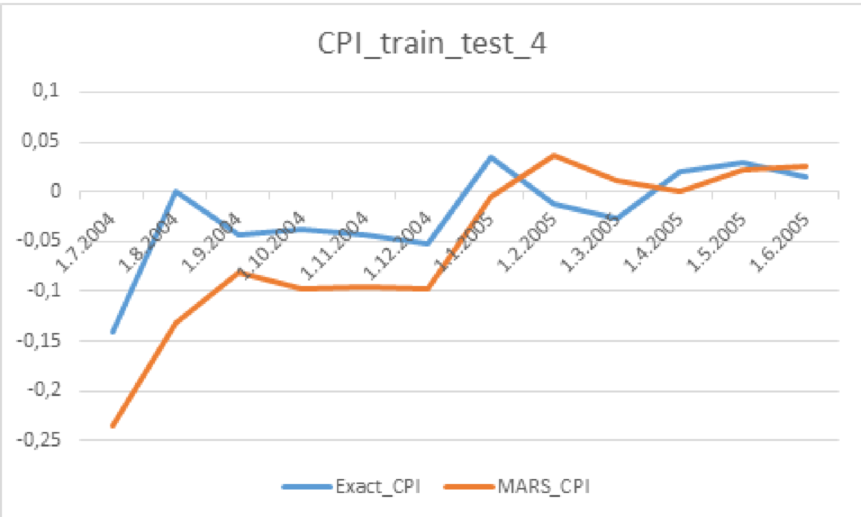

Figure 5.4: Modelling of CPI for the 2004-2005.

We see the train period for the second half of 2005 and test period for the first half of 2006 in Figure 5.5.

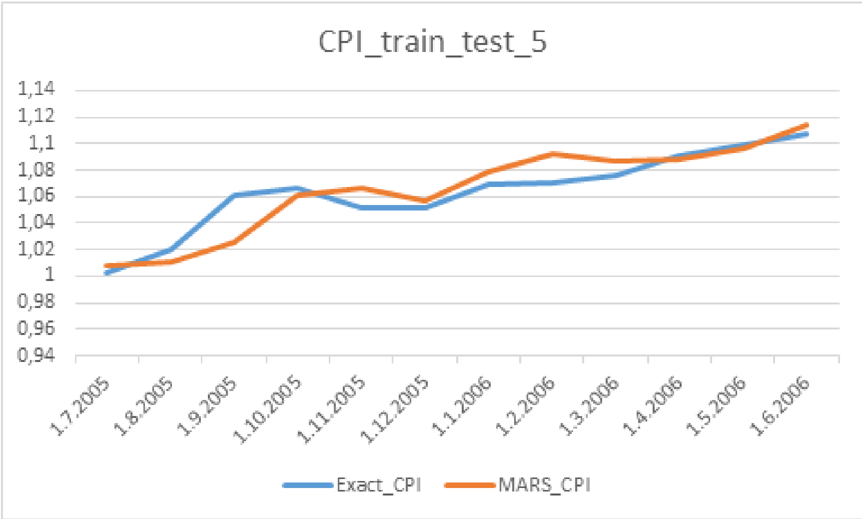

Figure 5.5: Modelling of CPI for the 2005-2006.

Figure 5.6 shows the total impact on CPI and the predicted impact through MARS for both the train and test periods. These periods consists of 6 months as other train and test periods too. Train period begins with the second half of 2006 and test period begins with the first half of 2007.

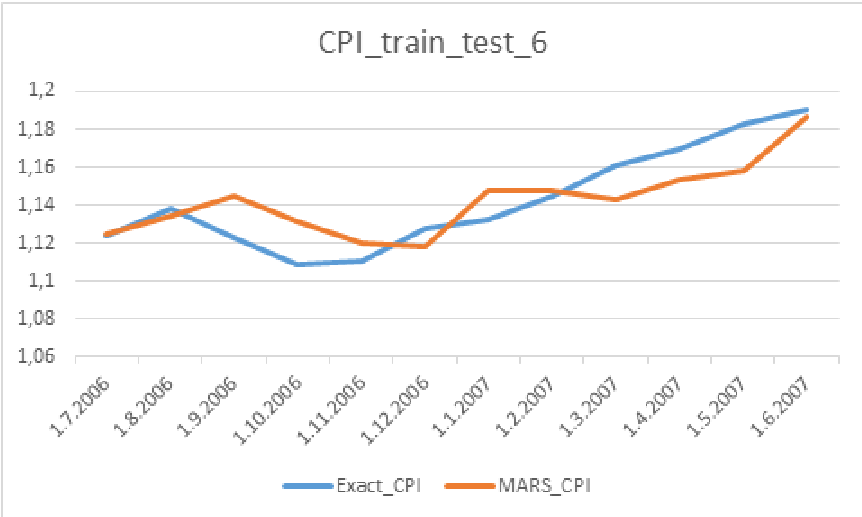

Figure 5.6: Modelling of CPI for the 2006-2007.

For the second half of 2007, we train our model and for the first half of 2008, we test our model. Figure 5.7 shows us the total impact on CPI and the predicted impact through MARS for both the train and test periods.

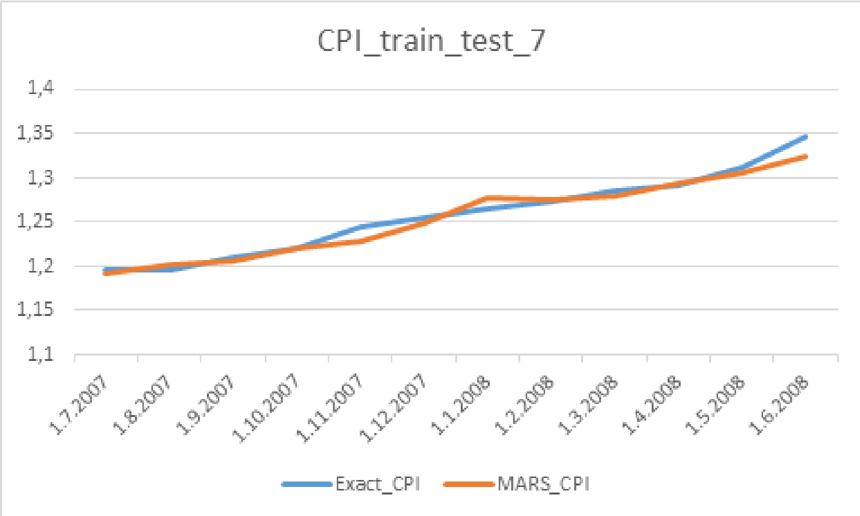

Figure 5.7: Modelling of CPI for the 2007-2008.

Figure 5.8 depicts the total impact on CPI and the predicted impact through MARS for both the train and test periods. Train period begins with the second half of 2008 and test period begins with the first half of 2009.

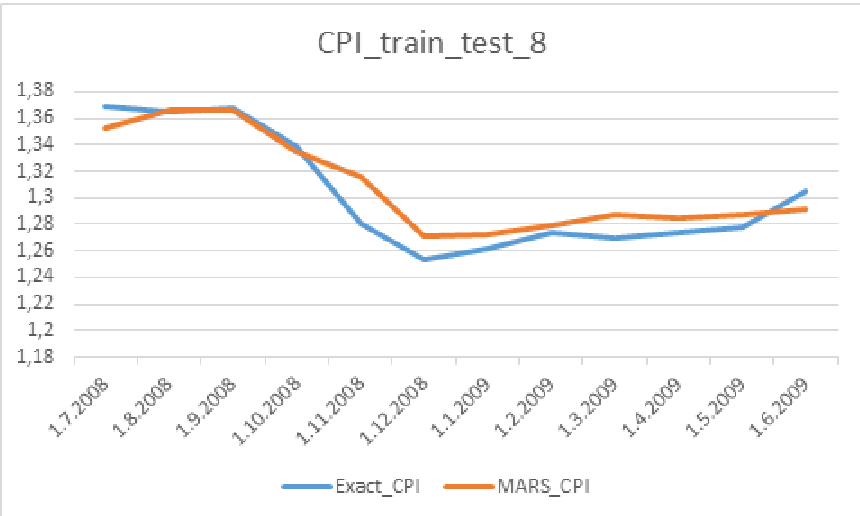

Figure 5.8: Modelling of CPI for the 2008-2009.

Figure 5.9 shows the total impact on CPI and the predicted impact through MARS for both the train and test periods. Train period begins with the second half of 2009 and test period begins with the first half of 2010.

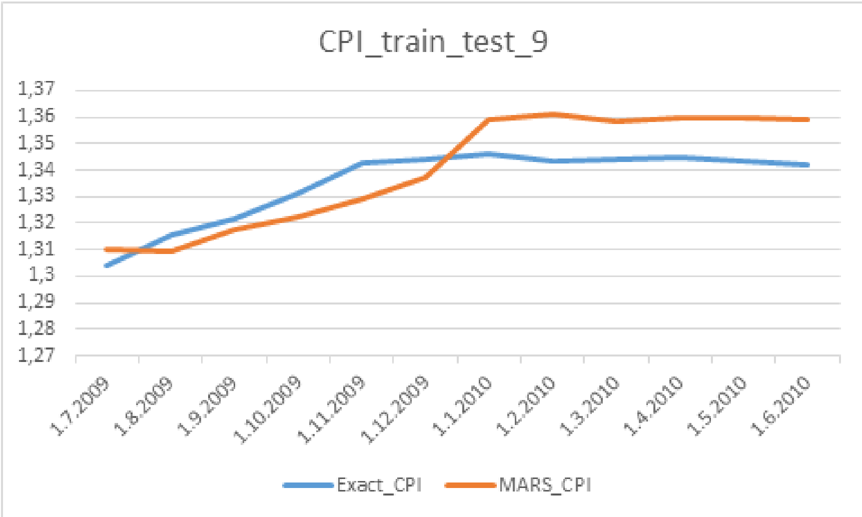

Figure 5.9: Modelling of CPI for the 2009-2010.

We see the train period for the second half of 2010 and test period for the first half of 2011 in Figure 5.10.

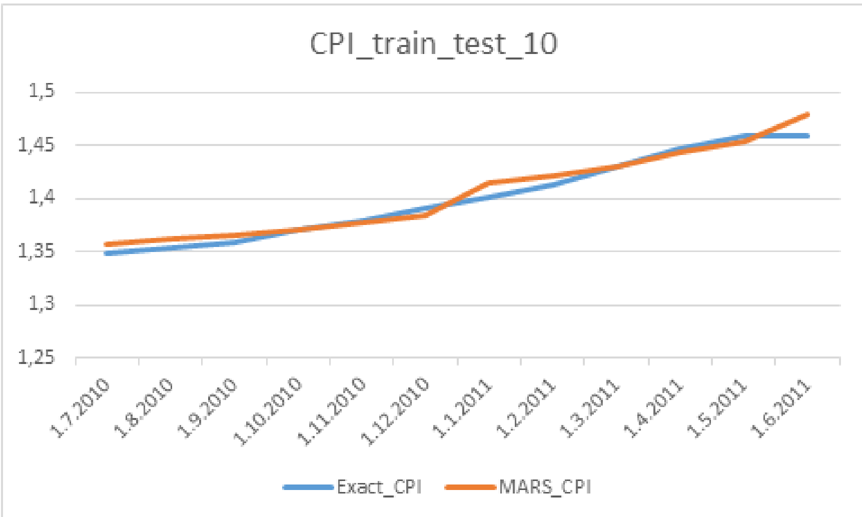

Figure 5.10: Modelling of CPI for the 2010-2011.

Finally, we see the train period for the second half of 2011 and test period for the first half of 2012 in Figure 5.11.

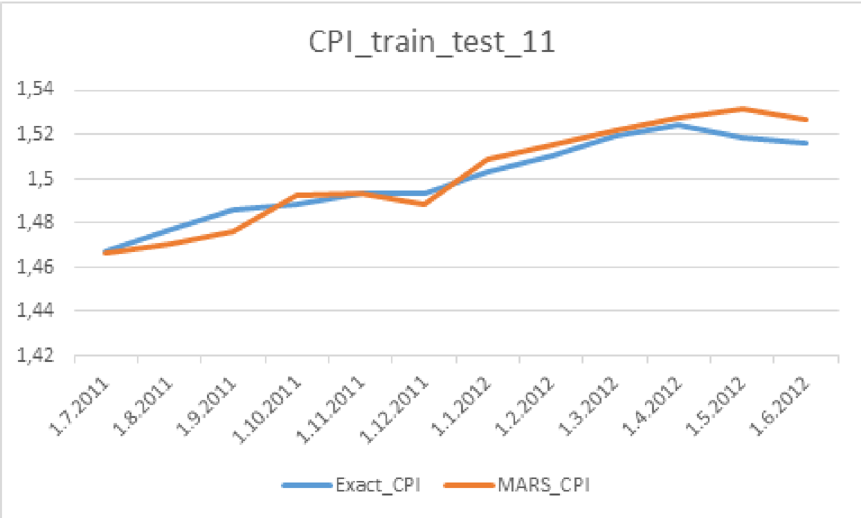

Figure 5.11: Modelling of CPI for the 2011-2012.

On the other hand, to display the total effect of Investor Sentiment, CPI, and time to Investor Sentiment Index, we obtain the following MARS basis functions:

$$\begin{aligned}
BF_1 &= \max\{0, S_{t-1} + 0.9321\},\\
BF_2 &= \max\{0, CPI_{t-1} - 0.915438\},\\
BF_3 &= \max\{0, 0.915438 - CPI_{t-1}\},\\
BF_4 &= \max\{0, S_{t-1} + 0.127611\} \cdot BF_3,\\
BF_5 &= \max\{0, -0.127611 - S_{t-1}\} \cdot BF_3,\\
BF_6 &= \max\{0, T_t - 2001\} \cdot BF_5,\\
BF_8 &= \max\{0, 2009 - T_t\} \cdot BF_2,\\
BF_9 &= \max\{0, T_t - 2001\} \cdot BF_1.
\end{aligned}$$

Then, the model has the form:

$S_t = -0.4157 - 1.27598 \cdot BF_2 + 5.81075 \cdot BF_4 - 3.95337 \cdot BF_6 + 0.250269 \cdot BF_8 + 0.119644 \cdot BF_9$.

Here, BF_1 is the standard BF and reflected BF for the input variable S_{t-1} (sentiment). Similarly, BF_2 and BF_3 are the standard BFs and reflected BFs for the input variable CPI_{t-1} (CPI), respectively. On the other hand, BF_4 and BF_5 involves BF_3 to indicate the interaction between the input variables S_{t-1} (Sentiment) and CPI_{t-1} (CPI). For BF_6, we can say that there is an interaction between the input variables S_{t-1} (Sentiment), CPI_{t-1} (CPI) and T_t (time).

For the first time period, the train period is the same as in the CPI; we train for the second half of 2001 and test for the first half of 2002. Figure 5.12 shows the total impact on sentiment and the predicted impact through MARS for both the train and test periods.

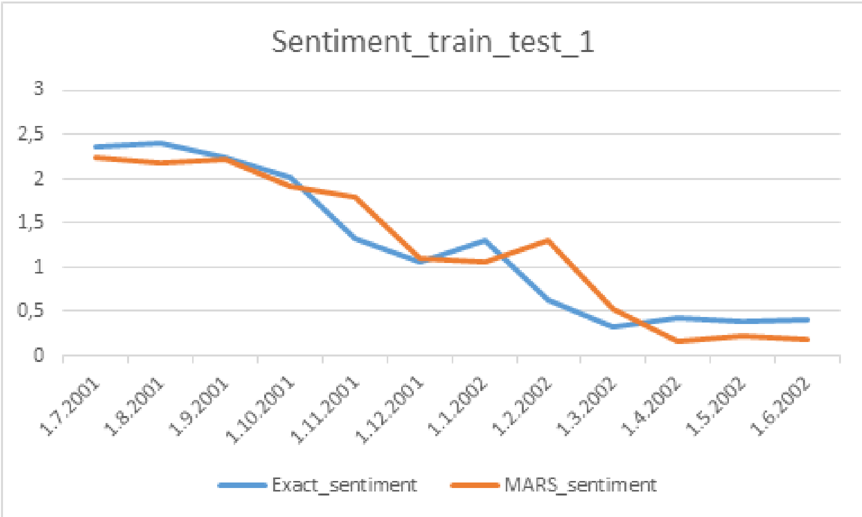

Figure 5.12: Modelling of Sentiment for the 2001-2002.

For the second graph, we train for the second half of 2002 and test for the first half of 2003. Figure 5.13 shows the total impact on Sentiment and the predicted impact through MARS for both train and test period. Our results show train period for 6 months and test period for 6 months.

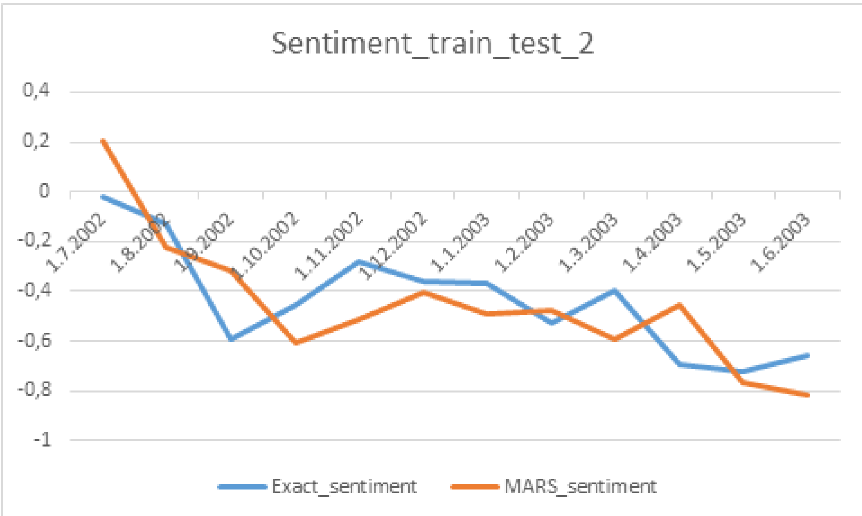

Figure 5.13: Modelling of Sentiment for the 2002-2003.

Figure 5.14 shows the total impact on Sentiment and the predicted impact through MARS for both train and test period. Here, the train period begins with the second half of 2003, lasts for 6 months and the test period begins with the first half of 2004, also lasts for 6 months.

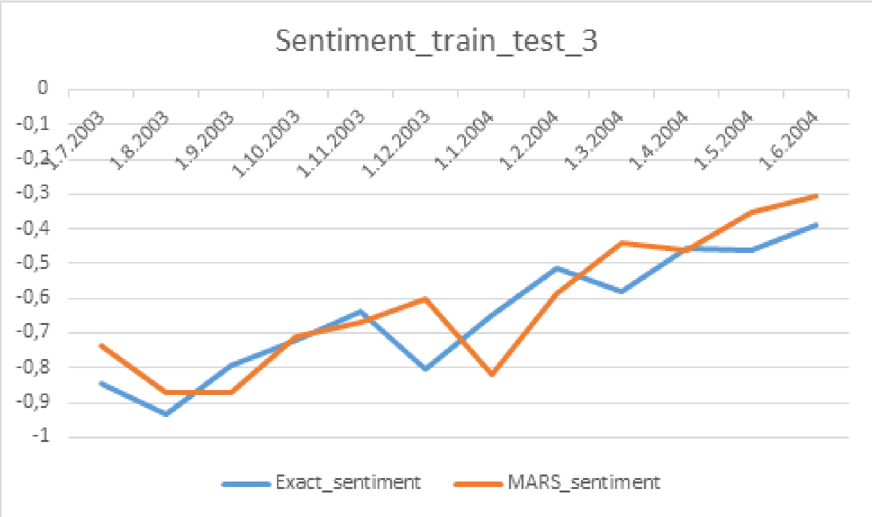

Figure 5.14: Modelling of Sentiment for the 2003-2004.

We see the train period for the second half of 2004 and test period for the first half of 2005 in Figure 5.15.

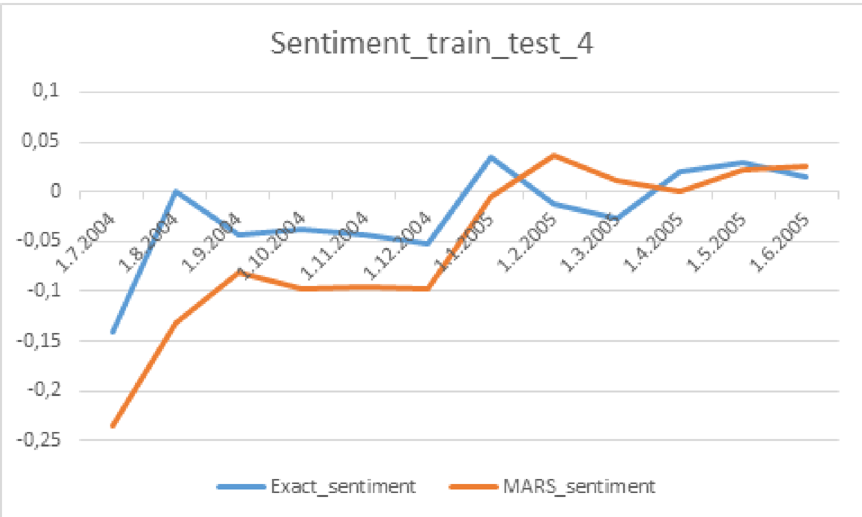

Figure 5.15: Modelling of Sentiment for the 2004-2005.

Figure 5.16 depicts the total impact on Sentiment and the predicted impact through MARS for both train and test period. Train period is for the second half of 2005 and test period for the first half of 2006.

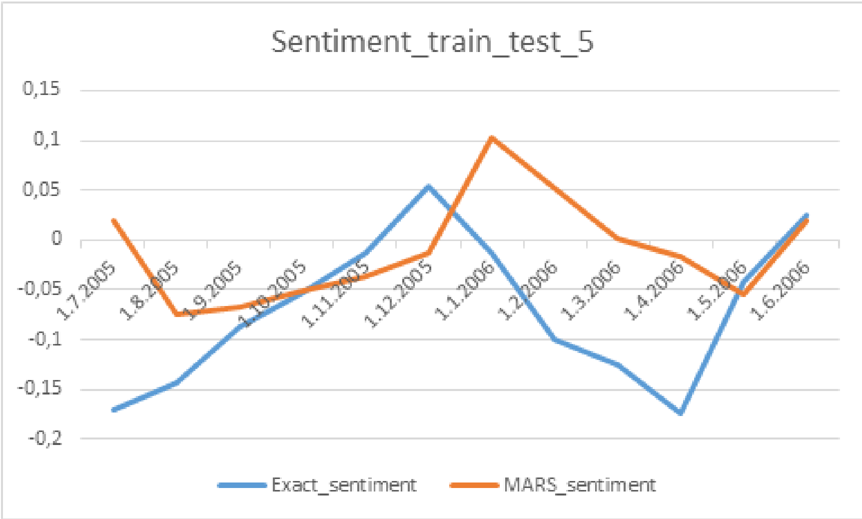

Figure 5.16: Modelling of Sentiment for the 2005-2006.

We see the train period for the second half of 2006 and test period for the first half of 2007 in Figure 5.17.

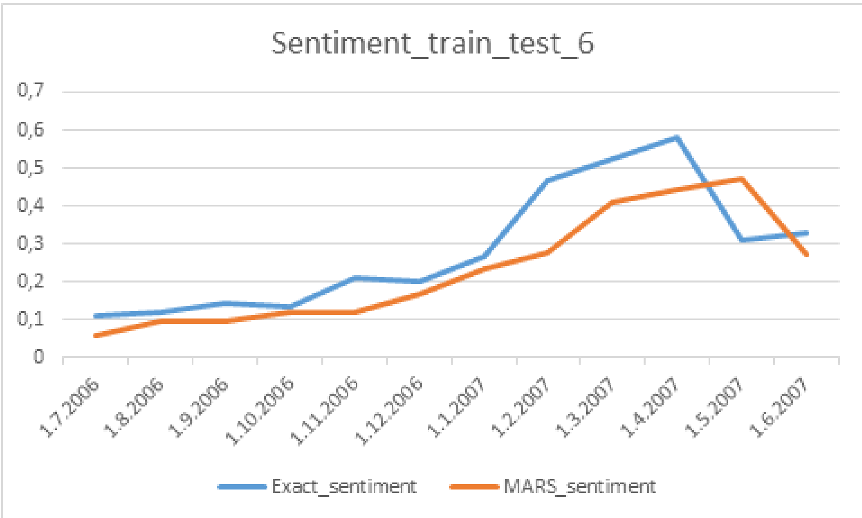

Figure 5.17: Modelling of Sentiment for the 2006-2007.

Figure 5.18 shows the total impact on Sentiment and the predicted impact through MARS for both train and test period. Here, the train period begins with the second half of 2007 and the test period begins with the first half of 2008.

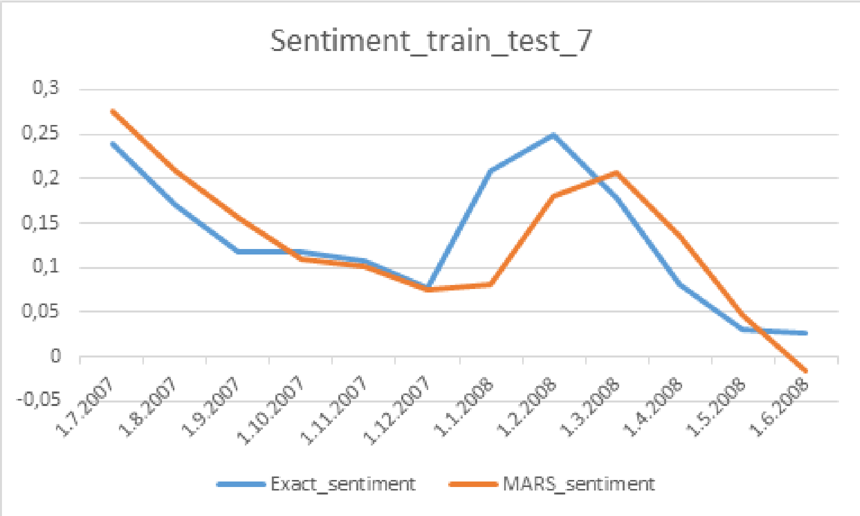

Figure 5.18: Modelling of Sentiment for the 2007-2008.

We see the train period for the second half of 2008 and test period for the first half of 2009 in Figure 5.19.

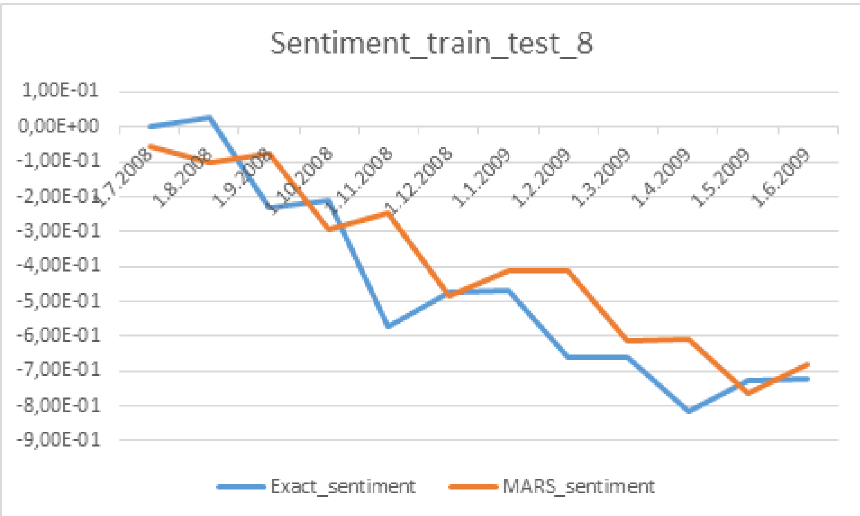

Figure 5.19: Modelling of Sentiment for the 2008-2009.

Figure 5.20 depicts the total impact on Sentiment and the predicted impact through MARS for both train and test period. Here, the train period begins with the second half of 2009 and the test period begins with the first half of 2010.

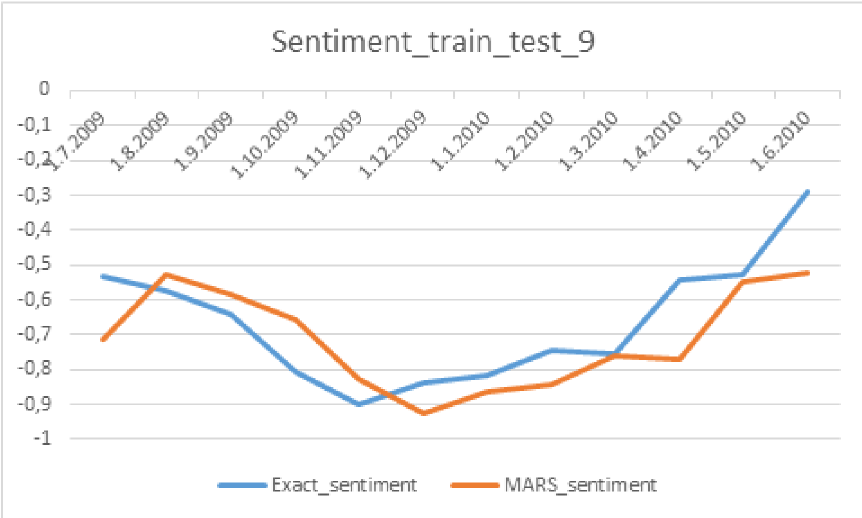

Figure 5.20: Modelling of Sentiment for the 2009-2010.

We see the train period for the second half of 2010 and test period for the first half of 2011 in Figure 5.21.

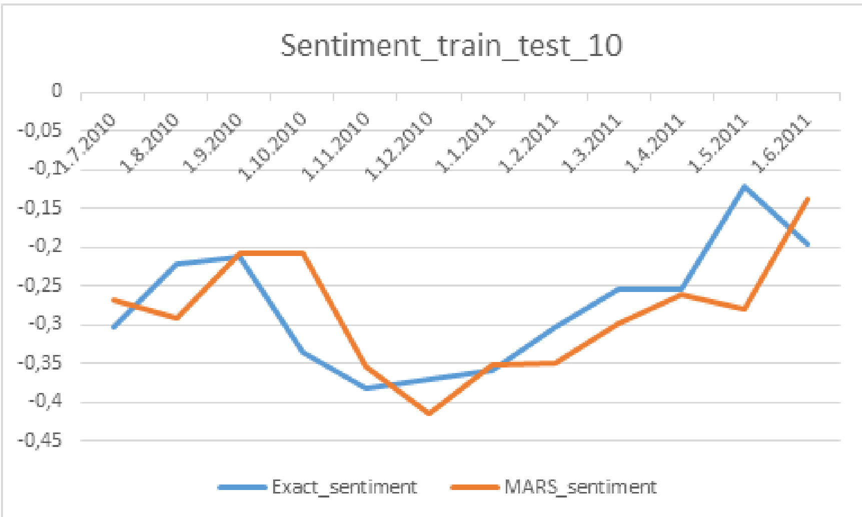

Figure 5.21: Modelling of Sentiment for the 2010-2011.

Finally, we see the train period for the second half of 2011 and test period for the first half of 2012 in Figure 5.22.

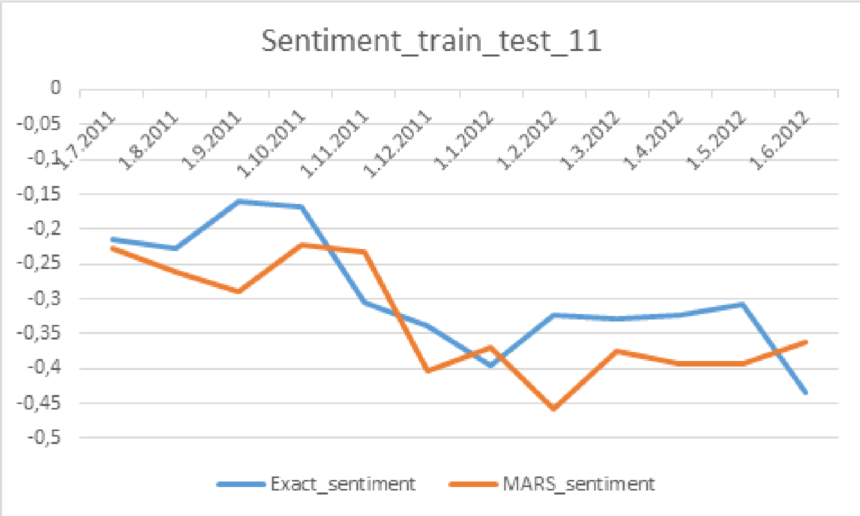

Figure 5.22: Modelling of Sentiment for the 2011-2012.

After we built our model, in Table 5.1 we obtain the statistical measures of the train and test results for CPI and Sentiment as well.

Table 5.1: Evaluation of the train and test results for both CPI and Sentiment with well-known statistical measures.

	CPI		Sentiment	
measures	train	test	train	test
R^2_{adj}	0,998	0,997	0,9668	0,8658
AAE	0,007	0,011	0,0865	0,1055
$RMSE$	0,01	0,013	0,1205	0,1468
r	0,999	0,999	0,9853	0,9429

If we look at our results for CPI, we see that the R^2_{adj} results are very close to each other for the train and test period and they are very close to 1, also for the r values we see that both the train and test results are identical. Besides, they are approximately equal to 1. For the AAE and $RMSE$ measures, we see that they are close to each other, and they are close to 0. According to all these measures, our MARS prediction reveals a very good performance.

Conversely, if we look at the measures for Sentiment, results are not as close to each other as in CPI. However, R^2_{adj} and r values are still nearby to 1 which is good. Besides, the error values AAE and $RMSE$ also close to each other.

According to all the results that we have obtained from the BFs, train-test graphs and statistical measures, we can make an inference that there is a strong relationship between the CPI, sentiment and time. For the time component, we received from our BFs that two years were mostly effective, especially, for the Sentimental sense, which are 2001 and 2009. In 2001, one of the biggest terrorist attacks occurred in the USA on the 11th of September. All the Americans were shocked, as a result, this situation affected all the people's sentiments very strongly, as we see from our graphs either. In fact, the main reason why we begin from the second half of 2001 is to see the impact of sentiment in the train period, explicitly. Also, in 2008 there was the so called *Great Recession*, which also affected the US and worldwide Sentiment levels, as we can observe from the graphs and the recession's reflection in 2009.

Remark. In order to understand the impact of time, first we obtained the statistical measures for this model *without* including the time component as another model variable. In other words, we applied MARS on our dynamics as if it was an autonomous system. There are no huge differences when time is included or omitted from the model. However, by adding time as a component, it allows us to represent the exact relation between sentiment, financial process, and time. In fact, we noticed that time is remarkably effective for the sentiment model. This is the objective why we constructed a coupled system of nonautonomous stochastic differential equations.

CHAPTER 6

CONCLUSION AND OUTLOOK

We know that SDEs are broadly used for applications in many areas. In our study, we concentrate on financial and "human factor" areas in sense of mutual interaction. Since processes, which belong to these fields, are exposed to random effects, we represent these processes with the SDEs. Especially in Financial Mathematics, there are numerous applications of SDEs such as in pricing of options or modelling of financial asset prices, interest rates, volatility and wealth.

Some of the researchers and expert traders are interested in finding out what kind of factors affect investors' decisions and the aspects of their investments. People sometimes make systematic errors in the way that they think: they are overconfident, they care a lot about the previous and recent experiences, etc. Behavioral Finance uses this body of knowledge and models in which some traders are not fully rational, because of their preferences or because of their mistaken beliefs [50].

As financial economists become used to consider the role of human behavior in driving stock prices, new terms have arisen to define these behaviors of investors. For the general case of various emotions of the investors expressed by the "Investor Sentiment" term [66].

In this thesis, the effect of financial processes and Sentiment of investors onto each other has been described by the SDEs as a coupled system. As a result, we obtain two-dimensional SDEs for assessing these two coupled differential equations. Afterwards, a parameter estimation procedure for the system of SDEs is introduced and carried out. This procedure is based on both discretization of the SDEs by using Milstein approximation and a refined application of MARS algorithm. MARS gives the opportunity for statistical inference of stochastic process parameters with the help of rather elementary basis functions [75].

In order to see the mutual influence of financial processes and Sentiment of traders on each other, we wanted to make an application of those via real-world datasets. These datasets from the financial sector and human factor are characterized by their massive quantity and fluctuation. By using MARS algorithm, we aimed to diminish this fluctuation as much as possible and to provide a smoother and more stable approximation for the data.

As far as it is derived from our research and its results, we can say that this thesis gives a new contribution for the representation of the coupled effects of financial processes and human behavior-sentiment- to each other, by constructing a coupled system of SDEs. In addition, to see the coupling effect of financial processes and Sentiment of investors, we also add *time* as a third independent variable. That is, we wanted to observe and investigate the influence of some financial process and Investor Sentiment including time to each other and, as a result, we constructed a coupled system. Furthermore, this thesis also gives a contribution by assessing the "human factor" by MARS algorithm, mathematical discretization and optimization, and the real-world application side.

In our thesis, we have used monthly dataset for two indexes. To see the relation between financial process and Investor Sentiment explicitly, we trained for half a year and test for the rest of the year. The reason for this procedure is, Sentiment can be more effective in short terms. As time passes, the degree of Sentiment will diminish. Therefore, our future works can contain a daily dataset to obtain a much more explicit strong relation between Investor Sentiment, financial process and time.

Apart from Behavioral Finance, we have made a research also about Neurofinance. This area investigates how peoples' brain activity and hormonal levels are shifting with respect to their financial decisions. According to many studies conducted in this area, numerous interesting results have already been obtained. In these studies, researchers generally used fMRI tools to measure subject's brain activities. Since we are not able to provide that equipment yet, we keep this desire for our future study.

In this thesis, to avoid from complexity, we have taken one Brownian motion for each SDE to increase the stability of our model. In future work, we may improve our calculations and construct a system of SDEs with mixed terms (b^{ij}) $(i \neq j)$ with m Brownian motions instead of two Brownian motions as we implied in Chapter 4. In addition to this extension, we can also construct a system of k SDEs with $k \geq 3$, in future investigations and applications.

We have used MARS method for the parameter estimation of our system of SDEs, and we might also employ Conic Multivariate Adaptive Regression Splines (CMARS), Robust Conic Multivariate Adaptive Regression Splines (RCMARS), Conic Generalized Partial Linear Models (CGPLM), and Robust Conic Generalized Partial Linear Models (RCGPLM) in our future studies. Those new versions of MARS allow for a stronger involvement of model-based, mathematical elements of methodology, and to robustify the model, allowing for a variance reduction [46, 74, 75, 80, 81]. Furthermore, we can use Neural Network to compare our results that we obtain from MARS and those versions of MARS.

We can also study on Stochastic Optimal Control subject to our systems of SDEs for a decision making that further involves the human factor, in our further studies.

Furthermore, we can continue our research going toward a "verification" and "falsification" of neuro financial, behavioral and sentiment-based theories. Eventually, we can examine the elements of the human factor such as empathy, love, perseverance, commitment, sympathy and impulse control. We can inquire on questions like, in which economic situations these elements of human factor occur, rise or decline, especially, what causes these emotions to reveal and disappear.

REFERENCES

[1] Bloomberg, https://www.bloomberg.com/quote/ CPURNSA:IND.

[2] MARS from Salford Systems, http://salfordsystems/products/mars/ (accessed 22 June 2017).

[3] R. Aster, B. Borchers, and C. Thurber, Parameter estimation and inverse problems, International Geophysics, 90, pp. 1–296, 2005.

[4] R. J. Aumann, Rationality and bounded rationality, Games and Economic behavior, 21(1-2), pp. 2–14, 1997.

[5] L. Bachelier, *Louis Bachelier's theory of speculation: the origins of modern finance*, Princeton University Press, 2011.

[6] M. Baker and J. Wurgler, Investor sentiment and the cross-section of stock returns, The Journal of Finance, 61(4), pp. 1645–1680, 2006.

[7] M. Baker and J. Wurgler, Investor sentiment and the cross-section of stock returns, http://pages.stern.nyu.edu/ jwurgler/data/, 2006.

[8] M. Baker and J. Wurgler, Investor sentiment in the stock market, The Journal of Economic Perspectives, 21(2), pp. 129–151, 2007.

[9] N. Barberis and R. Thaler, A survey of behavioral finance, Handbook of the Economics of Finance, 1, pp. 1053–1128, 2003.

[10] D. Bathia and D. Bredin, An examination of investor sentiment effect on G7 stock market returns, The European Journal of Finance, 19(9), pp. 909–937, 2013.

[11] J. P. Bishwal, *Parameter estimation in stochastic differential equations*, Springer, 2008.

[12] S.-K. Bormann, Sentiment indices on financial markets: What do they measure?, Technical report, Economics Discussion Papers, 2013.

[13] A. Çevik, G.-W. Weber, B. M. Eyüboğlu, K. K. Oğuz, A. D. N. Initiative, et al., Voxel-mars: a method for early detection of alzheimer's disease by classification of structural brain mri, Annals of Operations Research, pp. 1–27, 2017.

[14] M. Cícha, Czech yield curve, Research Centre for Competiveness of the Czech Economy, Working Paper, (11), 2010.

[15] L. Cobb, Stochastic differential equations for the social sciences, Mathematical frontiers of the social and policy sciences, (54), p. 37, 1981.

[16] D. Colander, *The complexity vision and the teaching of economics*, Edward Elgar Publishing, 2000.

[17] J. B. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann, Noise trader risk in financial markets, Journal of Political Economy, 98(4), pp. 703–738, 1990.

[18] K. L. Fisher and M. Statman, The sentiment of investors, large and small, Leavey School Faculty, Santa Clara University, 99, pp. 00–11, 1999.

[19] J. Fox, *Applied regression analysis, linear models, and related methods*, Sage Publications, Inc, 1997.

[20] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, volume 1, Springer series in statistics Springer, Berlin, 2001.

[21] C. Frydman, What drives peer effects in financial decision-making? neural and behavioral evidence, Technical report, Working paper, University of South California, 2015.

[22] C. Frydman and C. F. Camerer, The psychology and neuroscience of financial decision making, Trends in Cognitive Sciences, 20(9), pp. 661–675, 2016.

[23] C. D. Frydman, *Essays in neurofinance*, Ph.D. thesis, California Institute of Technology, Pasadena, California, 2012.

[24] Z. Gao, H. Ren, and B. Zhang, Googling investor sentiment around the world, Technical report, Working Paper, Chinese University of Hong Kong, 2016.

[25] J. Gonzales, *Simulation and Parametric Estimation of SDEs*, School of Mathematics, The University of Manchester, Magical Books Project, August 23, 2012.

[26] J. Hengelbrock, E. Theissen, and C. Westheide, Market response to investor sentiment, Journal of Business Finance & Accounting, 40(7-8), pp. 901–917, 2013.

[27] D. Huang, F. Jiang, J. Tu, and G. Zhou, Investor sentiment aligned: A powerful predictor of stock returns, The Review of Financial Studies, 28(3), pp. 791–837, 2015.

[28] T. P. Hubbard and Y. Saglam, *Stochastic Processes, Ito Calculus, and Applications in Economics*, Department of Economics, University of Iowa, March 03, 2006.

[29] S. M. Iacus, *Simulation and inference for stochastic differential equations: with R examples*, Springer Science & Business Media, 2009.

[30] C. P. Index, US Bureau of Labor Statistics, Washington, DC, 2011.

[31] M. Jeanblanc, M. Yor, and M. Chesney, *Mathematical methods for financial markets*, Springer Science & Business Media, 2009.

[32] A. Karimov, *Modelling Illiqudity Premium and Bid-Ask Prices of the Financial Securities*, PhD. thesis at IAM, METU, Ankara, Turkey, December, 2016.

[33] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments*, Springer Science & Business Media, 2012.

[34] C. M. Kuhnen and B. Knutson, The neural basis of financial risk taking, Neuron, 47(5), pp. 763–770, 2005.

[35] E. Kürüm, G.-W. Weber, and C. Iyigun, Early warning on stock market bubbles via methods of optimization, clustering and inverse problems, Annals of Operations Research, 2014.

[36] D. Lamberton and B. Lapeyre, *Introduction to stochastic calculus applied to finance*, CRC press, 2007.

[37] T.-S. Lee, C.-C. Chiu, Y.-C. Chou, and C.-J. Lu, Mining the customer credit using classification and regression tree and multivariate adaptive regression splines, Computational Statistics & Data Analysis, 50(4), pp. 1113–1130, 2006.

[38] Y.-C. Lei, *The trading volume trend, investor sentiment, and stock returns*, Ph.D. thesis, Louisiana State University, 2005.

[39] A. W. Lo, D. V. Repin, and B. N. Steenbarger, Fear and greed in financial markets: A clinical study of day-traders, Technical report, National Bureau of Economic Research, 2005.

[40] S. M. Lyons, Introduction to stochastic differential equations, School of Informatics, University of Edinburgh, Scotland, 11, p. 20, 2013.

[41] D. Montgomery, E. Peck, and G. Vining, Introduction to linear regression analysis, Wiley series in probability and statistics, 2006.

[42] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*, John Wiley & Sons, 2015.

[43] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*, John Wiley & Sons, 2010.

[44] N. M. N. Muhammad et al., Behavioural finance vs traditional finance, Adv. Manage, 2(6), pp. 1–10, 2009.

[45] B. Øksendal, Stochastic differential equations, in *Stochastic differential equations*, pp. 65–84, Springer, 2003.

[46] A. Özmen, *Robust Optimization of Spline Models and Complex Regulatory Networks*, Springer, 2016.

[47] H. A. Rahman, A. Bahar, M. K. Bazli, N. Rosli, M. Salleh, and G.-W. Weber, Levenberg-Marquardt parameter estimation of stochastic logistic model, MATEMATIKA UTM (Mathematics Journal of Universiti Teknologi Malaysia), 25(2), pp. 91–106, 2009.

[48] H. A. Rahman, A. Bahar, N. Rosli, and M. M. Salleh, Parameter estimation of stochastic differential equation, Sains Malaysiana, 41(12), pp. 1635–1642, 2012.

[49] J. K. Rilling and A. G. Sanfey, The neuroscience of social decision-making, Annual review of psychology, 62, pp. 23–48, 2011.

[50] J. R. Ritter, Behavioral finance, Pacific-Basin Finance Journal, 11(4), pp. 429–437, 2003.

[51] M. Sadeghnia, A. H. Hooshmand, and H. Niko, Behavioral finance and neurofinance and research conducted in this area, Interdisciplinary Journal of Contemporary Research in Business, 4, 2013.

[52] S. Särkkä, *Applied Stochastic Differential Equations*, Department of Electrical Engineering and Automation (EEA), Aalto University, Finland, November 21, 2012.

[53] T. Sauer, Numerical solution of stochastic differential equations in finance, in *Handbook of computational finance*, pp. 529–550, Springer, 2012.

[54] M. Schmeling, Investor sentiment and stock returns: Some international evidence, Journal of empirical finance, 16(3), pp. 394–408, 2009.

[55] G. Schneider, P. F. Craigmile, and R. Herbei, Maximum likelihood estimation for stochastic differential equations using sequential kriging-based optimization, arXiv preprint arXiv:1408.2441, 2014.

[56] P. Sephton, Forecasting recessions: Can we do better on MARS, Review, 83, 2001.

[57] M. Z. Shariff, J. Al-Khasawneh, and M. Al-Mutawa, Risk and reward: A neurofinance perspective, International Review of Business Research Papers, 8(6), pp. 126–133, 2012.

[58] R. B. Silberstein and G. E. Nield, Brain activity correlates of consumer brand choice shift associated with television advertising, International Journal of Advertising, 27(3), pp. 359–380, 2008.

[59] L. Squire, D. Berg, F. E. Bloom, S. Du Lac, A. Ghosh, and N. C. Spitzer, *Fundamental neuroscience*, Academic Press, 2012.

[60] K. J. Stewart and S. B. Reed, Consumer price index research series using current methods, 1978-98, Monthly Lab. Rev., 122, p. 29, 1999.

[61] P. Taylan and G.-W. Weber, Organization in finance prepared by stochastic differential equations with additive and nonlinear models and continuous optimization, Organizacija, 41(5), pp. 185–193, 2008.

[62] P. Taylan, G.-W. Weber, and A. Beck, New approaches to regression by generalized additive models and continuous optimization for modern applications in finance, science and technology, Optimization, 56(5-6), pp. 675–698, 2007.

[63] P. Taylan, G.-W. Weber, and F. Yerlikaya, Continuous optimization applied in mars for modern applications in finance, science and technology, in *ISI Proceedings of 20th Mini-EURO Conference Continuous Optimization and Knowledge-Based Technologies*, pp. 317–322, 2008.

[64] K. H. Teka, *Parameter estimation of the Black-Scholes-Merton model*, Ph.D. thesis, Kansas State University, 2013.

[65] P. C. Tetlock, Giving content to investor sentiment: The role of media in the stock market, The Journal of Finance, 62(3), pp. 1139–1168, 2007.

[66] R. H. Thaler, The end of behavioral finance, Financial Analysts Journal, pp. 13–23, 2010.

[67] R. Tibshirani and L. Wasserman, *Nonparametric Regression*, Carnegie Mellon University, Department of Statistics, 2013.

[68] J. Timmer, Parameter estimation in nonlinear stochastic differential equations, Chaos, Solitons & Fractals, 11(15), pp. 2571–2578, 2000.

[69] K. Tseng, Behavioral finance, bounded rationality, neuro-finance, and traditional finance, Investment Management and Financial Innovations, 3(4), pp. 7–18, 2006.

[70] U. Uygur and O. Taş, The impacts of investor sentiment on different economic sectors: Evidence from istanbul stock exchange, Borsa Istanbul Review, 14(4), pp. 236–241, 2014.

[71] D. Vasile, T. C. Sebastian, et al., Neurofinance - Getting an Insight into the Trader's Mind, Neuroscience, 27(31), pp. 8159–8160, 2007.

[72] D. Wang, *Model selection and estimation in generalized additive models and generalized additive mixed models*, North Carolina State University, 2013.

[73] X. Wang, *Online Nonparametric Estimation of Stochastic Differential Equations*, Ph.D. thesis, The University of Western Ontario, 2015.

[74] G.-W. Weber, P. Taylan, K. Yıldırak, and Z.-K. Görgülü, Financial regression and organization, Special Issue on Optimization in Finance, DCDIS-B, 17(16), pp. 149–174, 2010.

[75] F. Yerlikaya-Özkurt, *Refinements, Extensions and Modern Applications of Conic Multivariate Regression Splines*, PhD. thesis at IAM, METU, Ankara, Turkey, December, 2013.

[76] Y. Yılmaz, *Daily Natural Gas Consumption Prediction by MARS and CMARS Models for Residential Users in Ankara*, MSC. thesis at IAM, METU, Ankara, Turkey, August, 2015.

[77] S. C. Yu, A. Reinstein, and N. T. Churyk, An overview of investor sentiment in stock market.

[78] C. Zhang, Defining, modeling, and measuring investor sentiment, University of California, Berkeley, Department of Economics, 2008.

[79] F. Özkurt Yerlikaya, I. Batmaz, and G. Weber, *A review of conic multivariate adaptive regression splines CMARS: A powerful tool for predictive data mining*, Preprint, IAM, METU, Ankara, Turkey, 2012-15.

[80] F. Özkurt Yerlikaya and G. Weber, Identification of stochastic differential equations by conic optimization of multivariate adaptive regression splines, Preprint 2013-20, IAM, METU, Ankara, Turkey, 2013.

[81] A. Özmen, *Robust Conic Quadratic Programming in Applied to Quality Improvement - A Robustification of CMARS*, MSC thesis at IAM, METU, Ankara, Turkey, September, 2010.

[82] M. Zouaoui, G. Nouyrigat, and F. Beer, How does investor sentiment affect stock market crises? Evidence from panel data, Financial Review, 46(4), pp. 723–747, 2011.