1-1-2012

Regime-Switching Behaviour In US Equity Indices: Two State Model With Kalman Filter Tracking And Finite State Machine Trading System

Timothy Little
Ryerson University

REGIME-SWITCHING BEHAVIOUR IN US EQUITY INDICES: TWO STATE MODEL WITH KALMAN FILTER TRACKING AND FINITE STATE MACHINE TRADING SYSTEM

by

TIMOTHY LITTLE
BASc University of Toronto (2003)

A thesis
presented to Ryerson University
in partial fulfillment of the
requirement for the degree of
Master of Applied Science
in the Program of
Electrical and Computer Engineering.

Toronto, Ontario, Canada, 2012

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

BORROWER'S PAGE

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Name	Signature	Address	Date

Regime-Switching Behaviour in US Equity Indices: Two State Model with Kalman Filter Tracking and Finite State Machine Trading System

Master of Applied Science 2012

TIMOTHY LITTLE

Electrical and Computer Engineering

Ryerson University

Abstract

This thesis presents a time varying regime-switching model for US equity index daily returns. The parameters of the model are estimated recursively with the Kalman filter. We demonstrate our model and parameter estimation technique are effective by demonstrating improvements in model fit compared to alternate models. Information from our model is used to build a Finite State Machine trading system with back-tested performance in excess of $15,000\%$ above a buy and hold strategy for the DOW Jones Industrial average from 1928-2012. Similar results are found for both the S&P 500 index and the NASDAQ Composite index over a long period. Our model succeeds at identifying profitable investment opportunities and improving model fit with a minimum of parameters.

Acknowledgements

I would like to thank my friends and family, my wife Margaret without whose support the completion of this thesis would not have been possible. My parent's for demonstrating the value of hard work and supporting my endeavours.

Thank you to my supervisor, Dr. Xiao-Ping Zhang, whose support and technical knowledge enabled me to develop an understanding of the research process.

I would also like to thank every member of my thesis defence committee. I appreciate their time, efforts, and contributions to this work.

Special thanks go to my colleagues in Communications and Signal Processing Applications Laboratory (CASPAL). It has been a pleasure working with all of you, best of luck in the future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Objectives

There is probably no other field of research that can more directly lead to financial gain than accurately predicting movements in the financial markets. The lure of financial success has drawn academics from a variety of fields to study the behaviour of asset prices, returns and risk characteristics.

Successful investing requires identifying opportunities to make positive expected value bets. A time series model that predicts future investment returns accurately is useful for identifying such opportunities. The model developed in this thesis makes one-step ahead forecasts of investment returns for several US equity indices. We show the model is effective for predicting future returns in general then apply the predictions of the model to investment decision making with a finite state machine (FSM) trading system. Our regime switching model is well represented by the FSM. State transitions result in transactions and outputs from our model are used to define the current state. The exceptional back-tested results of the finite state machine trading system strengthens our argument for a two-regime model and our Kalman filter estimation technique.

The model explored in this thesis is a simple non-linear model of equity index returns. By examining conditional return distributions, we observe an interesting asymmetry in US daily equity index returns. We find that the return distribution given the return of the previous day is positive is significantly different from the return distribution given the return of the previous day is negative. The size of our dataset gives strong statistical evidence. This result motivates the exploration of a regime-switching model.

As a visual demonstration of the asymmetry in returns, we plot the accumulated capital for three investment strategies. One strategy simply buys the DOW Jones Index and holds until the end of the period. The second strategy invests only when the return of the previous day is positive and invests in cash otherwise. The third strategy invests only when the return of the previous day is negative. Figure 1.1 compares the cumulative returns for the three strategies.

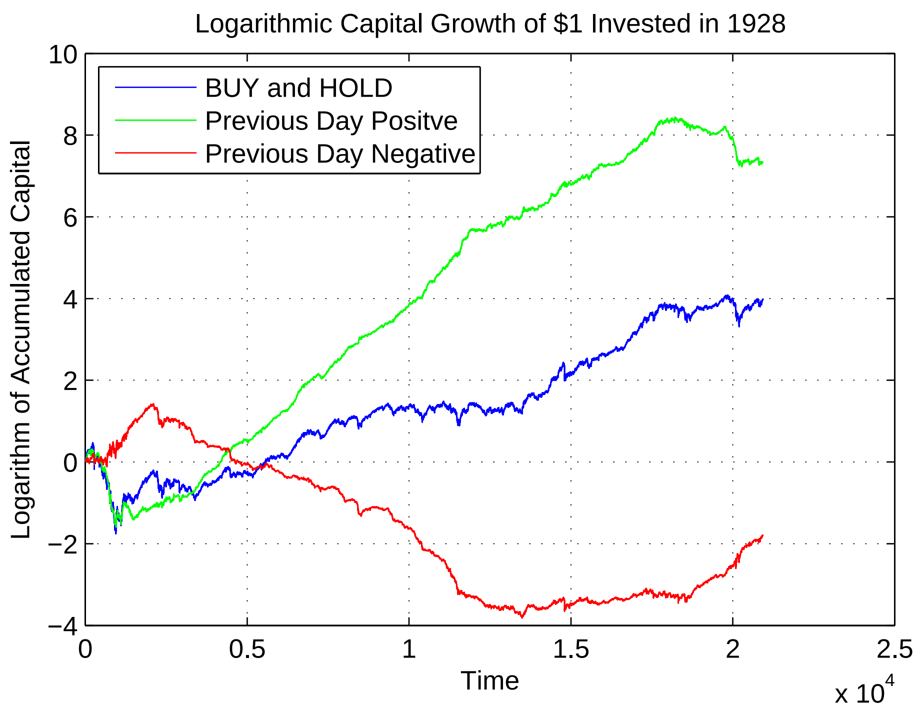

Figure 1.1: Capital Growth Over Time DOW Jones

From Figure 1.1, it is clear that vastly superior returns could have been achieved by selectively investing only when the return of the prior day is positive. This strategy provides much higher returns than a buy and hold strategy over the same period. Stocks are risky assets, by investing only when the previous return is positive, the investor is exposed to market risk only about half of the time of the buy and hold strategy while still providing higher returns. The remaining time the investor holds cash with zero return. Additional gains could be achieved by investing in short term Treasury bills when not invested in the market. The variance of returns for the strategy is lower than buy and hold. As a result, an investor achieves a higher return with lower risk by investing only when the previous day is up.

We use the observed mean asymmetry in returns to justify a two-state Model. We define a non-linear threshold function to control the state-switching based on the sign of the previous return. We then create a model using two regimes. Within each regime, the returns follow a random walk with drift but we allow the drift parameters to differ depending on the state. The parameters to be estimated are then the state dependent drift and variance.

Using empirical observations we suggest that the state dependent drift changes over time. We divide the data into smaller sub periods of non-overlapping 2000 point intervals and examine the dataset in smaller increments. We observe that the asymmetry is present in each sub period as established by hypothesis testing. We find that the state dependent expected return can vary significantly, in particular during periods of financial crisis.

In response to these observations we modify our original model with constant parameters and allow the drift parameters to vary over time. Prior studies have used autoregressive models for time varying returns in the class of regime-switching SETAR models [1]. We instead use a Kalman filter approach and allow the parameters to vary over time. The Kalman filter approach is preferable because it is adaptive, recursive and provides optimal estimates under the assumption of Gaussian noise.

By assuming the state dependent drift parameters themselves follow a random walk, the model can adapt to changing conditions. Conditions in financial markets can change dramatically in response to liquidity shocks, geopolitical events, natural and man made disasters among many other possible events. A model that can adapt to changing conditions is therefore well suited to financial applications.

The Kalman filter is a recursive algorithm. Each time new information becomes available it is used to produce a new estimate. When operating in the financial markets, new data becomes available every second of the day. By utilizing the Kalman filter we define a specific method for incorporating new data as it becomes available as required for realistic financial applications.

To estimate the parameters of our two-state over time, we apply the Kalman filter to track the time-varying parameters. From the parameter estimates we can find the expected return given the state of the system. We test our model and compare it to other candidate models. We demonstrate that our two-state model with Kalman filter tracking outperforms the other models considered.

Finally, we use our model to build a finite state trading machine. Trading in financial markets involves repeatedly identifying and profiting from positive expected returns. This requires a model to provide a prediction as well as a set of logical conditions describing trading actions. The parameter estimates from the Kalman filter are used to estimate the expected return given the state. The state machine uses the current state combined with the predictions of the model to make investing decisions. State transitions are triggered by new data becoming available and updated predictions from the Kalman filter estimation of state dependent expected returns. The trading system rotates between stocks and cash in an effort to maximize expected returns. When the expected return is positive, the system invests, when negative the system holds cash.

Running the trading system on three major US equity indices shows exceptional performance compared to buy and hold over the same period. Given the common belief that market timing is futile, the results demonstrate evidence to the contrary.

1.2 Literature Review

The model developed in this thesis is a modified version of the threshold autoregressive (AR) model [2] (TAR) introduced by Tong. The TAR model groups a set of time series observations into regimes using a threshold function. Within each regime, the series follows a different model. The parameters of an AR model can then be different for the different regimes. This approach provides additional flexibility for modelling non-linear phenomena and was originally used to model sunspot numbers and Canadian lynx data, two non-linear datasets.

The self-excited threshold autoregressive model (SETAR) was explored by Tong [3]. In the SETAR model the regime switching is controlled by a piecewise function of lagged values of the time series itself. The model developed in this thesis similarly uses a self-excited threshold function. Hanson explores the topic of threshold selection and model specification in [4]. Hansen [5] looks at the topic of inference in TAR models and testing for threshold co-integration in vector error correction models [6].

Various refinements and extensions of the basic SETAR model have also been studied. Zakoian [7] created the Threshold GARCH by incorporating GARCH variance specifications in the threshold model. Li and Lam [8] modelled mean and volatility asymmetry together in a stochastic volatility model creating the threshold stochastic volatility model. Regime-switching GARCH model are used for short term load forecasting for power applications in [9].

Asymmetry in stock returns was studied in Lam [10] and Li and Lam [11]. The results in [11] show that the conditional mean could depend significantly on the rise and fall of the market on the previous day. We studied this finding for the DOW Jones Industrial Average over a long period of time and came to the same conclusion. Numerous studies have attempted to establish the profitability of certain trading rules for a given dataset. Laurent demonstrated the predictive power candlestick chart patterns [12]. Brock et al.[13] demonstrated the effectiveness of a moving average rule based trading system for both in and out-of-sample testing. Andrew Lo attempts to answer a more general question regarding the conditional return distributions dependent on detection of a predefined chart pattern. He uses hypothesis testing to present statistical evidence that price patterns have some information content [14].

What technical analysis studies often lack is a model of returns that explains any positive results if found. Ready [15] questions the results of previous studies and the effectiveness of technical trading rules arguing that any positive results are likely the result of data snooping.

Chong and Lam [1] test a trading system based on a SETAR model for equity returns in the US markets demonstrating strong performance. Similarly, we develop a trading system based on a model or returns. This distinguishes our method from many studies of trading system performance because the trading actions we take follow the predictions of the model.

Time-varying parameters have been investigated for the Capital Asset Pricing Model [16] (CAPM) in [17]. In [18], Kim and Zumwalt analysed the variation of securities and portfolio returns in up and down markets, suggesting two states with different beta coefficients. In [19], the authors use Kalman filtering techniques to estimate time varying parameters for the CAPM model. Maximum likelihood estimation for stationary autoregressive moving average processes is discussed in [20]. [21] discuss maximum likelihood estimation of process noise variance in dynamic systems. In [22], the author investigates a piecewise constant model for beta using the Kalman filter to track beta over time.

Finite state machines for gambling decision making are developed by Feder [23]. The decision making process involves determining the optimal bet sizing given the state of a system. When applied to financial transactions, this process can be used to determine optimal use of leverage dynamically altering the position sizing in response to the predicted mean and variance of returns. Kelly [24] uses information theory to find the optimal bet size under uncertainty by viewing the gambling problem as communication over a noisy channel.

Summers [25] reasons that markets can be destabilized in the short term due to positive feedback trading. Positive feedback trading involves buying when prices rise and selling when they fall, exaggerating price swings. In Summers' model, rational investors buy when prices are low and sell when they high having the effect of dampening price swings. Summers hypothesises that the interactions of these two groups of investors create short-term herd behaviour with longer term reversion to fundamental values. Perhaps positive feedback trading strategies are the reason for the observed high frequency momentum in this thesis.

1.3 Contributions

This thesis builds a simple version of the SETAR model for equity index returns. Autoregressive models explain the time variation of returns by fitting the series to a linear combination of lagged values of the series itself. This is the approach taken in [1]. Our approach has no autoregressive terms, instead the time-varying behaviour is captured by allowing the parameters of the model to vary over time. Our model is very simple with few parameters. The use of the Kalman filter allows us to track the parameters over time rather than rely on a more complicated model with a large set of parameters. The result is a simple yet effective model that can adapt to changing market conditions.

We simplify trading system logic using a four state finite state machine. The finite state machine is well suited to act as a control system for a stock market trading system. Studies of technical analysis rules [13] usually specify a set of rules specifying trade entry and exit conditions. Similarly, the FSM developed makes trading decisions based on the current state and expected return predicted by the model. Entry and exit conditions for each state specify trading actions to be taken. The backtested results show impressive gains for the trading system over the period studied

We add an additional hypothesis test to the threshold selection process. Prior methods of threshold selection did not require that the returns within each regime created by application of the threshold function be statistically different. The standard modelling procedure for SETAR models [26] compares competing models based on improving model fit while penalizing additional parameters. The basis for creating a multi-regime model is that the set of returns within each regime are different and need to be modelled differently. Thus, we add an appropriate hypothesis test to verify that the groups of returns created by the threshold function are in fact different as an additional criteria for threshold selection.

The contribution of this thesis is summarized as follows:

- Improve general SETAR(0) model building procedure by adding an additional hypothesis test to statistically justify the existence of a threshold non-linearity in the data prior to model building.

- Present a novel two-regime self-excited threshold model with time-varying parameters. This contrasts with earlier efforts to model time-varying returns with autoregressive models. Using the Kalman filter to track the parameters, we demonstrate the effectiveness of our method by showing improved fit compared to alternate models.

- We develop a FSM trading system that uses the parameter estimates from the Kalman filter to make investment decisions. The trading system is simple and compactly represented by the FSM. The system achieves exceptional backtested performance over the period 1928-2012. Following a buy and hold strategy, one dollar invested in 1928 would now be worth about 54 dollars. Following the FSM trading system, one dollar invested in 1928 would be worth 9570 dollars today, a dramatic increase.

1.4 Organization of Thesis

In Chapter 2 we discuss relevant background information on financial modelling and Kalman filter basics.

In chapter 3, we outline the empirical observations justifying the development of a non-linear threshold model. We further attempt to optimize the threshold to produce the best model, demonstrating the optimum threshold is effectively zero.

In chapter 4, we examine the behaviour of the expected returns within each state over time, showing that a dynamic model may be needed to describe the parameters of the model over time. We then track the mean return within each state using a standard Kalman filter assuming the parameters follow a random walk process.

In chapter 5, we develop a finite state machine trading system that uses the predictions of the Kalman filter to make investment decisions. We show profitable performance to this trading system over the period studied far in excess of a buy and hold strategy over the same period.

In chapter 6, we discuss the results and conclude the thesis with suggestions for future work.

Chapter 2

Background

This chapter presents relevant background material used in this thesis. We begin by presenting some simple but effective time-series models used to model investment returns. We discuss general non-linear models and progress to information specific to our threshold model. A description of the Kalman filter is also presented.

2.1 Random Walk Model

Perhaps one of the most influential papers in finance concerns the random character of stock prices, formally introduced by Bachelier [27]. He theorized that stock prices are generated by the stochastic process given in Equation 2.1.

$$x_t = x_{t-1} + \epsilon_t \tag{2.1}$$

In the model, x_t represents the price of the stock at time t while ϵ_t is a random variable often assumed to be zero mean IID Gaussian noise. This model is called the random walk model. The noise term represents the price changes at time t. In the finance literature, these price changes are referred to as returns. To apply the model requires estimating the statistical properties of the set of asset returns. The random walk is a martingale model, knowledge of past events has no effect on future outcomes. A martingale can also be described as a fair game offering little hope to an investor looking for advantageous times to invest.

2.2 Investment Returns

The model presented in this thesis is developed for daily asset returns. The investment return is the percentage price change from the previous day. It represents the percentage return an investor would receive for holding the asset for one day. Successful investing involves identifying investments with positive expected returns and betting on them repeatedly. In quantitative finance, the convention is to use continuously compounded returns rather than simple returns [28]. The one day continuously compounded return r_t is given in Equation 2.2 where x_t represents the price at time t.

$$r_t = log(x_t/x_{t-1}) \tag{2.2}$$

An advantages of using log returns is they are time additive meaning that the two-period log return is identical to the sum of the each period's log return. In contrast, the simple return is not time additive. Log returns are also mathematically convenient particularly when dealing with integration and differentiation. In this thesis we use log returns for all time series under consideration. We use adjusted daily closing prices in our analysis. The adjustments are due to dividend payments the investor would receive if they owned the asset that day. Adjusted prices therefore reflect the total return available to the shareholders, capital gains and dividends.

2.3 Geometric Random Walk with Drift

Figure 2.1 shows the DOW Jones Industrial Average raw price history from 1928 to July 2012.

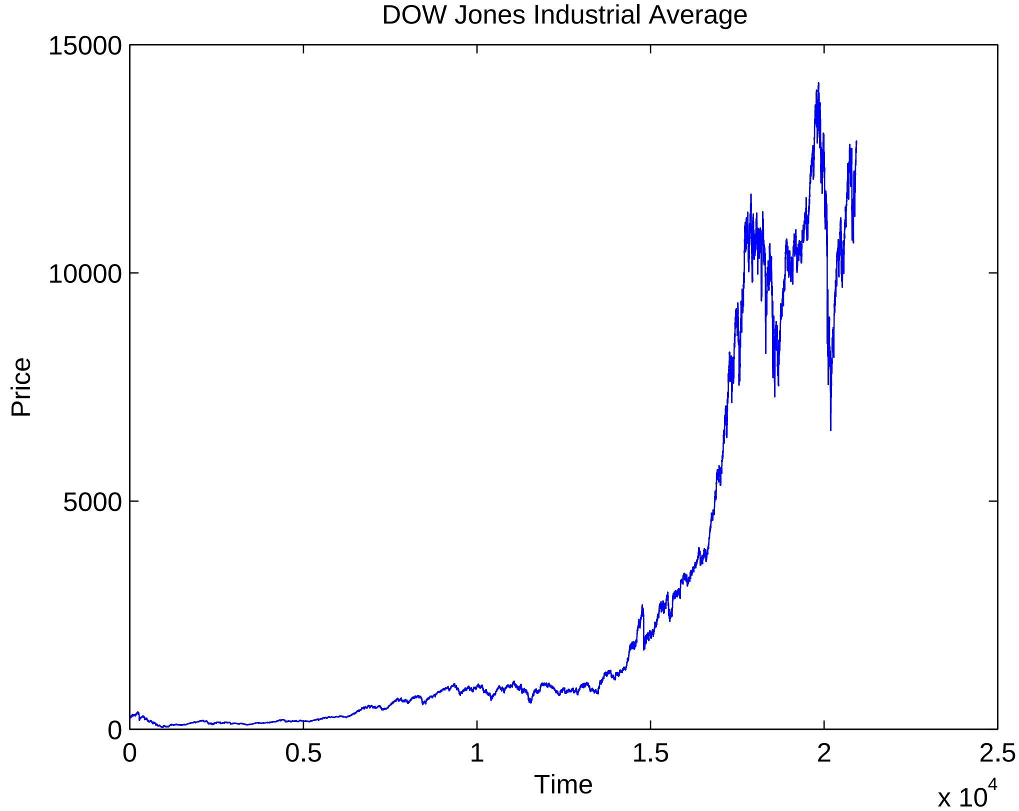

Figure 2.1: DOW Jones Raw Price

From the DOW Jones price history it is clear that the index has experienced an exponential upward drift over time. As the price increases, absolute volatility also increases. As a result, the geometric random walk with drift is a plausible model to describe the DOW Jones price history and many other equity indices. The DOW Jones logarithmic price series is shown in Figure 2.2

The model for the log returns r_t in this case is given in Equation 2.3.

$$r_t = \phi + \epsilon_t \tag{2.3}$$

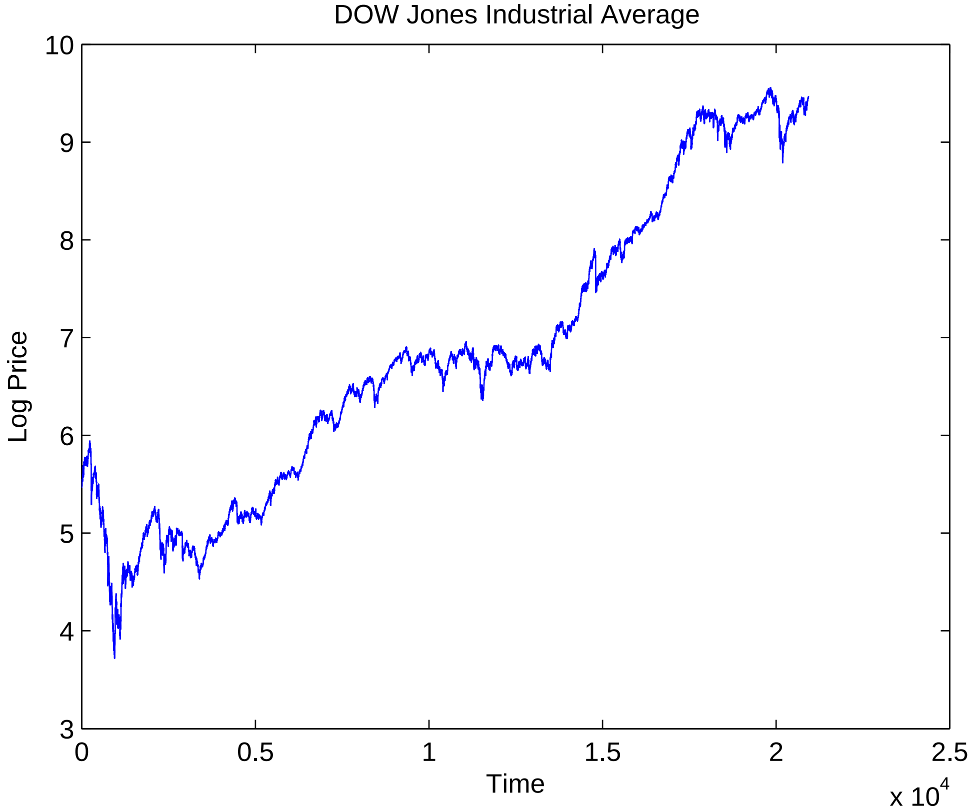

Figure 2.2: DOW Jones Logarithmic Price

In the geometric random walk with drift model, r_t is the log return, ϕ is the drift parameter and ϵ is a zero mean IID Gaussian noise process.

2.4 Non-Linear Time Series Models

While linear models are generally preferred, not all processes are linear. When presented with evidence of non-linear behaviour a non-linear model can be constructed in an attempt to better describe the data. In this thesis, we find a threshold non-linearity in the data and thus find use for a non-linear model.

Any time series model where the mean and/or variance of a random variable is described by a non-linear function is a non-linear model. In this thesis the random variable under consideration is the time varying mean. We find an asymmetry in the mean and create a model with two regimes. The threshold function divides the data into two groups and is in given in Equation 2.4.

$$s_t = \begin{cases} 0, & \text{if } r_{t-1} < 0 \\ \\ 1, & \text{if } r_{t-1} \geq 0 \end{cases} \tag{2.4}$$

The threshold function defines values of s_t and controls the state-switching. The threshold function depends on the return of the previous day r_{t-1}. In the above stated equation, the value of the threshold is zero.

2.5 Non-Linear Threshold Models

The class of threshold autoregressive models are considered to be locally linear [2]. Tong introduced the formal specification of the threshold autoregressive model. This type of model is in the family of state-dependent models. The state-switching is controlled by a threshold function that depends on observable endogeneous or exogeneous variables [28]. The state is therefore observable. Within each state, the process follows an autoregressive process. One of the simplest TAR models is the self-exciting threshold autoregressive. In this case, the threshold function is based on lagged values of the time-series itself and is used for the model developed in this thesis. Tong [3] developed the formal specification of SETAR model. The two-state SETAR model is given in Equation 2.5.

$$r_t = \begin{cases} \phi_{0,1} + \sum_{j=1}^{p_1} \phi_{j,1} r_{t-j} + \epsilon_{t,1}, & \text{if } r_{t-1} < 0 \\ \\ \phi_{0,2} + \sum_{j=1}^{p_2} \phi_{j,2} r_{t-j} + \epsilon_{t,2}, & \text{if } r_{t-1} > 0 \end{cases} \tag{2.5}$$

Where r_t is the return on day t, the set of $\phi_{i,j}$ are autoregressive terms for regime i and p_i is the order of the AR model in regime i. The noise terms $\epsilon_{t,i}$, is IID zero mean Gaussian noise in regime i. In our model, we apply a simple SETAR model, dividing the return series into a set of regimes and allowing the model parameters to vary by regime. The SETAR model allows the expected return to vary over time and attempts to use a linear combination of lagged values of the series to predict future values. In contrast, we capture the time variation using time varying parameters in a simple AR(0) model.

2.6 Kalman Filter

The Kalman filter method was originally developed by Rudolf Kalman [29]. The algorithm is designed to compute forecasts of the state of a system in an iterative fashion. The Kalman Filter can be applied a time series model represented in the form of state space system. The Kalman filter recursively estimates the state of a system corrupted by noise by using the measurements which are linearly related to the state and also corrupted by noise.

To implement the Kalman filter, the state equation and observation equation must be specified. Equation 2.6 and Equation 2.7 give the state and observation equations for a general Kalman filter application.

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{a}_t + \mathbf{M}_t \eta_t \tag{2.6}$$

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{x}_t + \epsilon_t \tag{2.7}$$

Where $\mathbf{x}$ is the state vector at a discrete time period t. $\mathbf{F}_t$ is the state transition matrix that is applied to the state vector of the previous period $\mathbf{x}_{t-1}$. $\mathbf{a}_t$ is the control input vector. The control input matrix is $\mathbf{B}_t$. $\mathbf{v}_t$ is generally a serially uncorrelated and independent noise with $\mathbf{v}_t \sim N(0, \mathbf{R}_t)$. $\mathbf{M}_t$ is the matrix applied to this noise. $\mathbf{y}_t$ is the observation vector, $\mathbf{A}_t$ is the observation matrix applied to the state vector $\mathbf{x}_t$ at time period t. $\mathbf{w}_t$ is the observation noise independent and serially uncorrelated following distribution $\mathbf{w}_t \sim N(0, \mathbf{Q}_t)$. Furthermore, for all time periods the observation and state equation disturbances are uncorrelated with each other and also with the initial state vector $\mathbf{x}_0$. In this system of equation $\mathbf{a}_t$ is non-random input to the system. The matrices $\mathbf{F}_t$, $\mathbf{B}_t$, $\mathbf{A}_t$ and $\mathbf{M}_t$ depend on the model specification. This representation of state space model with state noise matrix is described in [30].

The recursive process of Kalman Filter involves the forecast of the future state based on the previous updated state estimate and the new observed data. Because of its recursive nature, the Kalman Filter can be applied in real time. When studying time-varying phenomena, the Kalman filter can be used to track parameters in a model in an iterative fashion. Our model has time-varying parameters and as such, tracking using the Kalman filter is used.

Chapter 3

Establishing non-Linear Behaviour in the Return Distribution

In this chapter, we use empirical results to establish a threshold non-linearity in the return distribution of the US equity indices. The basic idea of a threshold model is piecewise linearisation of the observed series through the introduction of a threshold function. We divide the time series under consideration into groups based on some past information. We have chosen to use a self-excited threshold function with lag 1. We need to determine the number of regimes in the model and the threshold values. We combine hypothesis testing and model fitting to accomplish this. When considering a threshold value, hypothesis testing is first employed to verify that the returns within each regime created are different with statistical significance from the set of returns in the other regimes. We then find the threshold value(s) and other model parameters that produce the best fitting model by minimizing the mean squared error.

3.1 Dataset

The dataset under consideration in this thesis consists of the daily adjusted closing prices of the DOW Jones Industrial Average over a long period. The dates studied range from Jan 1928-July 2012, approximately 21,000 datapoints. This dataset was selected because it covers wide range of market conditions including several market crashes, world wars and a significant portion of the past century. Modern portfolio theory establishes that a diversified portfolio of equities has superior risk return characteristics compared to individual stocks [16]. To invest in equities, an investor should buy an index like the DOW Jones Industrial Average representing a diversified portfolio. In this thesis the investor has two investment options, invest in the DOW Jones Industrial Average or invest in cash.

Additional indices are also used to test the model. Two other major US equity indices are also studied. The NASDAQ Composite daily adjusted closing prices from Feb 1971 to July 2012 and the S& P 500 Index daily adjusted closing prices Jan 1950 to July 2012. The NASDAQ Composite Index is a market capitalization weighted index of all the stocks listed on the NASDAQ stock exchange. The index is mainly composed of technology companies. The S& P 500 Index is composed of the five hundred largest US companies by market capitalization and some additional criteria. This is a very broad index that is often used to represent the performance of the US equity markets as a whole.

3.2 Threshold Selection Process

We begin by defining a piecewise constant threshold function for an arbitrary number of states n in Equation 3.1.

$$s_t = \begin{cases} 1, & \text{if } r_{t-1} < \gamma_1 \\ 2, & \text{if } r_{t-1} \geq \gamma_1 \; AND \; r_{t-1} < \gamma_2 \\ ... & \\ n, & \text{if } r_{t-1} \geq \gamma_{n-1} \end{cases} \tag{3.1}$$

Where r_{t-1} is the return of the previous day, the set $\gamma_1, \gamma_2, ..., \gamma_{n-1}$ are the threshold values and the value of s_t represents the state or regime. When applied to an input series, the threshold function creates n groups of observations. We let $\gamma_0, \gamma_1, ..., \gamma_n$ denote a linearly ordered subset of real numbers to represent our set of thresholds. We require $\gamma_0 < \gamma_1 < ...\gamma_{n-1}$ and set γ_0 as $-\infty$ and γ_l as $+\infty$. The set of regimes are defined in Equation 3.2.

$$\Gamma_j = (\gamma_j - 1, \gamma_j] \tag{3.2}$$

The set of regimes define a partition of the set of real numbers $\mathbb{R}$.

$$\mathbb{R} = \Gamma_1 \cup \Gamma_2 \cup ... \cup \Gamma_n \tag{3.3}$$

Clearly, including the threshold function in a model only makes sense if supported by empirical evidence. We need to establish the existence of a threshold non-linearity in the DOW Jones return series, determine the number of regimes and find the set of threshold values.

We alter the standard modelling procedure for the class of SETAR models [26] by requiring that any groups created by the piecewise linearisation are statistically different to any neighbouring groups.

The common modelling approach for SETAR(m) models with m states [4] involves beginning with a two state model, or SETAR(2) model, and searching for the set of model parameters that produce the minimum MSE, these parameters include both the value of the threshold and the AR coefficients of the model. This approach finds the threshold value and other model parameters simultaneously. An F-test is then used to compare the SETAR(m-1) model to the SETAR(m) model. The SETAR(1) model is a linear model, thus an F-test comparing the SETAR(1) model and the SETAR(2) model serves as a test of the non-linear model versus a linear alternative.

In contrast to this approach, we divide the model building process into two steps. We limit our set of possible threshold values to those for which adjacent groups pass an appropriate hypothesis test at a strict level of confidence. Separating the original time series into groups should require strong evidence that the groups created are in fact different from adjacent groups. Thus, instead of searching for the threshold values and model parameters simultaneously, we use hypothesis testing to determine a set of possible threshold values first. Then using only these values, proceed the model selection process involving least squares parameter estimation and compare candidate models.

For our purposes, we are interested in finding differences in mean and possibly variance between groups of returns. As a result we use Welch's t-test for the statistical hypothesis test to compare adjacent groups of returns [31]. The test is designed to compare the means of two samples having possibly unequal variances. The null hypothesis states that the two sample means are equal. Rejecting the null would give evidence to support a threshold non-linearity. For more general applicability, using the non-parametric distribution free Kolmogorov-Smirnoff test as used by Lo [14] when comparing sets of returns may be more appropriate.

The threshold function divides the set of returns into two or more groups. For a two state model the conditional probability density functions are given by $f_r(r|S=0)$ and $f_r(r|S=1)$. The hypothesis test is stated as follows:

$\mathbf{H}_0$: $f_r(r|S=0)$ and $f_r(r|S=1)$ have equal means

$\mathbf{H}_1$: $f_r(r|S=0)$ and $f_r(r|S=1)$ have different means

If the null hypothesis is true there is little justification to proceed with a model based on asymmetry in mean. If the alternate hypothesis is true we have evidence to support a two-regime model.

3.3 DOW Return Series Threshold Search

3.3.1 Two State Model

As previously noted, we begin by determining the set of possible values of the threshold γ_1 for a two state model. These are values of the threshold that separate the original series into distinctly different groups as determined by Welch's t-test with $\alpha = 0.05$. The threshold function is given in Equation 3.4.

$$s_t = \begin{cases} 0, & \text{if } r_{t-1} < \gamma_1 \\ \\ 1, & \text{if } r_{t-1} \geq \gamma_1 \end{cases} \tag{3.4}$$

To evaluate a potential threshold value γ_1, we begin by fixing the threshold to some value. Using this threshold we then separate the return distribution distributions into two groups. The hypothesis test is used to compare the mean return within the regimes.

Figure 3.1 plots the $p - value$ of a the two sample t-test comparing the two groups of returns created by application of the threshold function for a range of threshold values.

It is clear from Figure 3.1 of the hypothesis test over a range of thresholds that in the case of a two state model we should confine the threshold between -1% and 1%. Any value chosen in this range satisfies our requirement for a statistically significant difference in mean between the two groups.

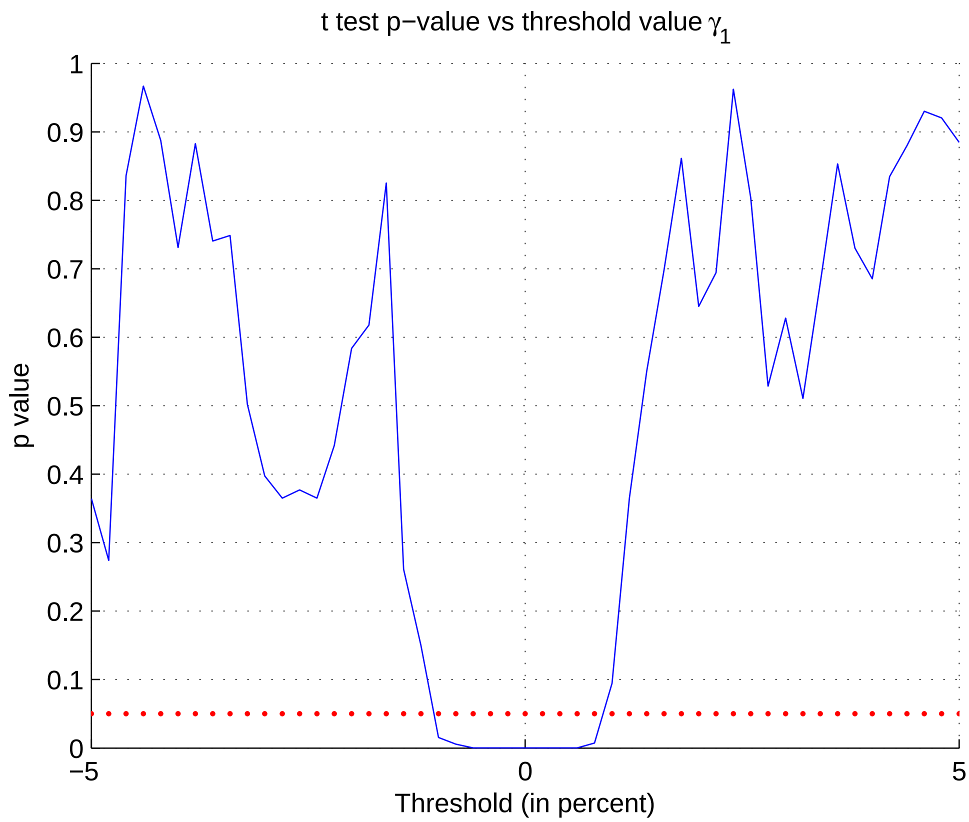

Figure 3.1: t-test p-values for a range of threshold values

3.3.2 Three State Model

We now consider the case of a three state model with two threshold values and compare results to the two-regime case. The threshold function for the three state model requires two thresholds to divide the time series into three groups and is given in Equation 3.5.

$$s_t = \begin{cases} 0, & \text{for } r_{t-1} < \gamma_1 \\ \\ 1, & \text{for } r_{t-1} \geq \gamma_1 \; And \; r_{t-1} < \gamma_2 \\ \\ 2, & \text{for } r_{t-1} \geq \gamma_2 \end{cases} \tag{3.5}$$

For the three regime case search for possible threshold values for which adjacent groups have different means with statistical significance at $\alpha = 0.05$. As an example, for $\gamma_1 = -1\%$ and $\gamma_2 = 1\%$ we divide the return distribution into three groups as shown in Equation 3.6.

$$s_t = \begin{cases} 0, & \text{for } r_{t-1} < -0.01 \\ 1, & \text{for } r_{t-1} \geq -0.01 \; And \; r_{t-1} < 0.01 \\ 2, & \text{for } r_{t-1} \geq 0.01 \end{cases} \tag{3.6}$$

We then perform two hypothesis tests comparing adjacent groups, comparing $f_r(r|S = 0)$ with $f_r(r|S = 1)$ and $f_r(r|S = 1)$ compared to $f_r(r|S = 2)$. The hypothesis test used is the two sample t-test. The hypothesis tests performed are stated as follows:

Hypothesis Test 1

$\mathbf{H}_0$: $f_r(r|S = 0)$ and $f_r(r|S = 1)$ have equal means

$\mathbf{H}_1$: $f_r(r|S = 0)$ and $f_r(r|S = 1)$ have different means

Hypothesis Test 2

$\mathbf{H}_0$: $f_r(r|S = 1)$ and $f_r(r|S = 2)$ have equal means

$\mathbf{H}_1$: $f_r(r|S = 1)$ and $f_r(r|S = 2)$ have different means

We perform a grid search of a large set of possible pairs of threshold values γ_1 and γ_2 performing the above stated hypothesis tests. We find that the range for the first threshold γ_1 should be between -0.7% and -0.1%. The second threshold, γ_2 is in the range 0.1% to 0.6%. The hypothesis testing performed has allowed us to limit our range of possible threshold values for the two and three regime cases.

3.4 Threshold Model for DOW Jones Data

Using the sets of potential thresholds determined in the previous section, we turn our attention to determining the best fitting model for the one and two threshold cases. This involves determining the threshold values and model parameters that produce the best model as defined by the lowest mean squared error. We begin by fixing our threshold, then estimate the parameters of the model in Equation 3.7. We then determine the mean squared error of the resulting model given the threshold value and the estimated parameter values. We then repeat this process over a range of threshold values to find the set of parameters that produce the best fitting model by minimizing the mean squared error using the entire dataset.

$$r_t = \begin{cases} \phi_0 + \epsilon_{t,0}, & \text{if } r_{t-1} < \gamma_1 \\ \phi_1 + \epsilon_{t,1}, & \text{if } r_{t-1} \geq \gamma_1 \end{cases} \tag{3.7}$$

Within each state the returns follow a random walk with drift, the drift parameters are ϕ_0 and ϕ_1. The noise terms are given by $\epsilon_{t,0}$ and $\epsilon_{t,1}$ and are assumed to be IID zero mean Gaussian noise. The threshold value is γ_1. Parameters are estimated with least squares.

We display our results for the two-regime model for a range of threshold values in Figure 3.2.

From the Figure 3.2 we observe a distinct improvement in model performance for a threshold of effectively zero. This is a very interesting result.

The model parameters for our best fitting two-state model are summarized in 3.1 with percentage returns used.

Table 3.1: Two-Regime Model Parameter Summary

Parameter	Value
ϕ_1	0.067
σ_{ϵ_1}	1.07
ϕ_0	-0.034
σ_{ϵ_2}	1.26

The key observation is that the expected return given the previous return is positive is positive, while the expected return given the previous return is negative is negative.

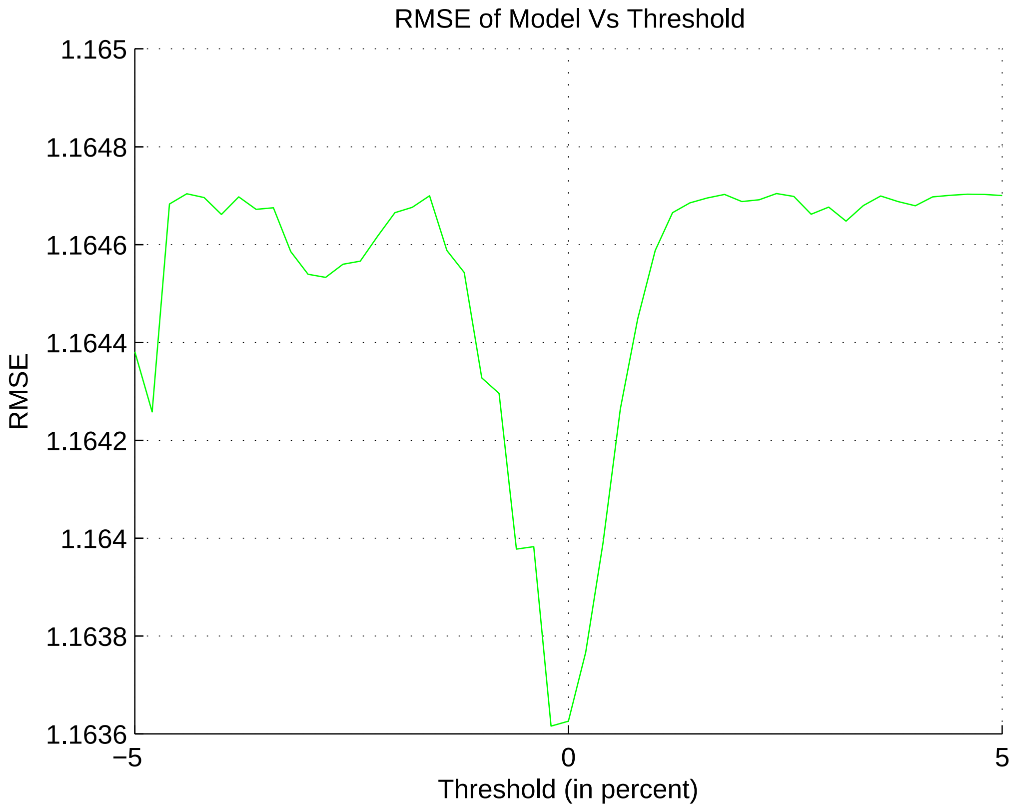

Figure 3.2: RMSE of two-state model for a range of threshold values DOW

The linear alternative model would be a one state random walk with drift. Using percentage returns, the RMSE of the linear alternative is 1.1647 while the best two state model lowers the RMSE to 1.1636.

For the three state model, we find that among our set of acceptable thresholds the lowest RMSE is produced for $\gamma_1 = -0.2\%$ and $\gamma_2 = 0.2\%$, The RMSE error for this model is 1.1635. We summarize the resulting model fit for the one, two and three regime cases in Figure 3.3. As shown in the plot, minimal gains in modelling efficiency are gained by adding a third regime and thus we prefer the two-regime model.

The selected model with estimated parameters for percentage returns is stated in Equation 3.8.

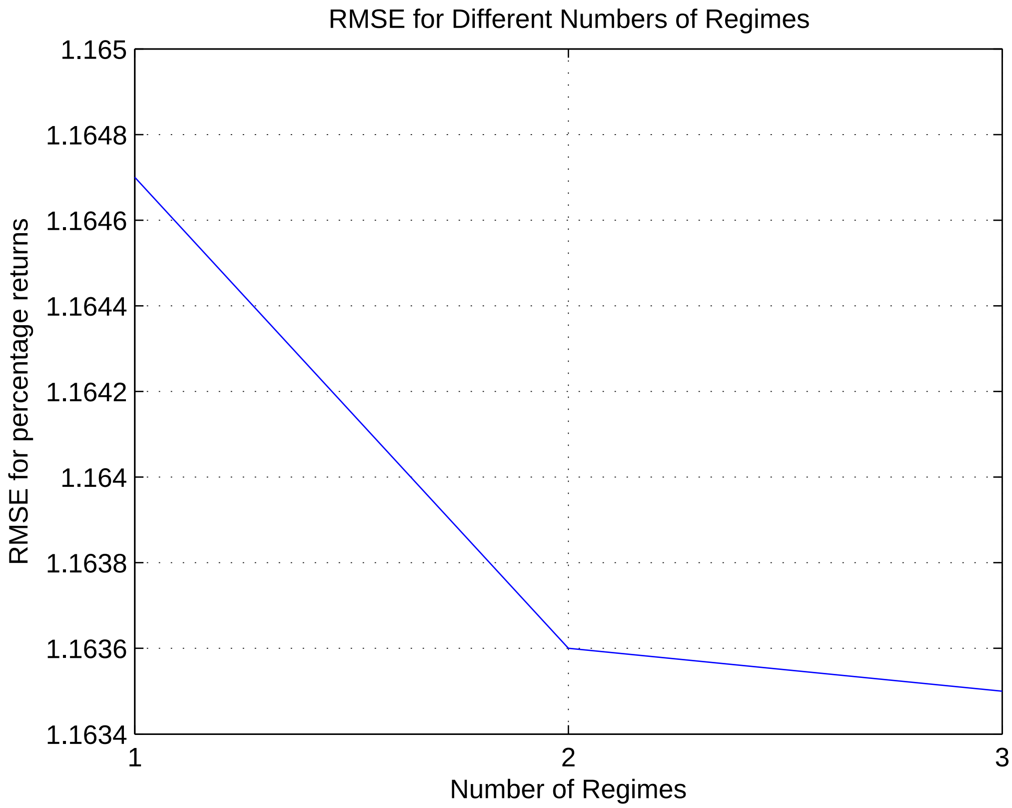

Figure 3.3: Comparing RMSE for one, two and three regime models

$$r_t = \begin{cases} 0.067 + \epsilon_{t,1}, & \text{if } r_{t-1} \geq 0, \ \epsilon_{t,1} \sim N(0, 1.145) \\ -0.034 + \epsilon_{t,2}, & \text{if } r_{t-1} < 0, \ \epsilon_{t,2} \sim N(0, 1.58) \end{cases} \tag{3.8}$$

The estimated parameters for the two-state model show that the higher return regime also has lower variance, making it both a higher return and lower risk state in terms of mean-variance risk.

We also want to test to see if an AR(1) model within each regime may be better than the AR(0) model used so far. The parameters were fitted with least squares for an AR(1) threshold model for the DOW dataset and the resulting model is presented in Equation 3.9 for percentage returns.

$$r_t = \begin{cases} -0.0655 - 4.14 r_{t-1} + \epsilon_{t,1}, & \text{if } r_{t-1} < 0, \ \epsilon_{t,1} \sim N(0, 1.15) \\ 0.0695 - 0.36 r_{t-1} + \epsilon_{t,2}, & \text{if } r_{t-1} \geq 0, \ \epsilon_{t,2} \sim N(0, 1.58) \end{cases} \tag{3.9}$$

It is interesting to note that both regimes show negative autocorrelations at lag 1 but the effect is much more pronounced following a negative day. The RMSE for the AR(1) two regime model is 1.1634, a small improvement on the AR(0) version which a MMSE of 1.1636. As the gains from using an AR(1) are minimal we prefer the simpler AR(0) model within each regime.

3.5 Threshold Model for Additional Stock Indices

The above procedure applied to modelling the DOW Jones data was applied to the NASDAQ Composite and S& P 500 Index.

For the NASDAQ Composite, a threshold of zero for a two regime model is found to produce the best model. Figure 3.4 displays the model error for a range of threshold values using percentage returns.

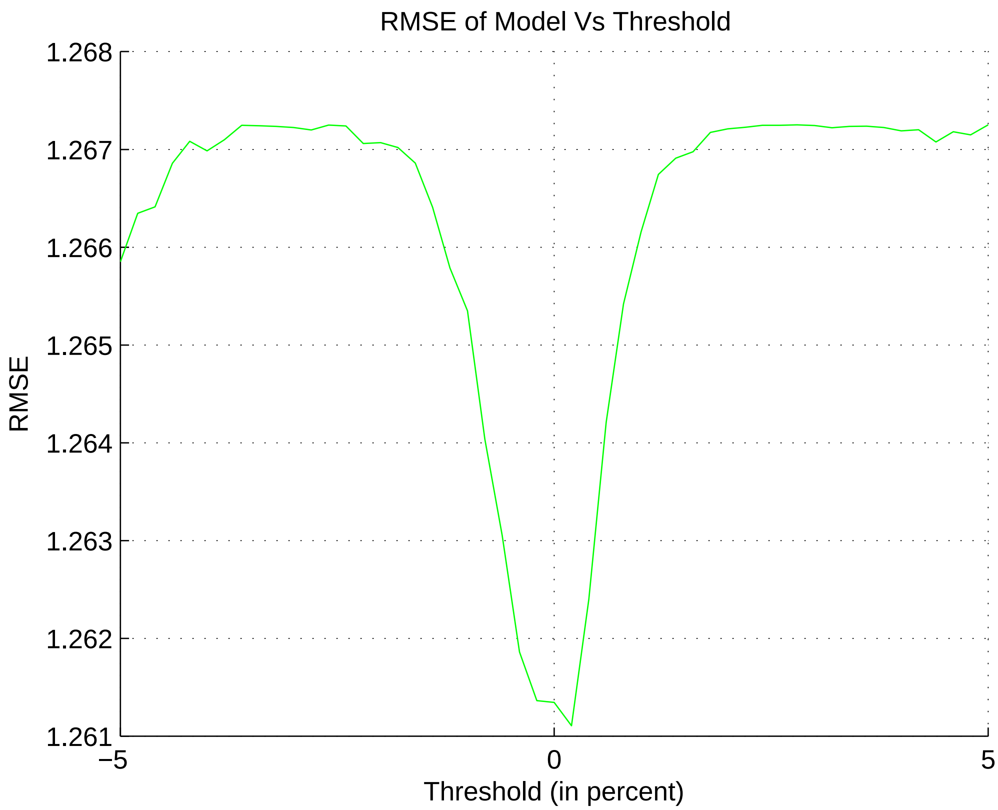

Figure 3.4: RMSE of two-regime model VS threshold for NASDAQ Comp

Similar results are found for the S& P 500 as shown in Figure 3.5. From these results, it is clear that a threshold model can be considered for all three major US equity indices and the value of the threshold is consistently zero.

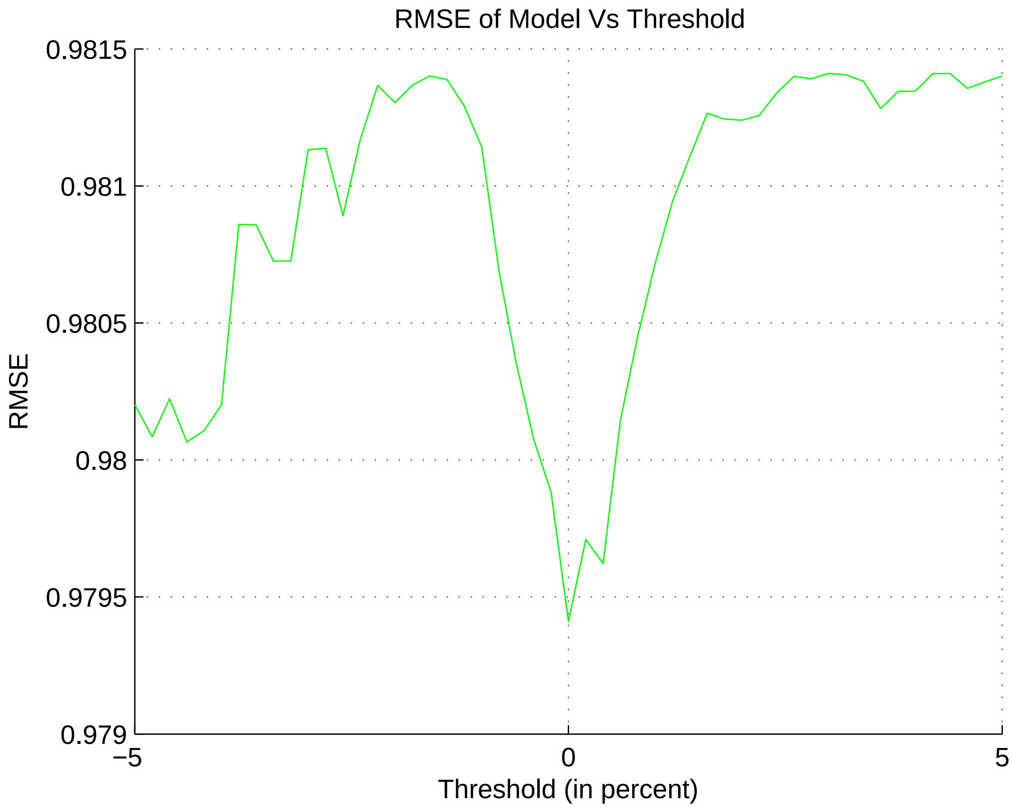

Figure 3.5: RMSE of two-regime model VS threshold for S& P 500

3.6 Chapter Summary

In this chapter, we provide empirical evidence to demonstrate the existence of an asymmetry in the daily return series of the DOW Jones Industrial Average conditional on the sign of the previous return. When the previous day is positive we get a positive expected return for the current day. When the previous day is negative we get a negative expected return for the current day. We use this observation to motivate the development of a two regime model. The threshold value is found to be approximately zero. We demonstrate improved model fit when compared to an alternative linear model. Our modelling procedure also includes a hypothesis testing step to ensure the threshold function separates the data into distinct groups. We further show that the same model is appropriate for both the S& P 500 Index and the NASDAQ Composite Index with the same threshold value of zero.

Chapter 4

Examining the Time Varying Behaviour of Returns

In the previous chapter we established a threshold non-linearity in the DOW Jones return series over a very long period. As a result we model the return process using two different states. In this chapter we perform a similar test but examine the mean return within each state for each consecutive 2000 point sub period in the original dataset. The purpose is to look for time varying behaviour. The empirical data demonstrates that the mean return within each state varies considerably over time while remaining different for the two states in each sub period. In particular, the mean return within each state changes considerably during the crash of 1929 and the recent financial crisis of 2008. The finding will motivate us to consider time-varying parameters in our model that adapt to changing conditions. We further extend our analysis to the NASDAQ Composite and S& P 500 datasets and find similar results.

4.1 Conditional Mean Returns Over Time

In the previous chapter we examined the expected return given the sign of the previous return using the entire dataset. In this chapter we examine the same hypothesis but for smaller time periods within the dataset. The data is divided into groups of consecutive 2000 point intervals. This size of interval was used because we wanted to examine the behaviour over time while keeping a large number of points in each interval for hypothesis testing. Each period is labelled P_i for $i = 1$ to $i = 11$ for the DOW Jones dataset as shown in Equation 4.1.

$$P_i = \{r_{2000(i-1)+1}, r_{2000i}\} \tag{4.1}$$

The threshold function defined by s_t is then used to classify the samples into two conditional distributions for each sub-period. We compare the mean values of the two groups by performing Welch's t-test for each set of points. The purpose of this analysis is twofold, we will verify if the observed asymmetry of returns is persistent over time and determine if the model parameters may vary over time. The results of this test are displayed in Table 4.1. The table compares the mean returns, $E(r_t|s_t = 0)$ and $E(r_t|s_t = 1)$ for the two regimes in each sub-period and provides results of the hypothesis test comparing the means of the two samples. All returns are expressed in percentage terms.

The hypothesis test used is Welch's t-test for two samples.

The threshold function divides the set of returns into two or more groups with conditional probability density functions $f_R(r|S = 0)$ and $f_R(r|S = 1)$. For a two regime model the hypothesis test is stated as follows:

$\mathbf{H}_0$: $f_R(r|S = 0)$ and $f_R(r|S = 1)$ have equal means

$\mathbf{H}_1$: $f_R(r|S = 0)$ and $f_R(r|S = 1)$ have different means

The results convincingly demonstrate a mean asymmetry between the two groups of returns is persistent over the sub periods studied. In nine out of eleven test periods, the

Table 4.1: State Dependent Returns by Sub Period

Period	$E(r_t)$	$E(r_t \mid s_t = 1)$	$E(r_t \mid s_t = 0)$	count $s_t = 1$	count $s_t = 0$	H	p-value
1	-0.0155	-0.1078	0.085	1043	957	0	0.05438
2	-0.0083	0.0939	-0.116	1027	973	1	0.00005
3	0.0302	0.1215	-0.075	1071	929	1	0.00005
4	0.0391	0.1341	-0.079	1110	890	1	0.00000
5	0.0227	0.1154	-0.081	1055	945	1	0.00000
6	0.0024	0.1878	-0.175	980	1020	1	0.00000
7	0.0107	0.0555	-0.035	1010	990	1	0.02326
8	0.0509	0.0932	0.002	1071	929	0	0.08687
9	0.0575	0.1026	0.005	1076	924	1	0.01990
10	0.0087	-0.0413	0.063	1040	960	1	0.02701
11	0.0013	-0.1171	0.141	505	429	1	0.01793

hypothesis test yields 95 percent confidence the two samples are in fact different. In eleven out of eleven test periods, the hypothesis test yields 90 percent confidence the two samples are different. The bar chart 4.1 compares the mean return within each regime by sub period. These are given by $E(r|S = 0)$ and $E(r|S = 1)$. The expected return for all returns during each sub period is given by $E(r)$.

Interestingly, at the start and end of the dataset, the behaviour of the two states appears to invert, positive return days precede expected negative returns and negative return days precede expected positive returns which is the opposite of the rest of the dataset.

The empirical results suggest that the observed return asymmetry is persistent over time but the expected returns within each state may vary over time. As a consequence of these findings, we modify the model employed in Chapter 2 to include time varying parameters. Our model with time varying drift is stated in Equation 4.2.

$$r_t = \begin{cases} \phi_{0,t} + \epsilon_{0,t}, & \text{if } r_{t-1} < 0 \\ \phi_{1,t} + \epsilon_{1,t}, & \text{if } r_{t-1} \geq 0 \end{cases} \tag{4.2}$$

The values of the parameters $\phi_{0,t}$ and $\phi_{1,t}$ vary with time and ϵ_t is an IID Gaussian noise process.

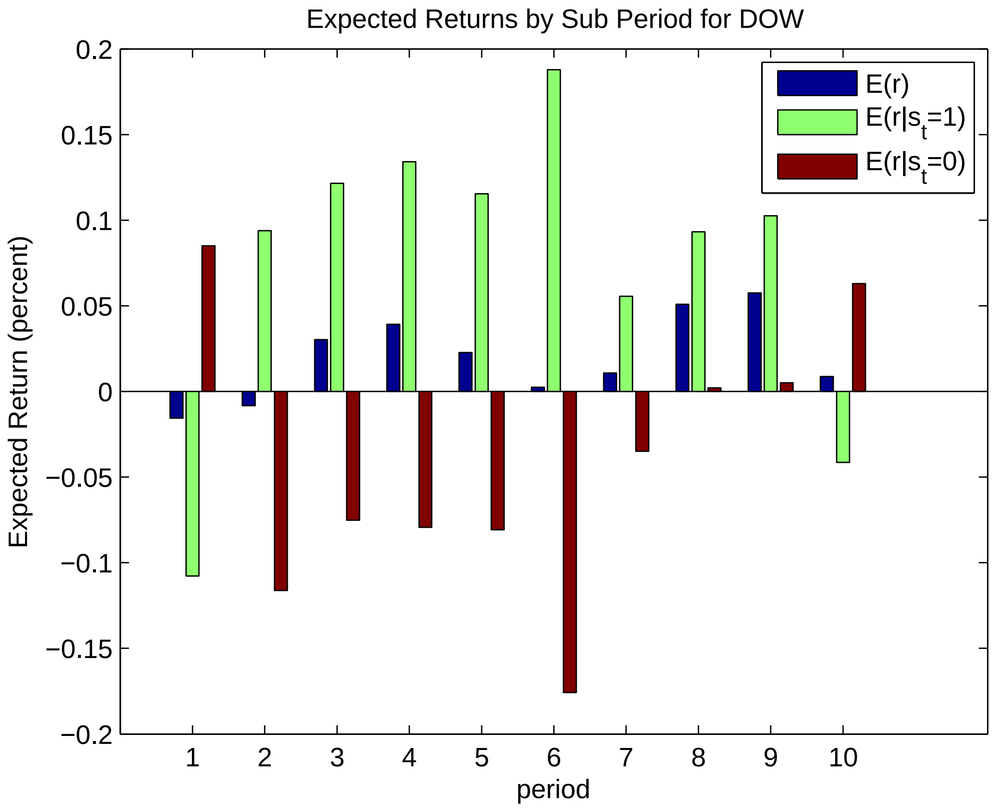

Figure 4.1: Expected Returns by Regime for DOW Over Time

4.2 Kalman Filter for Time-Varying Parameter Estimation

The Kalman filter is a recursive algorithm that can be used to optimally estimate parameters under the assumption of Gaussian noise. In this thesis, our interest is estimating the mean return within each state. From the analysis presented in this chapter we have established that the parameters of the model likely vary over time. In order to apply the Kalman filter to this problem we will need to specify a process that determines the time evolution of the parameters to be estimated. We will choose the random walk process for the parameters $\phi_{0,t}$ and $\phi_{1,t}$ with process equations given in Equation 4.3.

$$\phi_{0,t} = \phi_{0,t-1} + w_t \qquad \phi_{1,t} = \phi_{1,t-1} + v_t \tag{4.3}$$

From the general state Equation in 2.6, the matrices to be specified are $\mathbf{F}_t$, $\mathbf{B}_t$ and $\mathbf{M}_t$. Equation 4.4 provides the values of these matrices for our model.

$$\mathbf{F}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{B}_t = 0, \mathbf{M}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.4}$$

The state vector defines the parameters to be estimated. The state vector for our model is given in Equation 4.5.

$$X_t = \begin{bmatrix} \phi_{0,t} \\ \phi_{1,t} \end{bmatrix} \tag{4.5}$$

The process noise vector is given by η_t in Equation 4.6.

$$\eta_t = \begin{bmatrix} w_t \\ v_t \end{bmatrix}. \tag{4.6}$$

The process noise terms, w_t and v_t are assumed to be zero mean IID Gaussian noise. Since we want to track the parameters we prefer to make few assumptions regarding the process by which the parameters evolve and use the simple random walk process.

We assume the observation noise ϵ_t is the same for both states and is an IID Gaussian process with zero mean. The standard SETAR model allows the observation noise variance to change depending on the state. By assuming the noise variance is the same for both states, we can use a single observation equation to estimate both state parameters simultaneously. We do this by including the state information in the observation vector. Our observation equation is given in Equation 4.7.

$$r_t = \phi_{0,t} + \phi_{1,t} s_t + \epsilon_t \tag{4.7}$$

The state switching is controlled by s_t defined in Equation 4.8.

$$s_t = \begin{cases} 0, & \text{if } r_{t-1} < 0 \\ 1, & \text{if } r_{t-1} \geq 0. \end{cases} \tag{4.8}$$

The general Kalman filter observation Equation is given in 2.7.

$$r_t = A_t X_t + \epsilon_t \tag{4.9}$$

The observation vector A_t for our model is given in 4.10.

$$A_t = \begin{bmatrix} 1 & s_t \end{bmatrix} \tag{4.10}$$

The parameters are tracked over time using the Kalman filter. An advantage to using the Kalman filter algorithm is that it relies on past data to make future predictions making it useful for practical investment decision making. When new information is received it is compared to the current prediction, the parameter estimates are updated and a new prediction is made.

The Kalman filter is a recursive algorithm that gives the optimal estimate of Gaussian Random variables. In order to apply the Kalman filter to the estimation problem, asset returns must be assumed to be normally distributed and the various noise processes must be assumed to be zero mean and Gaussian.

We set the initial conditions as shown in 4.11 and 4.12.

$$X_{1|0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.11}$$

$$\Sigma_{1|0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{4.12}$$

The variances of the noise processes are estimated by recursively using the Kalman filter and maximum likelihood estimation. By using the Kalman filter repeatedly along with the log likelihood function, we test the likelihood of the model with different values for the noise variances to select the most likely set of values. Table 4.2 summarizes the values of the noise variances that produce the most likely model.

The small value of the process noise suggests the parameters vary slowly over time.

Table 4.2: Noise Variances

Parameter	Variance
σ_w^2	8×10^{-5}
σ_v^2	8×10^{-5}
σ_ϵ^2	1.3

4.3 Results for DOW Dataset

In the two-state model described in chapter 3, the regime dependent drift parameters are constant over time. This assumes that there is some true value of these parameters and the more data we have, the closer we will get to the true values. The parameter values are estimated by minimizing the mean squared error. In this chapter we have argued for time-varying parameters and assume the parameter values follow a random walk process. As a result, the parameter values are tracked over time. This approach assumes there is no constant true value of the parameters. Financial markets are dynamic, participants adapt and structural, political and economic changes occur. A model with time-varying parameters is consistent with an changing financial system.

We evaluate the dynamic model by comparing the mean squared prediction error of the Kalman filter to the simpler two-state model with static coefficients estimated from the entire sample. If the Kalman filter tracking is effective, allowing for time-varying parameters should produce a better model fit than using static parameters in a similar model.

By comparing the RMSE for each model we find that the performance of the time-varying model with Kalman filter estimation is superior to the static two state model described in chapter three. While the predictions of the best static two state model have RMSE of 1.1636, the dynamic model with Kalman filter tracking has a RMSE of 1.1615, a significant improvement over the static model.

In [1], the authors use a SETAR model with two-regimes with an AR(1) model in each regime. They use a rolling window technique to allow for time-varying parameters and test a trading system based on the results. They take a different approach to attempt to capture time varying returns use a rolling window estimation method. The rolling window technique uses the past w observations to estimate the model parameters, updated estimates are calculated with each new data point. It is a simple method of adding dynamic behaviour to the parameters of a model. We tested this model for four different window lengths $w = 50, 100, 150$ and 200. The best model resulted for a window length of $w = 200$. Unfortunately, this method was unable to beat the one-state random walk with drift model. The RMSE for the two state SETAR model with AR order 1 and window length $w = 200$ was 1.2239 which is considerably worse than the random walk with drift model.

The success of the Kalman filter technique employed in this thesis provides evidence to support the use of time varying coefficients in our model. Table 4.3 compares model fit for five models considered in this thesis for percentage returns.

Table 4.3: Model Fit Comparison

Model	Parameter Estimation	Static or Dynamic Parameters	MSE
Two State AR(1)	Rolling Window Least Squares	Dynamic	1.4977
One State AR(0)	Least Squares	Static	1.3565
Two State AR(0)	Least Squares	Static	1.3539
Two State AR(1)	Least Squares	Static	1.3534
Two State AR(0)	Kalman Filter	Dynamic	1.3490

The model with time-varying parameters estimated by the Kalman filter fits the data better than the static parameter model as well as a dynamic model with rolling window parameter estimation. This provides justification for the Kalman filter approach. The estimated values of the parameters over time are shown in Figure 4.2.

For the majority of the dataset the expected return given $r_{t-1} \geq 0$ is positive but undergoes significant change early and late in the dataset. The Kalman filter captures these changes in behaviour. Clearly, the expected returns are different depending on the state of the system.

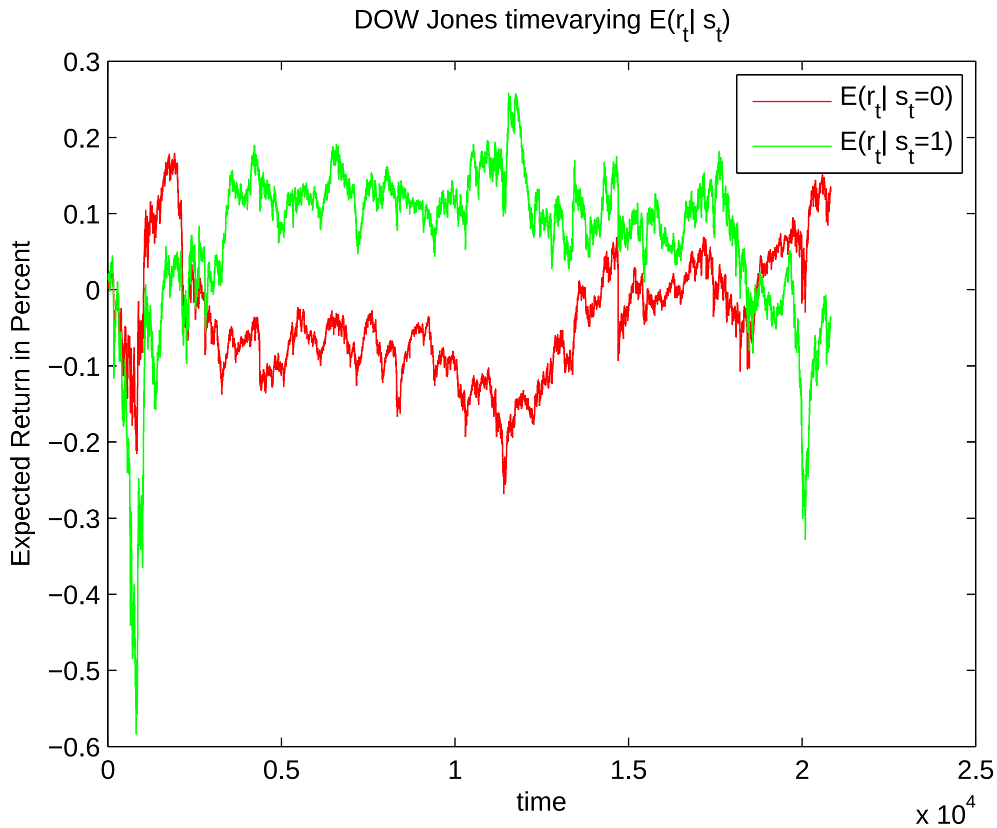

Figure 4.2: DOW Jones Expected Return Given $s_t = 1$ or $s_t = 0$

4.4 Dynamic Parameter Threshold Model for Additional Stock Indices

The approach used in this chapter was also applied to the NASDAQ Composite Index and the S & P 500 Index data. Each of these datasets was divided into 2000 point intervals and comparison of the state dependent expected returns were compared as with the DOW dataset. Figure 4.3 shows expected returns overs time for the S& P 500 index. The differences in mean for each sub period gives evidence to support a model with time-varying parameters and a threshold non-linearity as found for the DOW dataset.

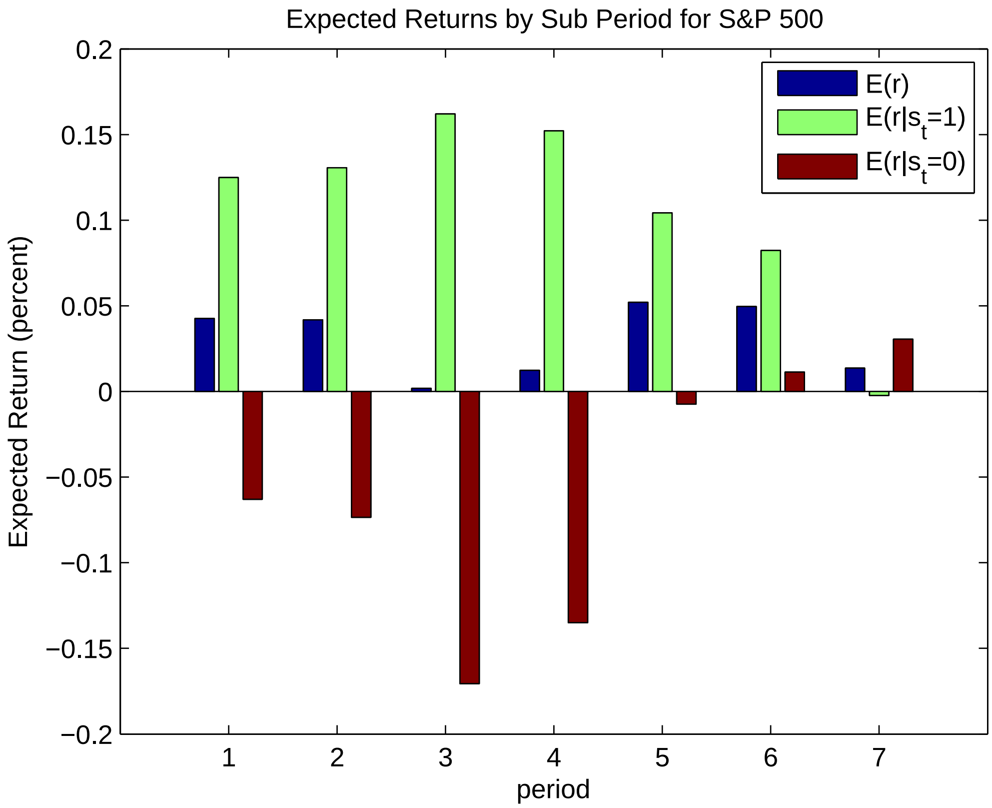

Figure 4.3: State Dependent Mean Return by Sub-Period S& P 500

The NASDAQ Composite Index displays the same behaviour. Figure 4.4 shows expected returns overs time for the NASDAQ Composite. The differences in mean for each sub period gives evidence to support a time-varying threshold non-linearity.

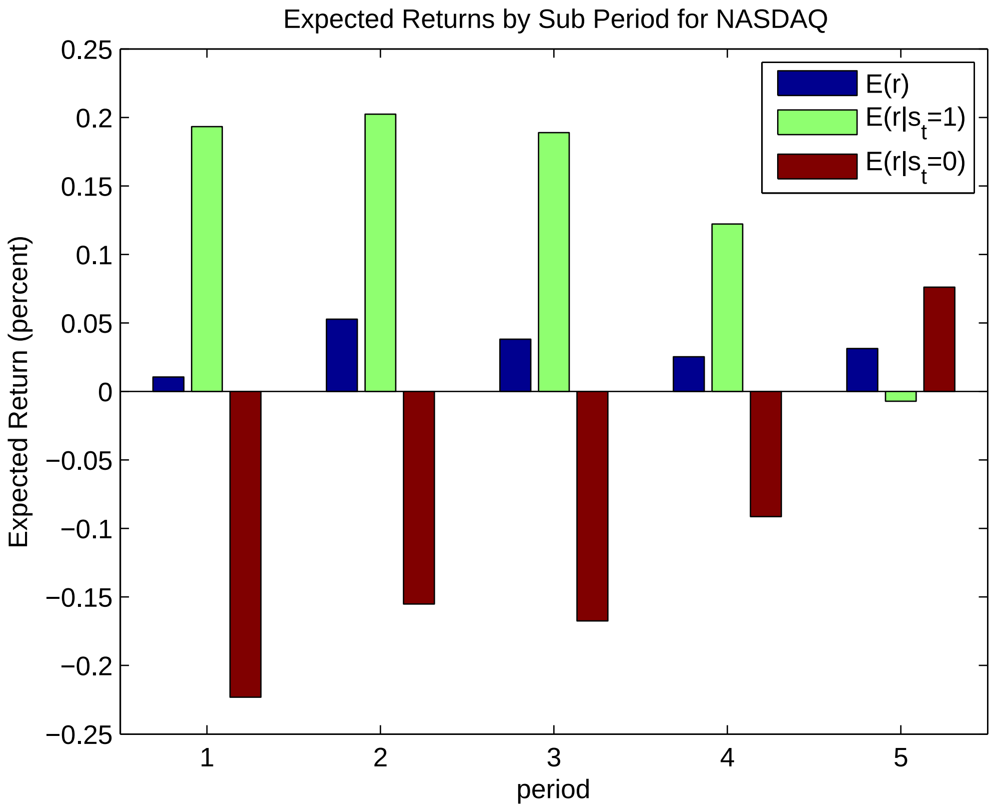

Figure 4.4: State Dependent Mean Return by Sub-Period NASDAQ Composite

The Kalman filter was applied to model returns for both the NASDAQ and S &P 500 datasets using the two regime model with time varying parameters. Maximum Likelihood is used to estimate the noise variances. A similar improvement in model performance is observed for both datasets as was observed for the DOW.

Table 4.4 compares results for different models of the S & P 500.

Table 4.4: Model Fit Comparison S & P 500

Model	Parameter Estimation	Static or Dynamic Parameters	MSE
One State AR(0)	Least Squares	Static	0.9632
Two State AR(0)	Least Squares	Static	0.9594
Two State AR(0)	Kalman Filter	Dynamic	0.9551

Table 4.5 compares results for different models of the NASDAQ Composite.

In both cases, the two regime model with dynamic parameters and Kalman filter tracking

Table 4.5: Model Fit Comparison NASDAQ

Model	Parameter Estimation	Static or Dynamic Parameters	MSE
One State AR(0)	Least Squares	Static	1.6061
Two State AR(0)	Least Squares	Static	1.5913
Two State AR(0)	Kalman Filter	Dynamic	1.5889

outperforms the other models considered.

4.5 Chapter Summary

In this chapter we provided empirical evidence that the parameters of our two state model vary over time by examining sub-periods of the entire dataset. We develop a time-varying two-state model where the model parameters follow a random walk process. We describe how a Kalman filter can be used to estimate the parameters over time. The predictions of the Kalman filter achieve a better fitting model than other candidate models as indicated by a lower MSE. The demonstrated model improvement supports our assumption that the model parameters are dynamic and that our estimation procedure is effective. The procedure is repeated for the S & P 500 and NASDAQ Composite data and results are similar to the DOW dataset.

Chapter 5

Finite State Machine Trading System

This chapter explores a trading system based on the model developed in the previous chapter. When evaluating a trading system it is important use information from the past and present for decision making at any point in time. We make an effort to use as little foresight as possible. The usefulness of the Kalman filter for this application is clear. We estimate the values of the model parameters at each point in time with one-step ahead forecasts and use the forecasts to make investing decisions. We utilize a finite state machine with two states. We make decisions based on the estimated values of the model parameters and the current state of the system. We test the system over a long period demonstrating exceptional performance compared to a buy and hold strategy over the same period.

5.1 Expected Returns

Our model estimates the logarithmic return conditional on the value of threshold function s_t defined in Equation 5.1.

$$s_t = \begin{cases} 0, & \text{if } r_{t-1} < 0 \\ \\ 1, & \text{if } r_{t-1} \geq 0. \end{cases} \tag{5.1}$$

The time-varying return is expressed very compactly in 5.2.

$$r_t = \phi_{0,t} + \phi_{1,t} s_t + \epsilon_t \tag{5.2}$$

$\phi_{0,t}$ and $\phi_{1,t}$ are parameters to be estimated and ϵ_t is IID Gaussian noise with zero mean. By taking the expected value of both sides of the equation, we arrive at

$$\begin{aligned} E(r_t|s_t = 0) = E(\phi_{0,t}) = \phi_{0,t} \\ E(r_t|s_t = 1) = E(\phi_{0,t}) + E(\phi_{1,t}) = \phi_{0,t} + \phi_{1,t} \end{aligned} \tag{5.3}$$

By determining the parameters $\phi_{0,t}$ and $\phi_{1,t}$, we get a prediction of the return given our knowledge of the state. We will use the predicted expected return to make investment decisions. The value of s_t is known at time t because it only depends on the value of the return at $t - 1$. As a result we use the values of $E(r_t|s_t = 0)$ and $E(r_t|s_t = 1)$ to guide our investing decisions. If the system is in state $s_t = 1$ and $E(r_t|s_t = 1) > 0$ we invest in the Dow Index for that day. Similarly, if the system is in state $s_t = 0$ and $E(r_t|s_t = 0) > 0$ we invest in the Dow Index for that day. Conversely, if the expected return given the value of s_t is negative, we do not invest in the market and instead hold our assets in cash assumed to generate zero interest.

The purpose of this exercise is to apply our model to actual financial decision making. This type of strategy is essentially a market timing strategy, determining which days to invest in the stock market and which days to hold cash instead.

The study undertaken in this thesis improves upon previous work by using time varying parameters and a Kalman filter for estimation and prediction. Our trading system test then becomes more realistic because we try to only use information available to us at the time a decision is made.

5.2 Finite State Machine

A finite-state machine is an abstract machine that can be in one of a finite number of states [32]. The machine is in only one state at a time. The state at any given time is called the current state. It can change from one state to another when initiated by a triggering event or condition, this is called a transition. A particular FSM is defined by a list of its states, and the triggering condition for each transition. The finite state machine developed here is based on our two state model with Kalman filter parameter tracking. The Kalman filter provides us with one step-ahead forecasts. The forecasts are used to make investment decisions. When the expected return predicted by the model is positive, we invest, when negative, we hold cash. To represent these trading actions as a FSM we use only two states: a Buy state and a Sell state.

Buy State

$(s_t = 0 \text{ AND } E(r_t|s_t = 0) \geq 0) \text{ OR } (s_t = 1 \text{ AND } E(r_t|s_t = 1) \geq 0)$

Sell State

$(s_t = 0 \text{ AND } E(r_t|s_t = 0) < 0) \text{ OR } (s_t = 1 \text{ AND } E(r_t|s_t = 1) \leq 0)$

State transitions are triggered when a new observation r_t is received. We then evaluate s_t and use a Kalman filter to predict $E(r_t|s_t = 0)$ and $E(r_t|s_t = 0)$. We use these values to find the current trading state, either **BUY** or **SELL**. Trading actions are associated with entry and exit from the two states.

The trading logic is implemented simply using the two-state FSM. In the two state FSM, when entering the **Buy State**, the trading decision Buy Index is made. When entering the **Sell State**, the trading decision Sell Index is made. State transitions are triggered by the predictions of our model and the state of the system. We perform rational trading actions given the state and the expected return given that state.

A visual summary of the FSM used for the trading system is given in Figure 5.1.

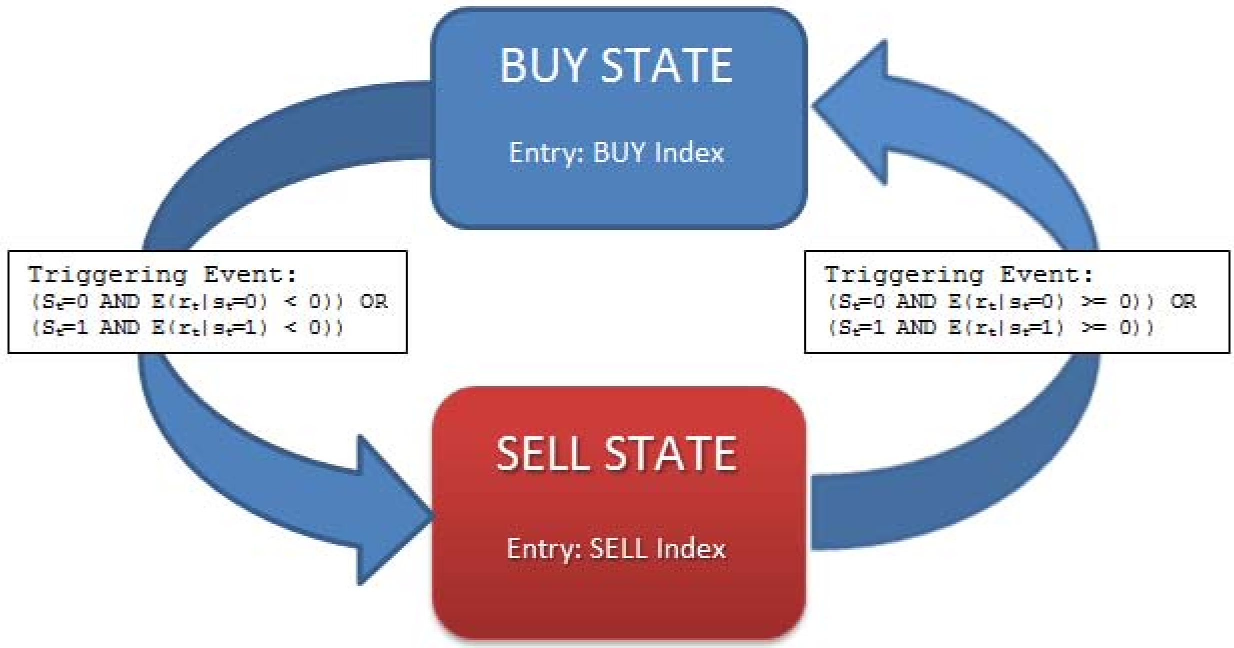

Figure 5.1: Finite State Machine Trading System

5.3 Trading System Results DOW

We implement the FSM described in the previous section for DOW Index data from January 1928 to July 2012. Our results indicate exceptional performance for the trading system over the period studied.

We implement two different trading systems to compare the results for our market timing strategy. The benchmark strategy is the simple buy and hold strategy employed by long term investors. The investor buys the DOW Jones Industrial Average at the beginning of the dataset and sells at the end.

Over the same period, we evaluate the performance of our FSM trading machine based on the two-regime model with time-varying parameters tracked by the Kalman filter. This is the trading system depicted in Figure 5.1.

The results show grossly superior performance for the dynamic model with Kalman filter parameter tracking compared to Buy and Hold. Table 5.1 compares the mean return, number of days invested, standard deviation of return and cumulative return for each of the two strategies.

Table 5.1: Kalman Filter Trading System Results DOW

Strategy	Mean Return(%)	Standard Dev	days invested	Cumulative Return
FSM Trading System	0.0808	1.02	11799	957000%
BUY and HOLD	0.0190	1.16	20934	5500%

What is remarkable is that the system invests just over half of the days from a series of returns with a positive expected value and dramatically outperforms the passive buy and hold strategy and the static parameter two state model. This demonstrates the effectiveness of this system in identifying trading opportunities with positive expected returns. The returns could be increased further by investing in a short term government bonds instead of cash when not in the market.

As a further demonstration of the trading system performance we compare the growth of one dollar invested using each strategy over the period studied in Figure 5.2.

It is clear from the plot that the dynamic strategy based on our two-regime model outperforms the benchmark in terms of cumulative returns.

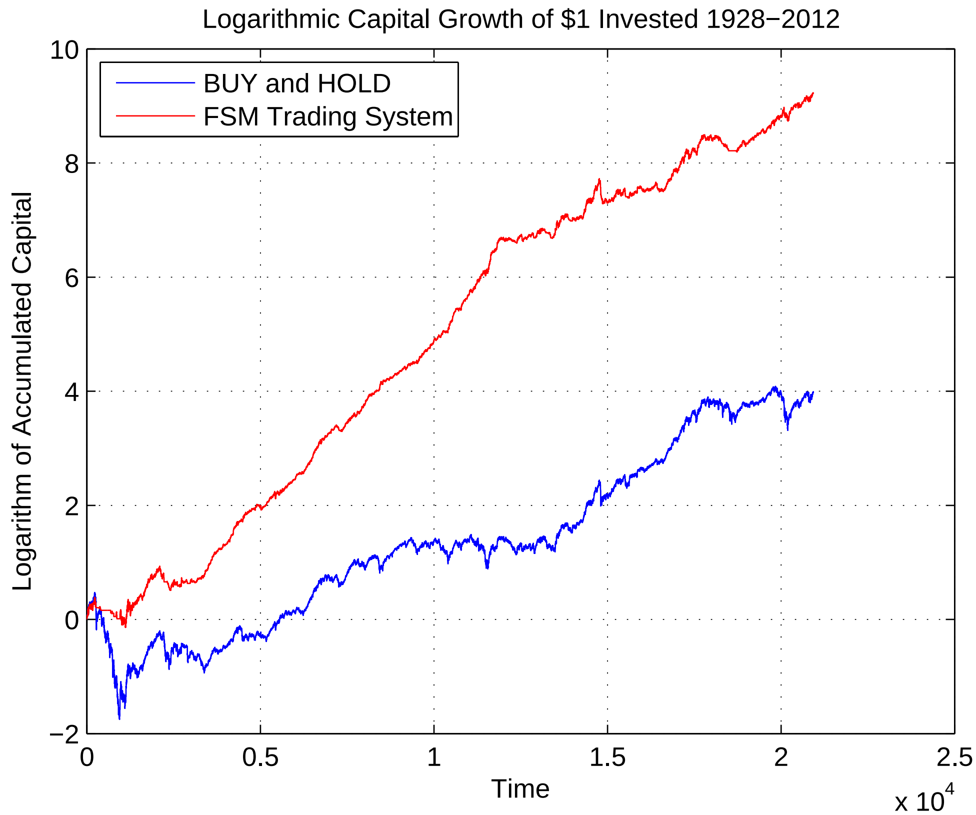

Figure 5.2: Capital Growth by Strategy DOW

5.4 Trading System Results NASDAQ and S& P

A similar trading system test was performed for the NASDAQ Composite and S& P 500 datasets. Table 5.2 shows the results for two strategies tested on the NASDAQ dataset. The strategy based on the two-regime model outperforms buy and hold by a significant margin.

Table 5.2: Kalman Filter Trading System Results NASDAQ

Strategy	Mean Return(%)	Standard Dev	days invested	Cumulative Return
FSM Trading System	0.13	1.01	6252	268300%
BUY and HOLD	0.032	1.27	10351	2830%

We compare the growth of one dollar invested using each strategy over the period studied in Figure 5.3.

It is clear from the plot that the FSM trading strategy based on our two-regime model

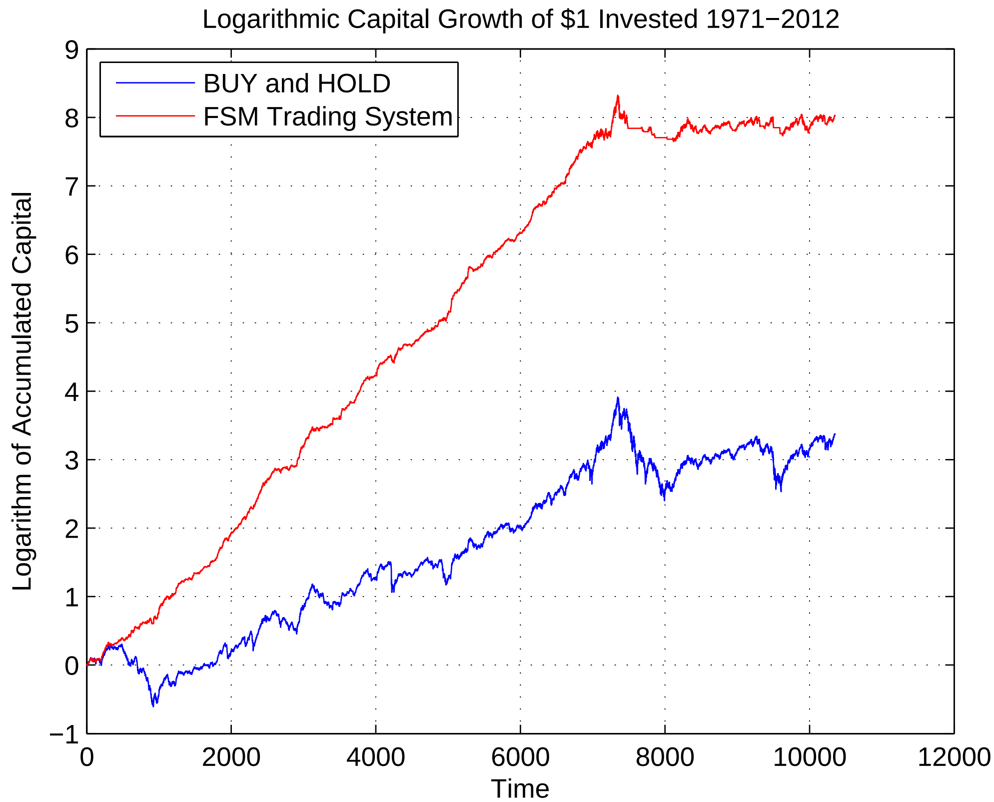

Figure 5.3: Capital Growth by Strategy NASDAQ

outperforms the benchmark.

Table 5.3 shows the results the two strategies tested on the S & P 500 dataset. The Finite State Machine strategy based on the two-regime model outperforms buy and hold considerably.

We compare the growth of one dollar invested using each strategy over the period studied in Figure 5.4.

It is clear from the plot that the dynamic strategy based on our two-regime model outperforms the benchmark in terms of cumulative returns.

Table 5.3: Kalman Filter Trading System Results S& P 500

Strategy	Mean Return(%)	Standard Dev	days invested	Cumulative Return
FSM Trading System	0.1	0.90	9051	888500%
BUY and HOLD	0.028	0.98	15632	8000%

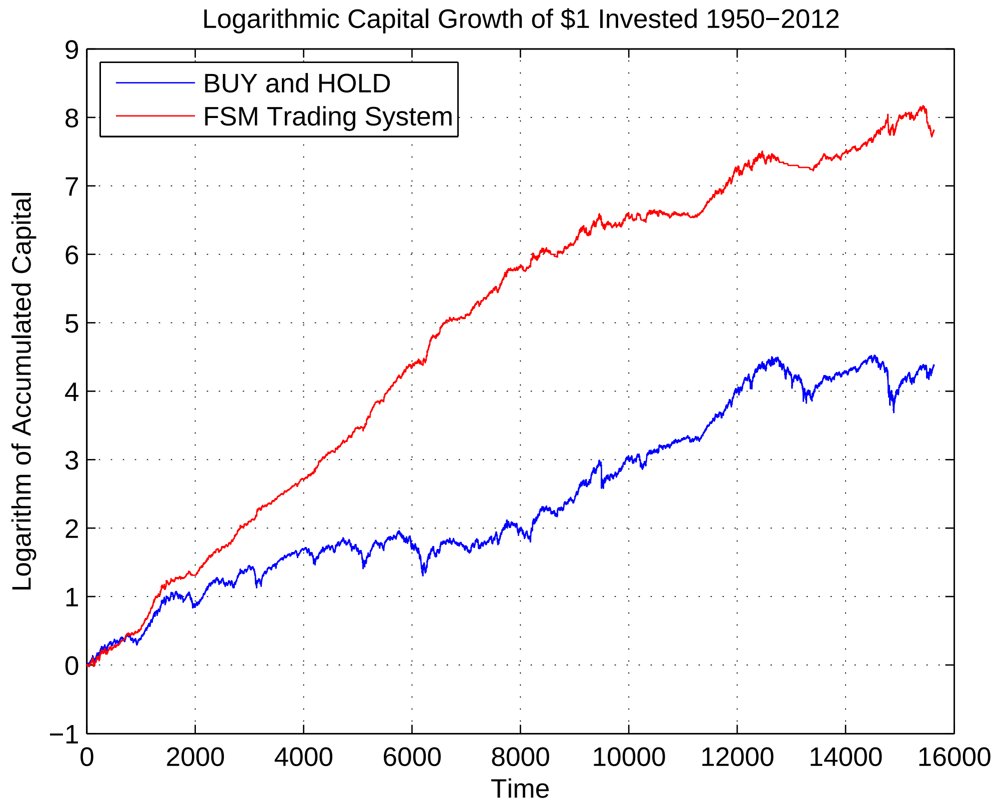

Figure 5.4: Capital Growth by Strategy S & P 500

5.5 Chapter Summary

In this chapter, we presented a Finite State Machine trading system that makes investment decisions given the predictions of our dynamic parameter two state model and the current state of the system. We demonstrate excellent performance on Dow Jones Index data over a long period, outperforming a passive buy and hold by a significant margin. Good results are also found for the NASDAQ Composite and S& P 500 datasets. Our dynamic model identifies trading opportunities with positive expected returns. The FSM trading system uses the predictions of our model to make rational investing decisions by investing only when the expected one-step ahead return is positive. The results demonstrate potentially lucrative results for this system.

Chapter 6

Conclusion and Future Work

In this thesis we presented empirical evidence from several US equity indexes over a long period to support a two-regime model of investment returns. We found compelling evidence that stock index return depends on the sign of the previous day's return. In response to these observations, we created a two regime model. Within each regime the investment returns follow a random walk with drift. The drift parameter is different within each regime. We examine this behaviour over time and found evidence to support time varying drift parameters.

As a result, we built a two regime model with time-varying parameters. Using a Kalman filter, we then track the drift parameters over time, assuming the parameters themselves follow a random walk. We show that the two-state model with time-varying drift parameters outperforms both the single regime random walk with drift and the two regime random walk with drift and constant parameters models. Our finite state machine trading system makes investment decisions based on the predictions of the model and is shown to perform exceptionally well over the time period studied.

Future work should focus on including more complicated variance specifications. Financial time-series have been shown to exhibit fat-tails [33]. GARCH models and stochastic volatility models have been developed to better describe the variance for financial time series. By improving our description of the variance, we could further improve our trading system by altering our position sizing to reflect both the expected return and expected variance.

This may require use of a particle filter in place of the Kalman filter used in this study.

The trading system results assume zero transactions costs and slippage. To implement the trading system discussed in this thesis would require a large number of transactions, each incurring some cost. Further investigation of the FSM trading system could take into consideration these costs. One difficulty is commissions for financial transactions have changed dramatically over the past 80 years. Electronic stock exchanges have reduced trading costs and tightened spreads in the past quarter century. Structural changes like the move to decimal prices on most exchanges from the more traditional fractional pricing also have an effect on trading costs.

Future studies could investigate other strategies based on the model. Perhaps require that the expected return be greater than commission costs or some other value before initiating a trade. The use of leverage could also be employed to increase gains. By investing in treasuries instead of cash then using the interest payments to cover interest costs paid to obtain leverage, the investor may be able to enhance gains further.

Bibliography

[1] T.L. Chong, , and T. Lam, "Predictability of Nonlinear Trading Rules in the US Stock Market," *Quantitative Finance*, vol. 10, no. 9, pp. 1067–1076, 2010.

[2] K S Lim H. Tong, "Threshold Autoregression, Limit Cycles and Cyclical Data," *Journal of the Royal Statistical Society*, vol. 42, no. 3, pp. 245–292, 1980.

[3] H. Tong, *Threshold Models in Non-Linear Time Series Analysis*, Number ISBN: 0387909184. Springer, New York, 1983.

[4] E.B. Hansen, "Inference When a Nuissance Parameter is not Identified Under the Null Hypothesis," *Econometrica*, vol. 64, pp. 413–430, 1996.

[5] E.B. Hansen, "Inference in Tar Models," *Studies in Non-Linear Dynamics and Econometrics*, vol. 2, pp. 1–14, 1997.

[6] E.B. Hansen and B. Seo, "Testing for Two-Regime Threshold Cointegration in Vector Error Correction Models," *Journal of Econometrics*, vol. 110, pp. 293–318, 2002.

[7] J.M. Zakoian, "Threshold Heteroscedastic Models," *Journal of Econometrics, Dynamics and Control*, vol. 18, pp. 931–955, 1994.

[8] W.K. Li M.K.P. So and K. Lam, "A Threshold Stochastic Volatility Model," *Journal of Forecating*, vol. 21, pp. 473–500, 2002.

[9] Q. Wan H. Chen, F. Li and Y. Wang, "Short term load forecasting using regime-switching garch models," in *Power and Energy Society General Meeting, 2011 IEEE*, july 2011, pp. 1 –6.

[10] P. Wong W. Li and K. Lam, "Price Changes and Trading Volume Relationship in the Hong Kong Stock Market," *Asia Pacific Journal of Management*, vol. 7, pp. 25–42, 1990.

[11] Li and Lam, "Modelling Asymmetry in Stock Returns by a Threshold ARCH Model," *The Statistician*, vol. 44, pp. 333–341, 1995.

[12] G. Caginal and H. Laurent, "The Predictive Power of Price Patterns," *Applied Mathematical Finance*, vol. 5, pp. 181–205, 1998.

[13] Lakonishok Brock and LeBaron, "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," *Journal of Finance*, vol. 47, pp. 1731–1764, 1992.

[14] A. Lo, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *Journal of Finance*, vol. 55, no. 4, pp. 1705–1764, 2000.

[15] M.J. Ready, "Profits from Technical Trading Rules," *Financial Management*, vol. 31, no. 3, pp. 43–61, 2002.

[16] H. Markowitz, "Portfolio Selection," *Journal of Finance*, vol. 7, pp. 77–91, 1952.

[17] F.J. Fabozzi and J.C. Francis, "Beta as a Random Coefficient," *Journal of Financial and Quantitive Analysis*, vol. 13, no. 1, pp. 101–116, 1978.

[18] M.K. Kim and J.K. Zumwalt, "An Analysis of Risk in Bull and Bear Markets," *Journal of Financial and Quantitative Analysis*, vol. 14, pp. 1015–1025, 1979.

[19] M. Gastaldi and A.Nardecchia, "The Kalman Filter Approach for Time-Varying β Estimation," *System Analysis Modelling Simulation*, vol. 43, pp. 1033–1042, 2003.

[20] L. L. Scharf J. P. Dugre and C. Gueguen, "Exact Likelihood for Stationary Autoregressive Moving Average Processes," *Signal Processing*, vol. 11, pp. 105–118, 1986.

[21] P. Gustafssonand M. Norrlöf P. Axelsson, U. Orguner, "ML Estimation of Process Noise Variance in Dynamic Systems," , no. LiTH-ISY-R-2969, Oct 2010.

[22] X.P. Zhang T. Rajbhandary, "Piecewise constant modeling and tracking of systematic risk in financial market," *Ryerson University*, p. MASc Thesis, 2011.

[23] M. Feder, "Gambling Using a Finite State Machine," *IEEE Transactions on Information Theory*, vol. 37, no. 5, pp. 1459–1465, 1991.

[24] J. L. Kelly Jr, "A New Interpretation of Information Rate," *Bell System Technical Journal*, vol. 35, no. 5, 1956.

[25] A. Shleifer R.J. Waldmann L.H. Summers, J. de Jong, "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *The Journal of Finance*, vol. 45, no. 2, pp. 379–395, 1990.

[26] J.D. Petruccelli N. Davies, J. Pemberton, "An Automatic Procedure for Identification, Estimation and Forecasting Univariate Self Exciting Threshold Autoregressive Models," *The Statistician*, vol. 37, pp. 199–204, 1988.

[27] L. Bachelier, "Theorie de la Speculation," *Annales Scientifiques de lEcole Normale Superieure*, vol. 3, no. 17, pp. 21–86, 1900.

[28] R.S. Tsay, *Analysis of Financial Time Series*, Wiley, New Jersey, 3rd edition, 2005.

[29] R. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[30] A.C. Harvey, *Forecasting Structural Time Series Models and the Kalman Filters*, Cambridge University Press, New York, 1 edition, 1989.

[31] B.L. Welch, "The Generalization of Students Problem When Several Different Population Variances are Involved," *Biometrika*, vol. 34, no. 1-2, pp. 28–35, 1947.

[32] T.L. Booth, *Sequential Machines and Automata Theory*, John WIley and Sons, New York, 1 edition, 1967.

[33] Y. Fang and D. Xu, "The Predictability of Asset Returns: An Approach Combining Technical Analysis and Time Series Forecasts," *Int. J. of Forecasting*, vol. 19, pp. 369–385, 2003.

Glossary

S& P 500:

A US equity index composed of the 500 largest US companies by market capitalization.

NASDAQ Composite:

A US equity index composed of the stocks listed on the NASDAQ stock exchange.

This index is dominated by technology companies and is weighted by market capitalization.

DOW Jones Industrial Average:

Price weighted stock index composed of US large cap stocks selected to the index by DOW Jones.

Finite State Machine (FSM):

An abstract machine that involves defining states and state transitions to accomplish some task.

Threshold Autoregressive (TAR):

Non-linear model that uses a threshold function to separates a time series into regimes and models each regime with a different autoregressive model.

Self-Excited Threshold Autoregressive (SETAR):

A TAR model where the threshold function is based on lagged values of the time-series itself.

VITA

NAME:	Timothy Little
PLACE OF BIRTH:	Oshawa, Ontario
YEAR OF BIRTH:	1979
POST-SECONDARY EDUCATION AND DEGREES:	University of Toronto
	1998-2003 Bachelor of Applied Science Engineering Physics
RELATED WORK EXPERIENCE:	Proprietary Trader US Equities Swift Trade Securities 2005-2008
	Data Analyst Detectent Inc 2008-2009
Publications:	Little,T., Zhang, X.P., "Regime-Switching Models for US Equities with Kalman Filter Tracking", Journal of Empirical Finance. (To be submitted)