SOME ASPECTS OF STATISTICAL ANALYSIS OF FINANCIAL TIME SERIES DATA

YUFENG LIN

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS
YORK UNIVERSITY
TORONTO, ONTARIO

February 2018

Abstract

Technical analysis is widely adopted by investors in practice. Moving average strategy is the simplest and most popular trading rule. This simple moving average strategy suffers a well-known drawback since its allocation is always either 100% or 0%. This rule is independent of the investor's risk tolerance level which is widely considered as an important factor in any investment activity. We first introduce an investor's specific risk tolerance into the moving average strategy. We then propose a single-asset generalized moving average crossover (SGMA) strategy and a multiple-asset generalized moving average crossover (MGMA) strategy. The SGMA and MGMA strategies allocate wealth among risky-assets and risk-free asset with the risk tolerance specified by investor. These trading strategies are evaluated on both simulation data and high-frequency exchange-traded fund (ETF) data. It is evident that both the SGMA and MGMA strategies can significantly increase the investor's expected utility of wealth and expected wealth.

Movements of stocks or equity indices are very important information for an investment decision. Empirical studies illustrate that the movements switch among different regimes or states. The Markov regime-switching model has important applications to this type analysis. However, parameters estimated under normality assumption might not be stable and the corresponding change-point detection algorithm might face some challenges when either the empirical distribution is heavy-tailed or observed data contain outliers. We relax the normality assumption and propose a generalized regime-switching (generalized RS) model. We then improve the corresponding change-point detection algorithm by using the generalized RS model. The change-point detection algorithm using the generalized RS model is tested on both simulation data and Hang Seng monthly index data from January 1988 to March 2015. Simulation studies show that the change-point detection algorithm using the generalized RS model can improve accuracy of identifying change-points when either the empirical distribution is heavy-tailed or observed data contain outliers. It is also evident that the identified change-points on Hang Seng monthly index data match the observed market behaviors.

Acknowledgements

First and foremost, I would like to express my deepest gratitude towards my supervisors, Professor Yuehua Wu and Professor Xiaogang Wang. Without their valuable guidance, caring, patience and persistent help, this dissertation would not have been possible.

I would like to extend my sincere appreciation to Professor Wei Liu and Professor Michael Chen for their great help as members of my supervisory committee. I would like to thank Professor Mike Zabrocki, Professor Alexey Kuznetsov, and Primrose Miranda for all their helps. I also would like to thank all my friends for their help, encouragement, and support. Special thanks go to Dr. Baifang Xing and Dr. Rong Zhu for their suggestions on my dissertation.

Last but not least, I deeply thank my parents, whose love and guidance are with me in whatever I pursue. I also thank my parents in-law, my sister and her husband. They were always supporting me with their best wishes. I wish to thank my loving and supportive wife, and my wonderful children, who provide unending inspiration.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Some Aspects of Single-Asset Portfolio Moving Average Strategy

Technical analysis is widely adopted by investors in practice. They believe that historical data contain important information that can be used to predict future movements of the market. Many empirical evidence including predicted performance of stock return demonstrates the usefulness of technical analysis (see Brock *et al.*, 1992; Lo and Mackinlay, 1999; Ang and Bekaert, 2006; Campbell and Thompson, 2008). Among many technical analysis techniques, moving average strategy is the simplest and most popular trading rule. Brock *et al.* (1992) appear to be the first to provide strong evidence of profitability by using moving average technique in analyzing daily Dow Jones Industrial Average (DJIA) data. Lo *et al.* (2000) provide future evidence based on different time series from financial market. These studies generate further research interests on moving average strategies. However, most of studies have been focused on validating the strategy using different data sets. The conclusions are mixed and inconclusive (see Ito, 1999; Gunaskarage and Power, 2001; Chang *et al.*, 2004; Tabak and Lima, 2009; Mohr *et al.*, 2013). Recent studies are focused on predictive power of moving average technique (see Burghandt and Walls, 2011; Cespa and Vives, 2012; Neely *et al.*, 2014).

The moving average in technical analysis follows an all-or-nothing investment strategy: When a *buy* signal is triggered by MA, the investor should allocate all of his/her wealth to the stock of interest; when a *sell* signal is triggered by MA, the investor should allocate nothing of wealth into the stock by selling all the current holdings. This simple moving average strategy suffers a well-known drawback since its allocation is always either 100% or 0%. This is independent of the investor's risk tolerance level which is widely considered as an important factor in any investment activity. We believe that the allocation amount should be a function of the investor specified risk tolerance and could lead to a much favorable investment outcome. Zhu and Zhou (2009) provide the first theoretical analysis for the simplest moving average strategy. However, their study focuses on how technical analysis such as moving average strategy can add value to commonly used allocation rules that invest fixed proportions of wealth in stock. To our best knowledge, there is no theoretical study in literature about a trading strategy that involves the risk tolerance specified by investor. This motives us to propose a single-asset generalized moving average crossover (SGMA) strategy in chapter 2.

1.2 Some Aspects of Multi-Asset Portfolio Moving Average Strategy

In order to avert significant loss, many investors might be more interested in an investment based on diversified investment portfolio, which contains more than one risky asset. They might want to allocate the wealth not only between risk-free asset and risky asset, but also among different risky assets. They might be interested in optimal trading strategies when they have enough capital to invest more than two assets. However, most of asset allocation studies focus on finding optimal portfolio choice under different modeling processes (see Buraschi *et al.*, 2010; Chiu and Wong, 2014; Lioui and Poncet, 2016; Legendre and Togola, 2016; Wu *et al.*, 2017). They do not study the optimal allocation in the context of using technical analysis strategies for multiple assets allocation.

Therefore we take a different approach and focus on finding optimal trading strategies based on technical analysis. The moving average is the simplest and most popular technical analysis trading rule. However, it faces difficulty when there are more than one investment signals because it is an all-or-nothing investment strategy.

The common approach is to assign equal weights when allocating the wealth among the risky assets, which is not optimal. This motives us to continue our study for multi-asset portfolio and propose a multi-asset generalized moving average crossover (MGMA) strategy in chapter 3.

1.3 Some Aspects of Markov Regime-Switching Model

In order to generate a much favorable investment outcome, many investors are interested in movements of stocks or equity indices that are widely considered as basic information in an investment activity. Many empirical studies illustrate that the movements of stocks or equity indices switch among different regimes or states. The Markov regime-switching model introduced by Hamilton (1989) and Hamilton and Susmel (1994) has important applications. The Markov regime-switching model is very useful because it allows model parameters to take different values to reflect intrinsic nature of different regimes. There are numerous applications of the Markov regime-switching model in economic or financial time series analysis (see Guérin and Marcellino, 2013; Zou and Chen, 2013; Bauwens *et al.*, 2014; Chen *et al.*, 2014).

Many researchers believe that market behaviors can be captured through probabilistic switching models for different regimes. Hardy (2001) compares different models and finds that a two-state Markov regime-switching log-normal (RSLN) model is able to fit S&P 500 monthly index data relatively well. However, parameters estimated under normality assumption might not be stable when either the empirical distribution is heavy-tailed or observed data contain outliers. Guo *et al.* (2011) find that using a single regime-switching model to model entire time series might not work well as financial market behavior changes at some points in time. They find that to use a regime-switching model with different parameters for time series before and after those time points can explain the financial market behavior relatively well. Hence, they define change-points as the time points that segment a time series if data in two neighboring segments are modeled by the same model with different parameters or different models. They also propose a change-point detection algorithm to identify change-points. However, the change-point detection algorithm faces some challenges when either the empirical distribution is heavy-tailed or observed data contain outliers. This motives us to propose a generalized regime-switching (generalized RS) model and use the generalized RS model to improve the change-point detection algorithm in chapter 4.

2 Statistical Modeling and Single-Asset Generalized Moving Average Crossover (SGMA) Strategy

In this chapter, we study moving average strategy for single-asset portfolio. Notice that the moving average in technical analysis suffers a well-known drawback since its allocation is always either 100% or 0%. We introduce the investor specified risk tolerance into the moving average strategy and propose a single-asset generalized moving average crossover (SGMA) strategy. The SGMA strategy is not an all-or-nothing investment strategy and involves the risk tolerance specified by investor. The SGMA strategy not only can increase the investor's expected utility of wealth, but also can increase the investor's expected wealth.

Chapter 2 is organized as follows. Section 2.1 introduces the model with the SGMA strategy. Section 2.2 provides all preliminary Lemmas for analytic results. We present main analytic results in section 2.3. An investment algorithm for single-asset portfolio is given in section 2.4. Section 2.5 and 2.6 provide simulation studies and real data analysis. The conclusion is presented in section 2.7.

2.1 The Model and The SGMA Strategy

Suppose that there are two assets in the market. First one is a risk-free asset. For example, cash or money market account with a constant interest rate r. Second one is a risky asset. For example, a stock or index representing the aggregated equity market. A single-asset portfolio only contains one risky asset, and wealth can be only allocated between one risk-free asset and one risky asset.

We follow Zhu and Zhou (2009) to define the model. The model follows fundamental model setting of financial research. Suppose that the price of risk-free asset P_t^f at any time t satisfies

$$dP_t^f = rP_t^f dt, \tag{2.1}$$

and the price of risky asset (stock) P_t at any time t satisfies

$$dP_t = (\mu_0 + \mu_1 X_t) P_t dt + \sigma_p P_t dB_t, \tag{2.2}$$

where μ_0, μ_1, σ_p are parameters, B_t is one-dimensional standard Brownian motion, and X_t is a predictive variable that could help to predict stock return. By Keim and Stambaugh (1986) and Stambaugh (1999), X_t is assumed to be a stationary process for $t \geq 0$ and satisfies

$$dX_t = (\theta_0 + \theta_1 X_t)\, dt + \sigma_x dZ_t, \tag{2.3}$$

where $\theta_0, \theta_1, \sigma_x$ are parameters, θ_1 is negative to ensure that X_t is a mean-reverting process, Z_t is one-dimensional standard Brownian motion correlated with B_t. $\rho \in [-1, 1]$ is a corresponding correlation coefficient between B_t and Z_t.

Let us first recall the moving average crossover (MA) strategy for single-asset portfolio. Suppose that P_t is the stock price and Y_t is the log transformed stock price, i.e., $Y_t = \log P_t$. Let $h > 0$ be a lag or lookback period. By Zhu and Zhou (2009), a continuous time version of the moving average of the log transformed stock price at any time t is defined as

$$M_t^{(h)} = \frac{1}{h} \int_{t-h}^{t} Y_u du, \tag{2.4}$$

i.e., the average log transformed stock price over time period $[t-h, t]$. They choose to use this definition because (1) the distribution of the moving average on original stock price is very complex and difficult to analyze, and (2) this definition is tractable for explicit solutions. Therefore, we also use this moving average definition in our research. Let $M_t^{(s,l)}$ be difference between $M_t^{(s)}$ and $M_t^{(l)}$, where $s > 0$ is a short term lookback period and $l > s$ is a long term lookback period, i.e.,

$$M_t^{(s,l)} = M_t^{(s)} - M_t^{(l)}. \tag{2.5}$$

Then the MA strategy $\tau_t = \tau(M_t^{(s,l)})$ is defined as

$$\tau_t = \begin{cases} 0, & \text{if } M_t^{(s,l)} \in \tilde{\Omega}_1, \\ \\ 1, & \text{if } M_t^{(s,l)} \in \tilde{\Omega}_2. \end{cases} \tag{2.6}$$

where $\tilde{\Omega}_i$ is defined as

$$\tilde{\Omega}_i = \begin{cases} (-\infty, 0)\,, & \text{if } i = 1, \\ \\ [0, \infty)\,, & \text{if } i = 2. \end{cases}$$

Now let us introduce the investor specified risk tolerance $\epsilon > 0$ into the moving average strategy and define the SGMA strategy. Let δ_1 and δ_2 be asset allocation parameters, which can be interpreted as long portion of stock ($\delta_i \geq 0$) and short portion of stock ($\delta_i < 0$). Define Ω_i as

$$\Omega_i = \begin{cases} (-\infty, -\epsilon)\,, & \text{if } i = 1, \\ \\ [-\epsilon, 0)\,, & \text{if } i = 2, \\ \\ [0, \epsilon]\,, & \text{if } i = 3, \\ \\ (\epsilon, \infty)\,, & \text{if } i = 4. \end{cases} \tag{2.7}$$

Then using $M_t^{(s,l)} = M_t^{(s)} - M_t^{(l)}$ as defined in Equation (2.5), we define the SGMA strategy $\eta_t = \eta(M_t^{(s,l)}, \epsilon)$ as

$$\eta_t = \begin{cases} 0, & \text{if } M_t^{(s,l)} \in \Omega_1, \\ \delta_2, & \text{if } M_t^{(s,l)} \in \Omega_2, \\ \delta_1, & \text{if } M_t^{(s,l)} \in \Omega_3, \\ 1, & \text{if } M_t^{(s,l)} \in \Omega_4. \end{cases} \tag{2.8}$$

This is well defined when $t \geq l$. For $t < l$, we define the SGMA strategy η_t as a pre-chosen constant λ. The SGMA strategy η_t is a market timing strategy that allocates wealth between cash and stock with the risk tolerance specified by the investor. Notice that the MA strategy for single-asset portfolio τ_t is a special case of the SGMA strategy η_t when $\delta_1 = 1$ and $\delta_2 = 0$. It is obvious that the SGMA strategy provides more investment options when either $\delta_1 \neq 1$ or $\delta_2 \neq 0$.

The MA strategy suffers a well-known drawback given that it is an all-or-nothing investment strategy. The SGMA strategy can overcome the drawback by finding optimal η_t that maximizes the investor's expected log-utility of wealth

$$\max_{\eta_t} E\left(\log W_T\right), \tag{2.9}$$

subject to a budget constraint

$$\frac{dW_t}{W_t} = rdt + \eta_t\left(\mu_0 + \mu_1 X_t - r\right)dt + \eta_t \sigma_p dB_t, \tag{2.10}$$

given an initial wealth W_0, a constant rate of interest r and an investment horizon T.

2.2 Preliminaries

All preliminary Lemmas in this section are used to derive analytical results in next section. Zhu and Zhou (2009) provide Lemmas 2.1 – 2.7 and 2.11 with limited proof. We verify these Lemmas by providing detailed proof. We also state new Lemmas 2.8 – 2.10 and 2.12 – 2.15. The proofs are also given.

Lemma 2.1 *Let X_t be the predictive variable in the market, then X_t is normally distributed with mean $E\left(X_t\right)$ and variance $Var\left(X_t\right)$, where*

$$X_t = e^{\theta_1 t} X_0 - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1 (t-u)} dZ_u,$$

and

$$E\left(X_t\right) = -\frac{\theta_0}{\theta_1} \quad \& \quad Var\left(X_t\right) = -\frac{\sigma_x^2}{2\theta_1}.$$

Proof.

$$\begin{aligned} d\left(e^{-\theta_1 u} X_u\right) &= -\theta_1 e^{-\theta_1 u} X_u du + e^{-\theta_1 u} dX_u \\ &= -\theta_1 e^{-\theta_1 u} X_u du + e^{-\theta_1 u}\left[\left(\theta_0 + \theta_1 X_u\right) du + \sigma_x dZ_u\right] \\ &= \theta_0 e^{-\theta_1 u} du + e^{-\theta_1 u} \sigma_x dZ_u, \end{aligned}$$

then

$$e^{-\theta_1 u}X_u \mid_0^t = e^{-\theta_1 t}X_t - X_0 = \int_0^t d\left(e^{-\theta_1 u}X_u\right) = \int_0^t \left[\theta_0 e^{-\theta_1 u}du + e^{-\theta_1 u}\sigma_x dZ_u\right]$$
$$= \int_0^t \theta_0 e^{-\theta_1 u}du + \int_0^t e^{-\theta_1 u}\sigma_x dZ_u = -\frac{\theta_0}{\theta_1}\left(e^{-\theta_1 t} - 1\right) + \int_0^t e^{-\theta_1 u}\sigma_x dZ_u,$$

which implies

$$X_t = e^{\theta_1 t}X_0 - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1(t-u)}dZ_u.$$

Given $E\left(dZ_u\right) = 0$ and $Var\left(dZ_u\right) = du$, it follows that X_t is normally distributed with mean

$$E\left(X_t\right) = E\left[e^{\theta_1 t}X_0 - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1(t-u)}dZ_u\right]$$
$$= e^{\theta_1 t}E\left(X_0\right) - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1(t-u)}E\left(dZ_u\right)$$
$$= e^{\theta_1 t}E\left(X_0\right) - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right),$$

and variance

$$Var\left(X_t\right) = Var\left[e^{\theta_1 t}X_0 - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1(t-u)}dZ_u\right]$$
$$= \left(e^{\theta_1 t}\right)^2 Var\left(X_0\right) + Var\left(-\frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right)\right) + \sigma_x^2 \int_0^t \left(e^{\theta_1(t-u)}\right)^2 Var\left(dZ_u\right)$$
$$= e^{2\theta_1 t}Var\left(X_0\right) + 0 + \sigma_x^2 \int_0^t e^{2\theta_1(t-u)}du$$
$$= e^{2\theta_1 t}Var\left(X_0\right) - \frac{\sigma_x^2}{2\theta_1}\left(1 - e^{2\theta_1 t}\right),$$

where $E\left(X_0\right)$ and $Var\left(X_0\right)$ are mean and variance of X_0. Under stationarity condition, i.e., $E\left(X_0\right) = E\left(X_t\right)$ and $Var\left(X_0\right) = Var\left(X_t\right)$, when $\theta_1 < 0$,

$$E\left(X_t\right) = -\frac{\theta_0}{\theta_1} \quad \& \quad Var\left(X_t\right) = -\frac{\sigma_x^2}{2\theta_1}.$$

The Lemma is proved.

Lemma 2.2 *Let* $Y_t = \log P_t$ *be the log transformed stock price, then* Y_t *is normally distributed with mean* $E(Y_t)$*, where*

$$Y_t = Y_0 + \int_0^t \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \sigma_p B_t,$$

and

$$E(Y_t) = Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right) t.$$

Proof.

$$\frac{dP_u}{P_u} = (\mu_0 + \mu_1 X_u)\, du + \sigma_p dB_u,$$

and $(du)^2 = o(du), dudB_u = o(du)$ and $(dB_u)^2 = du$, then

$$\left(\frac{dP_u}{P_u}\right)^2 = (\mu_0 + \mu_1 X_u)^2 (du)^2 + 2\sigma_p dB_u (\mu_0 + \mu_1 X_u)\, du + \sigma_p^2 (dB_u)^2 = \sigma_p^2 du.$$

Therefore

$$\begin{aligned} d(Y_u) = d(\log P_u) &= [\log P_u]' \, dP_u + \frac{1}{2} [\log P_u]'' \, (dP_u)^2 = \frac{dP_u}{P_u} - \frac{1}{2}\left(\frac{dP_u}{P_u}\right)^2 \\ &= (\mu_0 + \mu_1 X_u)\, du + \sigma_p dB_u - \frac{1}{2}\sigma_p^2 du = \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \sigma_p dB_u, \end{aligned}$$

which implies

$$\begin{aligned} \log P_u \mid_0^t = \log P_t - \log P_0 = Y_t - Y_0 &= \int_0^t d(\log P_u) \\ &= \int_0^t \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \int_0^t \sigma_p dB_u \\ &= \int_0^t \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \sigma_p B_t. \end{aligned}$$

Therefore

$$Y_t = Y_0 + \int_0^t \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \sigma_p B_t.$$

Given that $E(B_t) = 0$, by Lemma 2.1, it follows that Y_t with mean

$$\begin{aligned}
E(Y_t) &= E\left[Y_0 + \int_0^t (\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2) du + \sigma_p B_t\right] \\
&= E(Y_0) + \int_0^t \left(\mu_0 + \mu_1 E(X_u) - \frac{1}{2}\sigma_p^2\right) du + \sigma_p E(B_t) \\
&= Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right) t.
\end{aligned}$$

The Lemma is proved.

Lemma 2.3 *Let $M_t^{(h)}$ be the moving average based on lookback period $h > 0$, then $M_t^{(h)}$ is normally distributed with mean $E\left(M_t^{(h)}\right)$, where*

$$E\left(M_t^{(h)}\right) = Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right)\left(t - \frac{h}{2}\right).$$

Proof.

By Lemma 2.2,

$$\begin{aligned}
E\left(M_t^{(h)}\right) &= E\left[\frac{1}{h}\int_{t-h}^t Y_u du\right] \\
&= \frac{1}{h}\int_{t-h}^t E(Y_u) du = \frac{1}{h}\int_{t-h}^t \left(Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right) u\right) du \\
&= \frac{1}{h}\int_{t-h}^t Y_0 du + \frac{1}{h}\int_{t-h}^t \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right) u du \\
&= Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right)\frac{1}{2h}\left(t^2 - (t-h)^2\right) \\
&= Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right)\left(t - \frac{h}{2}\right).
\end{aligned}$$

The Lemma is proved.

Lemma 2.4 *Let Z_u and B_v be two-dimensional standard Brownian motion. Let ρ be the correlation between Z_t and B_t for any time t, then $Cov\left(Z_u, B_v\right)$ and $Cov\left(dZ_u, B_v\right)$ for any u, v are*

$$Cov\left(Z_u, B_v\right) = \rho \min\left(u, v\right),$$

and

$$Cov\left(dZ_u, B_v\right) = \begin{cases} \rho du, & \text{if } u < v, \\ 0, & \text{otherwise.} \end{cases}$$

Proof.

Given that (Z_u, B_u) is two-dimensional standard Brownian motion with correlation coefficient ρ, Z_u can be represented as $\rho B_u + \sqrt{1-\rho^2} B'_u$, where B_u and B'_u are independent, then

$$\begin{aligned} Cov\left(Z_u, B_v\right) &= Cov\left(\rho B_u + \sqrt{1-\rho^2} B'_u, B_v\right) \\ &= \rho Cov\left(B_u, B_v\right) + \sqrt{1-\rho^2} Cov\left(B'_u, B_v\right) \\ &= \rho \min\left(u, v\right), \end{aligned}$$

and

$$Cov\left(dZ_u, B_v\right) = Cov\left(Z_{u+du} - Z_u, B_v\right) = Cov\left(Z_{u+du}, B_v\right) - Cov\left(Z_u, B_v\right)$$
$$= \rho \min\left(u + du, v\right) - \rho \min\left(u, v\right)$$
$$= \begin{cases} \rho du, & \text{if } u < v, \\ 0, & \text{otherwise.} \end{cases}$$

The Lemma is proved.

Lemma 2.5 *Let X_t be the predictive variable in the market. Let B_v be one-dimensional standard Brownian motion, then $Cov\left(B_v, X_t\right)$ for any v, t is*

$$Cov\left(B_v, X_t\right) = \begin{cases} -\frac{\sigma_x \rho}{\theta_1}\left(e^{\theta_1 (t-v)} - e^{\theta_1 t}\right), & \text{if } v \leq t, \\ -\frac{\sigma_x \rho}{\theta_1}\left(1 - e^{\theta_1 t}\right), & \text{if } v > t. \end{cases}$$

Proof.

Since B_v, X_0 are independent, then $Cov\left(B_v, X_0\right) = 0$, by Lemma 2.1 and Lemma 2.4,

$$Cov\left(B_v, X_t\right) = Cov\left(B_v, e^{\theta_1 t} X_0 - \frac{\theta_0}{\theta_1}\left(1 - e^{\theta_1 t}\right) + \sigma_x \int_0^t e^{\theta_1 (t-u)} dZ_u\right)$$
$$= \sigma_x Cov\left(B_v, \int_0^t e^{\theta_1 (t-u)} dZ_u\right) = \sigma_x \int_0^t e^{\theta_1 (t-u)} Cov\left(B_v, dZ_u\right).$$

For $v \leq t$, then

$$Cov\left(B_v, X_t\right) = \sigma_x \int_0^v e^{\theta_1 (t-u)} Cov\left(B_v, dZ_u\right) + \sigma_x \int_v^t e^{\theta_1 (t-u)} Cov\left(B_v, dZ_u\right)$$
$$= \sigma_x \int_0^v e^{\theta_1 (t-u)} \rho du + 0 = -\frac{\sigma_x \rho}{\theta_1}\left(e^{\theta_1 (t-v)} - e^{\theta_1 t}\right),$$

and for $v > t$, then

$$Cov\left(B_v, X_t\right) = \sigma_x \int_0^t e^{\theta_1(t-u)}\rho du = -\frac{\sigma_x \rho}{\theta_1}\left(1 - e^{\theta_1 t}\right).$$

The Lemma is proved.

Lemma 2.6 *Let Z_t be one-dimensional standard Brownian motion, then for any u, v, $Cov\left(dZ_u, dZ_v\right) = 0$.*

Proof.

For $u \leq v$, then

$$Cov\left(dZ_u, dZ_v\right) = Cov\left(Z_{u+du} - Z_u, Z_{v+dv} - Z_v\right)$$

$$= Cov\left(Z_{u+du}, Z_{v+dv}\right) - Cov\left(Z_{u+du}, Z_v\right) - Cov\left(Z_u, Z_{v+dv}\right) + Cov\left(Z_u, Z_v\right)$$

$$= (u + du) - (u + du) - u + u = 0.$$

Similarly, for $u > v$, $Cov\left(dZ_u, dZ_v\right) = (v + dv) - v - (v + dv) + v = 0$. The Lemma is proved.

Lemma 2.7 *Let X_t be the predictive variable in the market, then $Cov\left(X_a, X_b\right)$ for any a, b is*

$$Cov\left(X_a, X_b\right) = -\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 |a-b|}.$$

Proof.

By Lemma 2.1,

$$
\begin{aligned}
& Cov\left(X_a, X_b\right) \\
& = Cov\left(e^{\theta_1 a} X_0, e^{\theta_1 b} X_0\right) + Cov\left(\sigma_x \int_0^a e^{\theta_1 (a-u)} dZ_u, \sigma_x \int_0^b e^{\theta_1 (b-v)} dZ_v\right) \\
& = e^{\theta_1 (a+b)} Var\left(X_0\right) + \sigma_x^2 e^{\theta_1 a} e^{\theta_1 b} Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^b e^{-\theta_1 v} dZ_v\right) \\
& = e^{\theta_1 (a+b)} Var\left(X_0\right) + \sigma_x^2 e^{\theta_1 (a+b)} Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^b e^{-\theta_1 v} dZ_v\right).
\end{aligned}
$$

For $a \leq b$, by Lemma 2.6,

$$
\begin{aligned}
& Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^b e^{-\theta_1 v} dZ_v\right) \\
& = Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^a e^{-\theta_1 v} dZ_v + \int_a^b e^{-\theta_1 v} dZ_v\right) \\
& = Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^a e^{-\theta_1 v} dZ_v\right) + Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_a^b e^{-\theta_1 v} dZ_v\right) \\
& = Var\left(\int_0^a e^{-\theta_1 u} dZ_u\right) + \int_0^a e^{-\theta_1 u} \int_a^b e^{-\theta_1 v} Cov\left(dZ_u, dZ_v\right) \\
& = \int_0^a e^{-2\theta_1 u} du + 0 = -\frac{1}{2\theta_1}\left(e^{-2\theta_1 a} - 1\right),
\end{aligned}
$$

similarly, for $a > b$,

$$
Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^b e^{-\theta_1 v} dZ_v\right) = \int_0^b e^{-2\theta_1 v} dv = -\frac{1}{2\theta_1}\left(e^{-2\theta_1 b} - 1\right),
$$

which implies

$$
Cov\left(\int_0^a e^{-\theta_1 u} dZ_u, \int_0^b e^{-\theta_1 v} dZ_v\right) = -\frac{1}{2\theta_1}\left(e^{-2\theta_1 \min(a,b)} - 1\right).
$$

Therefore

$$
\begin{aligned}
Cov\left(X_a, X_b\right) &= e^{\theta_1(a+b)} Var\left(X_0\right) + \sigma_x^2 e^{\theta_1(a+b)} \left[-\frac{1}{2\theta_1} \left(e^{-2\theta_1 \min(a,b)} - 1 \right) \right] \\
&= e^{\theta_1(a+b)} Var\left(X_0\right) - \frac{\sigma_x^2}{2\theta_1} \left[e^{(\theta_1(a+b) - 2\theta_1 \min(a,b))} - e^{\theta_1(a+b)} \right] \\
&= e^{\theta_1(a+b)} Var\left(X_0\right) - \frac{\sigma_x^2}{2\theta_1} \left[e^{\theta_1 |a-b|} - e^{\theta_1(a+b)} \right] \\
&= e^{\theta_1(a+b)} \left(-\frac{\sigma_x^2}{2\theta_1} \right) - \frac{\sigma_x^2}{2\theta_1} \left[e^{\theta_1 |a-b|} - e^{\theta_1(a+b)} \right] \\
&= -\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 |a-b|}.
\end{aligned}
$$

The Lemma is proved.

Lemma 2.8 *Let X_t be the predictive variable in the market. Let $Y_t = \log P_t$ be the log transformed stock price, then $Cov\left(Y_v, X_t\right)$ for any $v \leq t$ is*

$$
Cov\left(Y_v, X_t\right) = \left(\frac{\mu_1 \sigma_x^2}{2\theta_1^2} - \frac{\sigma_x \sigma_p \rho}{\theta_1} \right) \left(e^{\theta_1(t-v)} - e^{\theta_1 t} \right).
$$

Proof.

Given that $Y_0, \mu_0, -\frac{1}{2}\sigma_p^2$ are all constants, by Lemma 2.2,

$$
\begin{aligned}
Cov\left(Y_v, X_t\right) &= Cov\left(Y_0 + \int_0^v \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2 \right) du + \sigma_p B_v, X_t \right) \\
&= \int_0^v \mu_1 Cov\left(X_u, X_t\right) du + \sigma_p Cov\left(B_v, X_t\right),
\end{aligned}
$$

then for $v \leq t$, by Lemma 2.5 and Lemma 2.7,

$$\begin{aligned}
Cov\left(Y_v, X_t\right) &= \int_0^v \mu_1 \left(-\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 (t-u)}\right) du + \sigma_p \left(-\frac{\sigma_x \rho}{\theta_1}\left(e^{\theta_1 (t-v)} - e^{\theta_1 t}\right)\right) \\
&= -\frac{\mu_1 \sigma_x^2}{2\theta_1} e^{\theta_1 t} \int_0^v e^{-\theta_1 u} du - \frac{\sigma_p \sigma_x \rho}{\theta_1} e^{\theta_1 t}\left(e^{-\theta_1 v} - 1\right) \\
&= \frac{\mu_1 \sigma_x^2 e^{\theta_1 t}}{2\theta_1^2}\left(e^{-\theta_1 v} - 1\right) - \frac{\sigma_p \sigma_x \rho}{\theta_1} e^{\theta_1 t}\left(e^{-\theta_1 v} - 1\right) \\
&= \frac{\sigma_x e^{\theta_1 t}}{\theta_1}\left(\sigma_p \rho - \frac{\mu_1 \sigma_x}{2\theta_1}\right)\left(1 - e^{-\theta_1 v}\right) \\
&= \left(\frac{\mu_1 \sigma_x^2}{2\theta_1^2} - \frac{\sigma_x \sigma_p \rho}{\theta_1}\right)\left(e^{\theta_1 (t-v)} - e^{\theta_1 t}\right).
\end{aligned}$$

The Lemma is proved.

Lemma 2.9 *Let $Y_t = \log P_t$ be the log transformed stock price. Let B_v be one-dimensional standard Brownian motion, then $Cov\left(Y_t, B_v\right)$ is independent of v for any $t \leq v$, i.e.,*

$$Cov\left(Y_t, B_v\right) = -\frac{\mu_1 \sigma_x \rho}{\theta_1}\left(t - \frac{1}{\theta_1} e^{\theta_1 t} + \frac{1}{\theta_1}\right) + \sigma_p t.$$

Proof.

For $t \leq v$, by Lemma 2.2 and Lemma 2.5 with $u \leq t \leq v$,

$$\begin{aligned}
Cov\left(Y_t, B_v\right) &= Cov\left(Y_0 + \int_0^t \left(\mu_0 + \mu_1 X_u - \frac{1}{2}\sigma_p^2\right) du + \sigma_p B_t, B_v\right) \\
&= \int_0^t \mu_1 Cov\left(X_u, B_v\right) du + \sigma_p Cov\left(B_t, B_v\right) \\
&= \int_0^t \mu_1 \left(-\frac{\sigma_x \rho}{\theta_1}\left(1 - e^{\theta_1 u}\right)\right) du + \sigma_p \min\left(t, v\right) \\
&= -\frac{\mu_1 \sigma_x \rho}{\theta_1}\left(t - \frac{1}{\theta_1} e^{\theta_1 t} + \frac{1}{\theta_1}\right) + \sigma_p t.
\end{aligned}$$

The Lemma is proved.

Lemma 2.10 *Let $Y_t = \log P_t$ be the log transformed stock price, then $Cov\left(Y_u, Y_v\right)$ for any u, v is*

$$Cov\left(Y_u, Y_v\right) = \left(\sigma_p^2 + 2\theta_1 c_1\right) \min\left(u, v\right) + c_1 \left(1 - e^{\theta_1 v} + e^{\theta_1 |u-v|} - e^{\theta_1 u}\right),$$

where c_1 is a constant and independent of u, v, i.e.,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2}.$$

Proof.

Given that $Y_0, \mu_0, -\frac{1}{2}\sigma_p^2$ are all constants, by Lemma 2.2, for $u \geq v$,

$$\begin{aligned}
&Cov\left(Y_u, Y_v\right) \\
&= Cov\Big(Y_0 + \int_0^u \left(\mu_0 + \mu_1 X_a - \frac{1}{2}\sigma_p^2\right) da + \sigma_p B_u, \\
&Y_0 + \int_0^v \left(\mu_0 + \mu_1 X_b - \frac{1}{2}\sigma_p^2\right) db + \sigma_p B_v\Big) \\
&= Cov\left(\mu_1 \int_0^u X_a da + \sigma_p B_u, \mu_1 \int_0^v X_b db + \sigma_p B_v\right) \\
&= \mu_1^2 \int_0^u da \int_0^v db Cov\left(X_a, X_b\right) + \mu_1 \sigma_p \int_0^u Cov\left(X_a, B_v\right) da \\
&+ \mu_1 \sigma_p \int_0^v Cov\left(X_b, B_u\right) db + \sigma_p^2 Cov\left(B_u, B_v\right),
\end{aligned}$$

where by Lemma 2.5,

$$\begin{aligned}
&\int_0^u Cov\left(X_a, B_v\right) da \\
&= \int_0^v Cov\left(X_a, B_v\right) da + \int_v^u Cov(X_a, B_v) da \\
&= \int_0^v -\frac{\sigma_x \rho}{\theta_1}\left(1 - e^{\theta_1 a}\right) da + \int_v^u -\frac{\sigma_x \rho}{\theta_1}\left(e^{\theta_1 (a-v)} - e^{\theta_1 a}\right) da \\
&= \frac{\sigma_x \rho}{\theta_1}\left[\int_0^v e^{\theta_1 a} da - \int_0^v da + \int_v^u e^{\theta_1 a} da - \int_v^u e^{\theta_1 (a-v)} da\right] \\
&= \frac{\sigma_x \rho}{\theta_1}\left[\int_0^u e^{\theta_1 a} da - v - \int_v^u e^{\theta_1 (a-v)} da\right] \\
&= \frac{\sigma_x \rho}{\theta_1}\left[\frac{1}{\theta_1}\left(e^{\theta_1 u} - 1\right) - v - \frac{1}{\theta_1} e^{-\theta_1 v}\left(e^{\theta_1 u} - e^{\theta_1 v}\right)\right] \\
&= \frac{\sigma_x \rho}{\theta_1^2}\left[e^{\theta_1 u} - 1 - \theta_1 v - e^{\theta_1 (u-v)} + 1\right] \\
&= \frac{\sigma_x \rho}{\theta_1^2}\left[e^{\theta_1 u} - e^{\theta_1 (u-v)} - \theta_1 v\right],
\end{aligned}$$

and also by Lemma 2.5,

$$\begin{aligned}
&\int_0^v Cov\left(X_b, B_u\right) db \\
&= \int_0^v -\frac{\sigma_x \rho}{\theta_1}\left(1 - e^{\theta_1 b}\right) db \\
&= \frac{\sigma_x \rho}{\theta_1}\left[\int_0^v e^{\theta_1 b} db - \int_0^v db\right] \\
&= \frac{\sigma_x \rho}{\theta_1}\left[\frac{1}{\theta_1}\left(e^{\theta_1 v} - 1\right) - v\right] \\
&= \frac{\sigma_x \rho}{\theta_1^2}\left[e^{\theta_1 v} - 1 - \theta_1 v\right],
\end{aligned}$$

and $Cov(B_u, B_v) = \min(u, v) = v$. In addition, by Lemma 2.7 and $0 \le b \le u$,

$$\begin{aligned}
&\int_0^u da \int_0^v db Cov(X_a, X_b) \\
&= \int_0^u da \int_0^v db \left(-\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 |a-b|}\right) \\
&= -\frac{\sigma_x^2}{2\theta_1} \int_0^v db \int_0^u e^{\theta_1|a-b|} da = -\frac{\sigma_x^2}{2\theta_1} \int_0^v db \left[\int_0^b e^{\theta_1|a-b|} da + \int_b^u e^{\theta_1|a-b|} da\right] \\
&= -\frac{\sigma_x^2}{2\theta_1} \int_0^v db \left[\int_0^b e^{\theta_1(b-a)} da + \int_b^u e^{\theta_1(a-b)} da\right] \\
&= -\frac{\sigma_x^2}{2\theta_1} \int_0^v \left[\frac{1}{\theta_1}\left(e^{\theta_1 b} - 1\right) + \frac{1}{\theta_1}\left(e^{\theta_1(u-b)} - 1\right)\right] db \\
&= -\frac{\sigma_x^2}{2\theta_1^2} \int_0^v \left[e^{\theta_1 b} + e^{\theta_1(u-b)} - 2\right] db \\
&= -\frac{\sigma_x^2}{2\theta_1^2} \left[\frac{1}{\theta_1}\left(e^{\theta_1 v} - 1\right) - \frac{1}{\theta_1}\left(e^{\theta_1(u-v)} - e^{\theta_1 u}\right) - 2v\right] \\
&= -\frac{\sigma_x^2}{2\theta_1^3} \left[e^{\theta_1 v} - e^{\theta_1(u-v)} + e^{\theta_1 u} - 2v\theta_1 - 1\right],
\end{aligned}$$

which implies, for $u \ge v$,

$$\begin{aligned}
&Cov(Y_u, Y_v) \\
&= \mu_1^2 \left(-\frac{\sigma_x^2}{2\theta_1^3}\left[e^{\theta_1 v} - e^{\theta_1(u-v)} + e^{\theta_1 u} - 2v\theta_1 - 1\right]\right) \\
&+ \mu_1\sigma_p \left(\frac{\sigma_x\rho}{\theta_1^2}\left[e^{\theta_1 u} - e^{\theta_1(u-v)} - \theta_1 v\right]\right) + \mu_1\sigma_p \left(\frac{\sigma_x\rho}{\theta_1^2}\left[e^{\theta_1 v} - 1 - \theta_1 v\right]\right) + \sigma_p^2 v \\
&= \left(\frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2} - \frac{\mu_1^2\sigma_x^2}{2\theta_1^3}\right)\left(e^{\theta_1 u} + e^{\theta_1 v} - e^{\theta_1(u-v)} - 2v\theta_1 - 1\right) + \sigma_p^2 v \\
&= \left(\sigma_p^2 + \frac{\mu_1^2\sigma_x^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_p\rho}{\theta_1}\right) v + \left(\frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2}\right)\left(1 - e^{\theta_1 v} + e^{\theta_1(u-v)} - e^{\theta_1 u}\right) \\
&= \left(\sigma_p^2 + 2\theta_1 c_1\right) v + c_1\left(1 - e^{\theta_1 v} + e^{\theta_1(u-v)} - e^{\theta_1 u}\right),
\end{aligned}$$

where c_1 is a constant and independent of u, v, i.e.,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2}.$$

Similarly, for $u < v$,

$$Cov\left(Y_u, Y_v\right) = \left(\sigma_p^2 + 2\theta_1 c_1\right) u + c_1 \left(1 - e^{\theta_1 u} + e^{\theta_1 (v-u)} - e^{\theta_1 v}\right).$$

Therefore, $\forall u, v$,

$$Cov\left(Y_u, Y_v\right) = \left(\sigma_p^2 + 2\theta_1 c_1\right) \min\left(u, v\right) + c_1 \left(1 - e^{\theta_1 v} + e^{\theta_1 |u-v|} - e^{\theta_1 u}\right),$$

where c_1 is a constant and independent of u, v,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2}.$$

The Lemma is proved.

Lemma 2.11 *Let X_t be the predictive variable in the market. Let $M_t^{(h)}$ be the moving average based on lookback period $h > 0$, then $Cov\left(X_t, M_t^{(h)}\right)$ is*

$$Cov\left(X_t, M_t^{(h)}\right) = -\frac{c_1}{h\mu_1}\left(1 - e^{\theta_1 h}\right) - \frac{c_1 \theta_1}{\mu_1} e^{\theta_1 t},$$

where c_1 is a constant,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2}.$$

Proof.

By Lemma 2.8 and $u \le t$,

$$
\begin{aligned}
& Cov\left(X_t, M_t^{(h)}\right) \\
&= Cov\left(X_t, \frac{1}{h}\int_{t-h}^{t} Y_u du\right) = \frac{1}{h}\int_{t-h}^{t} Cov\left(Y_u, X_t\right) du \\
&= \frac{1}{h}\int_{t-h}^{t}\left(\frac{\mu_1\sigma_x^2}{2\theta_1^2} - \frac{\sigma_x\sigma_p\rho}{\theta_1}\right)\left(e^{\theta_1(t-u)} - e^{\theta_1 t}\right) du = \frac{1}{h}\frac{c_1\theta_1}{\mu_1}\int_{t-h}^{t}\left(e^{\theta_1(t-u)} - e^{\theta_1 t}\right) du \\
&= \frac{1}{h}\frac{c_1\theta_1}{\mu_1}\left[\int_{t-h}^{t} e^{\theta_1(t-u)} du - \int_{t-h}^{t} e^{\theta_1 t} du\right] = \frac{1}{h}\frac{c_1\theta_1}{\mu_1}\left[-\frac{1}{\theta_1}\left(1 - e^{\theta_1 h}\right) - e^{\theta_1 t} h\right] \\
&= -\frac{c_1}{h\mu_1}\left(1 - e^{\theta_1 h}\right) - \frac{c_1\theta_1}{\mu_1} e^{\theta_1 t},
\end{aligned}
$$

where c_1 is a constant,

$$c_1 = \frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2}.$$

The Lemma is proved.

Lemma 2.12 *Let X_t be the predictive variable in the market. Let $M_t^{(s,l)}$ be the difference between moving average $M_t^{(s)}$ and moving average $M_t^{(l)}$, where $s < l$ is lookback period, then $Cov\left(X_t, M_t^{(s,l)}\right)$ is independent of time t, i.e.,*

$$Cov\left(X_t, M_t^{(s,l)}\right) = \frac{c_1}{l\mu_1}\left(1 - e^{\theta_1 l}\right) - \frac{c_1}{s\mu_1}\left(1 - e^{\theta_1 s}\right),$$

where c_1 is a constant,

$$c_1 = \frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2}.$$

Proof.

By Lemma 2.11,

$$Cov\left(X_t, M_t^{(s,l)}\right) = Cov\left(X_t, M_t^{(s)} - M_t^{(l)}\right)$$
$$= Cov\left(X_t, M_t^{(s)}\right) - Cov\left(X_t, M_t^{(l)}\right)$$
$$= \left(-\frac{c_1}{s\mu_1}\left(1 - e^{\theta_1 s}\right) - \frac{c_1\theta_1}{\mu_1}e^{\theta_1 t}\right) - \left(-\frac{c_1}{l\mu_1}\left(1 - e^{\theta_1 l}\right) - \frac{c_1\theta_1}{\mu_1}e^{\theta_1 t}\right)$$
$$= \frac{c_1}{l\mu_1}\left(1 - e^{\theta_1 l}\right) - \frac{c_1}{s\mu_1}\left(1 - e^{\theta_1 s}\right),$$

which shows it is independent of time t, and

$$c_1 = \frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2}.$$

The Lemma is proved.

Lemma 2.13 *Let $M_t^{(s,l)}$ be the difference between moving average $M_t^{(s)}$ and moving average $M_t^{(l)}$, where $s < l$ is lookback period, then $M_t^{(s,l)}$ is normally distributed with mean $E\left(M_t^{(s,l)}\right)$, which is independent of time t, i.e.,*

$$E\left(M_t^{(s,l)}\right) = \frac{1}{2}(l-s)\left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right).$$

Proof.

Since $M_t^{(s,l)}$ is a linear combination of $M_t^{(s)}$ and $M_t^{(l)}$, by Lemma 2.3,

$$E\left(M_t^{(s,l)}\right) = E\left(M_t^{(s)} - M_t^{(l)}\right) = E\left(M_t^{(s)}\right) - E\left(M_t^{(l)}\right)$$
$$= \left[Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right)\left(t - \frac{s}{2}\right)\right] - \left[Y_0 + \left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right)\left(t - \frac{l}{2}\right)\right]$$
$$= \frac{1}{2}(l-s)\left(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2\right),$$

and $M_t^{(s,l)}$ is also normally distributed. The Lemma is proved.

Lemma 2.14 *Let $M_t^{(s)}$ and $M_t^{(l)}$ be the moving averages based on lookback period s and l, where $s < l$, then $Cov\left(M_t^{(s)}, M_t^{(l)}\right)$ is*

$$Cov\left(M_t^{(s)}, M_t^{(l)}\right) = \left[\left(\sigma_p^2 + 2\theta_1 c_1\right)t - \frac{c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1 (t-s)}\right) - \frac{c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1 (t-l)}\right)\right]$$
$$+ c_2\left(s, l\right),$$

where c_1 is a constant,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2},$$

and $c_2\left(s, l\right)$ is a constant, which depends on s and l, but does not depend on t, i.e.,

$$c_2\left(s, l\right) = c_1 - \frac{c_1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1 (l-s)} + 2\theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(\frac{s^2}{6l} + \frac{l}{2}\right).$$

The special cases,

$$Cov\left(M_t^{(s)}, M_t^{(s)}\right) = \left[\left(\sigma_p^2 + 2\theta_1 c_1\right)t - \frac{2c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1 (t-s)}\right)\right] + c_2\left(s, s\right),$$
$$Cov\left(M_t^{(l)}, M_t^{(l)}\right) = \left[\left(\sigma_p^2 + 2\theta_1 c_1\right)t - \frac{2c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1 (t-l)}\right)\right] + c_2\left(l, l\right),$$

where

$$c_2\left(s, s\right) = c_1 - \frac{2c_1}{s^2\theta_1^2}\left(1 - e^{\theta_1 s} + \theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2s}{3},$$
$$c_2\left(l, l\right) = c_1 - \frac{2c_1}{l^2\theta_1^2}\left(1 - e^{\theta_1 l} + \theta_1 l\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2l}{3}.$$

Proof.

Since $s \leq l$, then $t-s \geq t-l$, then by Lemma 2.10,

$$Cov\left(M_t^{(s)}, M_t^{(l)}\right) = Cov\left(\frac{1}{s}\int_{t-s}^{t} Y_u du, \frac{1}{l}\int_{t-l}^{t} Y_v dv\right) = \frac{1}{sl}\int_{t-s}^{t} du \int_{t-l}^{t} Cov\left(Y_u, Y_v\right) dv$$

$$= \frac{1}{sl}\int_{t-s}^{t} du \left[\int_{t-l}^{u} Cov\left(Y_u, Y_v\right) dv + \int_{u}^{t} Cov\left(Y_u, Y_v\right) dv\right]$$

$$= \frac{1}{sl}\int_{t-s}^{t} du \Bigg[\int_{t-l}^{u} \left(\left(\sigma_p^2 + 2\theta_1 c_1\right) v + c_1\left(1 - e^{\theta_1 v} - e^{\theta_1 u} + e^{\theta_1(u-v)}\right)\right) dv$$

$$+ \int_{u}^{t} \left(\left(\sigma_p^2 + 2\theta_1 c_1\right) u + c_1\left(1 - e^{\theta_1 v} - e^{\theta_1 u} + e^{\theta_1(v-u)}\right)\right) dv\Bigg]$$

$$= \frac{1}{sl}\int_{t-s}^{t} \Bigg[\frac{\sigma_p^2 + 2\theta_1 c_1}{2}\left(-\left(u-t\right)^2 + 2tl - l^2\right) + c_1\left(1 - e^{\theta_1 u}\right) l - \frac{c_1}{\theta_1}\left(e^{\theta_1 t} - e^{\theta_1(t-l)}\right)$$

$$+ \frac{c_1}{\theta_1}\left(e^{\theta_1(t-u)} + e^{\theta_1(u-t+l)} - 2\right)\Bigg] du$$

$$= \frac{1}{sl}\Bigg[\frac{\sigma_p^2 + 2\theta_1 c_1}{2}\left(-\frac{s^3}{3} + 2tls - l^2 s\right) + c_1\Bigg(ls - \frac{l}{\theta_1}\left(e^{\theta_1 t} - e^{\theta_1(t-s)}\right)$$

$$- \frac{s}{\theta_1}\left(e^{\theta_1 t} - e^{\theta_1(t-l)}\right)\Bigg) + \frac{c_1}{\theta_1}\left(\frac{1}{\theta_1}\left(e^{\theta_1 s} - 1\right) + \frac{1}{\theta_1}\left(e^{\theta_1 l} - e^{\theta_1(l-s)}\right) - 2s\right)\Bigg]$$

$$= \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(t - \frac{s^2}{6l} - \frac{l}{2}\right) + c_1\Bigg(1 - \frac{1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1(l-s)} + 2\theta_1 s\right)$$

$$- \frac{1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1(t-s)}\right) - \frac{1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1(t-l)}\right)\Bigg)$$

$$= \left[\left(\sigma_p^2 + 2\theta_1 c_1\right) t - \frac{c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1(t-s)}\right) - \frac{c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1(t-l)}\right)\right]$$

$$+ \left[c_1 - \frac{c_1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1(l-s)} + 2\theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(\frac{s^2}{6l} + \frac{l}{2}\right)\right]$$

$$= \left[\left(\sigma_p^2 + 2\theta_1 c_1\right) t - \frac{c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1(t-s)}\right) - \frac{c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1(t-l)}\right)\right] + c_2\left(s, l\right),$$

where c_1 is a constant,

$$c_1 = \frac{\mu_1^2 \sigma_x^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \sigma_p \rho}{\theta_1^2},$$

and $c_2(s,l)$ is a constant, which depends on s and l, but does not depend on t, i.e.,

$$c_2(s,l) = c_1 - \frac{c_1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1 (l-s)} + 2\theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(\frac{s^2}{6l} + \frac{l}{2}\right).$$

The Lemma is proved.

Lemma 2.15 *Let $M_t^{(s,l)}$ be the difference between moving average $M_t^{(s)}$ and moving average $M_t^{(l)}$, where $s < l$ is lookback period, then $Var\left(M_t^{(s,l)}\right)$ is independent of time t, i.e.,*

$$Var\left(M_t^{(s,l)}\right) = c_2(s,s) + c_2(l,l) - 2c_2(s,l),$$

where $c_2(s,s)$, $c_2(l,l)$ and $c_2(s,l)$ are constants, which depend on s and l, i.e.,

$$c_2(s,l) = c_1 - \frac{c_1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1 (l-s)} + 2\theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(\frac{s^2}{6l} + \frac{l}{2}\right),$$

and

$$c_2(s,s) = c_1 - \frac{2c_1}{s^2\theta_1^2}\left(1 - e^{\theta_1 s} + \theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2s}{3},$$
$$c_2(l,l) = c_1 - \frac{2c_1}{l^2\theta_1^2}\left(1 - e^{\theta_1 l} + \theta_1 l\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2l}{3}.$$

Proof.

By Lemma 2.14,

$$Var\left(M_t^{(s,l)}\right) = Var\left(M_t^{(s)} - M_t^{(l)}\right) = Cov\left(M_t^{(s)} - M_t^{(l)}, M_t^{(s)} - M_t^{(l)}\right)$$

$$= Cov\left(M_t^{(s)}, M_t^{(s)}\right) + Cov\left(M_t^{(l)}, M_t^{(l)}\right) - 2Cov\left(M_t^{(s)}, M_t^{(l)}\right)$$

$$= \left[\left(\sigma_p^2 + 2\theta_1 c_1\right) t - \frac{2c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1 (t-s)}\right)\right] + c_2\left(s, s\right)$$

$$+ \left[\left(\sigma_p^2 + 2\theta_1 c_1\right) t - \frac{2c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1 (t-l)}\right)\right] + c_2\left(l, l\right)$$

$$- \left[2\left(\sigma_p^2 + 2\theta_1 c_1\right) t - \frac{2c_1}{\theta_1 s}\left(e^{\theta_1 t} - e^{\theta_1 (t-s)}\right) - \frac{2c_1}{\theta_1 l}\left(e^{\theta_1 t} - e^{\theta_1 (t-l)}\right)\right] - 2c_2\left(s, l\right)$$

$$= c_2\left(s, s\right) + c_2\left(l, l\right) - 2c_2\left(s, l\right).$$

The Lemma is proved.

2.3 The Analytic Results

In order to find optimal η_t, we need derive the investor's expected log-utility of wealth $E\left(\log W_T\right)$. To derive $E\left(\log W_T\right)$, we need know joint distribution of $\left(X_t, M_t^{(s,l)}\right)^T$. Based on Lemmas 2.1, 2.12, 2.13 and 2.15, it is derived that $\left(X_t, M_t^{(s,l)}\right)^T$ are jointly normally distributed, i.e.,

$$\begin{pmatrix} X_t \\ M_t^{(s,l)} \end{pmatrix} \sim \mathrm{MN}\left[\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}\right], \tag{2.11}$$

and

$$
\begin{aligned}
b_1 &= -\frac{\theta_0}{\theta_1}, \\
b_2 &= \frac{1}{2}(l-s)(\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \frac{1}{2}\sigma_p^2), \\
\sigma_1^2 &= -\frac{\sigma_x^2}{2\theta_1}, \\
\sigma_{12} = \sigma_{21} &= \frac{c_1}{l\mu_1}(1-e^{\theta_1 l}) - \frac{c_1}{s\mu_1}(1-e^{\theta_1 s}), \\
\sigma_2^2 &= c_2\left(s,s\right) + c_2\left(l,l\right) - 2c_2\left(s,l\right), \qquad (2.12)
\end{aligned}
$$

where c_1 is a constant, i.e.,

$$
c_1 = \frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_p\rho}{\theta_1^2},
$$

and $c_2\left(s,l\right)$, $c_2\left(s,s\right)$ and $c_2\left(l,l\right)$ are constants, which depend on s and l, but do not depend on t, i.e.,

$$
\begin{aligned}
c_2\left(s,l\right) &= c_1 - \frac{c_1}{sl\theta_1^2}\left(1 - e^{\theta_1 s} - e^{\theta_1 l} + e^{\theta_1(l-s)} + 2\theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\left(\frac{s^2}{6l} + \frac{l}{2}\right), \\
c_2\left(s,s\right) &= c_1 - \frac{2c_1}{s^2\theta_1^2}\left(1 - e^{\theta_1 s} + \theta_1 s\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2s}{3}, \\
c_2\left(l,l\right) &= c_1 - \frac{2c_1}{l^2\theta_1^2}\left(1 - e^{\theta_1 l} + \theta_1 l\right) - \left(\sigma_p^2 + 2\theta_1 c_1\right)\frac{2l}{3}.
\end{aligned}
$$

Notice that both marginal and joint density functions are independent of time t. The marginal density functions of X_t and $M_t^{(s,l)}$ can be defined as

$$
\begin{aligned}
f_{X_t}(\epsilon) &= \frac{1}{\sigma_1}\phi(\frac{\epsilon - b_1}{\sigma_1}), \\
f_{M_t^{(s,l)}}(\epsilon) &= \frac{1}{\sigma_2}\phi(\frac{\epsilon - b_2}{\sigma_2}),
\end{aligned}
\qquad (2.13)
$$

where ϕ is standard normal density function, i.e.,

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2},$$

and cumulative distribution functions of X_t and $M_t^{(s,l)}$ can be defined as

$$\begin{aligned} F_{X_t}(\epsilon) &= P(X_t \leq \epsilon) = \Phi(\frac{\epsilon - b_1}{\sigma_1}), \\ F_{M_t^{(s,l)}}(\epsilon) &= P(M_t^{(s,l)} \leq \epsilon) = \Phi(\frac{\epsilon - b_2}{\sigma_2}), \end{aligned} \tag{2.14}$$

where Φ is standard normal cumulative distribution function, i.e.,

$$\Phi(u) = \int_{-\infty}^{u} \phi(v) dv = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}v^2} dv.$$

We state following Propositions for the SGMA strategy.

Proposition 2.1 *The expected value of η_t is independent of time t, i.e.,*

$$E(\eta_t) = 1 - \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \delta_1 \left[\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right] + \delta_2 \left[\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right].$$

Proof.

By the definition of η_t in Equation (2.8) and by Equation (2.14),

$$\begin{aligned} &E(\eta_t) \\ &= 1 \times P\left(M_t^{(s,l)} > \epsilon\right) + \delta_1 P\left(0 \leq M_t^{(s,l)} \leq \epsilon\right) + \delta_2 P\left(-\epsilon \leq M_t^{(s,l)} < 0\right) \\ &= \left(1 - F_{M_t^{(s,l)}}(\epsilon)\right) + \delta_1 \left(F_{M_t^{(s,l)}}(\epsilon) - F_{M_t^{(s,l)}}(0)\right) + \delta_2 \left(F_{M_t^{(s,l)}}(0) - F_{M_t^{(s,l)}}(-\epsilon)\right) \\ &= 1 - \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \delta_1 \left[\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right] + \delta_2 \left[\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right]. \end{aligned}$$

The Proposition is proved.

Proposition 2.2 *The expected value of η_t^2 is independent of time t, i.e.,*

$$E\left(\eta_t^2\right) = 1 - \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \delta_1^2\left[\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right] + \delta_2^2\left[\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right].$$

Proof.

By the definition of η_t in Equation (2.8) and by Equation (2.14),

$$\begin{aligned}
&E\left(\eta_t^2\right)\\
&= 1 \times P\left(M_t^{(s,l)} > \epsilon\right) + \delta_1^2 P\left(0 \le M_t^{(s,l)} \le \epsilon\right) + \delta_2^2 P\left(-\epsilon \le M_t^{(s,l)} < 0\right)\\
&= \left(1 - F_{M_t^{(s,l)}}(\epsilon)\right) + \delta_1^2\left(F_{M_t^{(s,l)}}(\epsilon) - F_{M_t^{(s,l)}}(0)\right) + \delta_2^2\left(F_{M_t^{(s,l)}}(0) - F_{M_t^{(s,l)}}(-\epsilon)\right)\\
&= 1 - \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \delta_1^2\left[\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right] + \delta_2^2\left[\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right].
\end{aligned}$$

The Proposition is proved.

Proposition 2.3 *Let $\Psi(u) = \int_{-\infty}^{u} v\phi(v)\,dv$, then the expected value of $\eta_t M_t^{(s,l)}$ is independent of time t, i.e.,*

$$\begin{aligned}
E\left(\eta_t M_t^{(s,l)}\right) = 1 &- \left[b_2\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \sigma_2\Psi\left(\frac{\epsilon - b_2}{\sigma_2}\right)\right]\\
&+ \delta_1\left[b_2\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) + \sigma_2\Psi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - b_2\Phi\left(\frac{-b_2}{\sigma_2}\right) - \sigma_2\Psi\left(\frac{-b_2}{\sigma_2}\right)\right]\\
&+ \delta_2\left[b_2\Phi\left(\frac{-b_2}{\sigma_2}\right) + \sigma_2\Psi\left(\frac{-b_2}{\sigma_2}\right) - b_2\Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right) - \sigma_2\Psi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right].
\end{aligned}$$

Proof.

Let $\Psi(u) = \int_{-\infty}^{u} v\phi(v)\,dv$ and let $t = \frac{m - b_2}{\sigma_2}$, then $m = b_2 + \sigma_2 t$ and $dm = \sigma_2 dt$.

We define $H_{M_t^{(s,l)}}(\epsilon)$ as

$$\begin{aligned}
H_{M_t^{(s,l)}}(\epsilon) &= \int_{m<\epsilon} m f_{M_t^{(s,l)}}(m)\, dm = \int_{-\infty}^{\epsilon} m \frac{1}{\sigma_2} \phi\left(\frac{m-b_2}{\sigma_2}\right) dm \\
&= \int_{-\infty}^{\frac{\epsilon-b_2}{\sigma_2}} (b_2+\sigma_2 t)\, \phi(t)\, dt = b_2 \Phi\left(\frac{\epsilon-b_2}{\sigma_2}\right) + \sigma_2 \Psi\left(\frac{\epsilon-b_2}{\sigma_2}\right),
\end{aligned}$$

then

$$\begin{aligned}
& E\left(\eta_t M_t^{(s,l)}\right) \\
&= 1 - H_{M_t^{(s,l)}}(\epsilon) + \delta_1 \left(H_{M_t^{(s,l)}}(\epsilon) - H_{M_t^{(s,l)}}(0)\right) + \delta_2 \left(H_{M_t^{(s,l)}}(0) - H_{M_t^{(s,l)}}(-\epsilon)\right) \\
&= 1 - \left[b_2 \Phi\left(\frac{\epsilon-b_2}{\sigma_2}\right) + \sigma_2 \Psi\left(\frac{\epsilon-b_2}{\sigma_2}\right)\right] \\
&+ \delta_1 \left[b_2 \Phi\left(\frac{\epsilon-b_2}{\sigma_2}\right) + \sigma_2 \Psi\left(\frac{\epsilon-b_2}{\sigma_2}\right) - b_2 \Phi\left(\frac{-b_2}{\sigma_2}\right) - \sigma_2 \Psi\left(\frac{-b_2}{\sigma_2}\right)\right] \\
&+ \delta_2 \left[b_2 \Phi\left(\frac{-b_2}{\sigma_2}\right) + \sigma_2 \Psi\left(\frac{-b_2}{\sigma_2}\right) - b_2 \Phi\left(\frac{-\epsilon-b_2}{\sigma_2}\right) - \sigma_2 \Psi\left(\frac{-\epsilon-b_2}{\sigma_2}\right)\right].
\end{aligned}$$

The Proposition is proved.

Proposition 2.4 *The expected value of $\eta_t X_t$ is independent of time t, i.e.,*

$$E(\eta_t X_t) = \left(b_1 - \frac{\sigma_{12} b_2}{\sigma_2^2}\right) E(\eta_t) + \frac{\sigma_{12}}{\sigma_2^2} E\left(\eta_t M_t^{(s,l)}\right),$$

where $E(\eta_t)$ satisfies Proposition 2.1 and $E\left(\eta_t M_t^{(s,l)}\right)$ satisfies Proposition 2.3.

Proof.

Based on the joint distribution for $\left(X_t, M_t^{(s,l)}\right)^T$, it is derived that

$$E\left(X_t \mid M_t^{(s,l)}\right) = b_1 + \rho^* \frac{\sigma_1}{\sigma_2}(m - b_2) = b_1 + \frac{\sigma_{12}}{\sigma_2^2}(m - b_2),$$

as $\rho^* = \frac{Cov\left(X_t, M_t^{(s,l)}\right)}{SD(X_t)SD\left(M_t^{(s,l)}\right)} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$. By Law of total expectation $E(X) = E(E(X \mid Y))$,

$$E(\eta_t X_t) = E\left(\eta_t E\left(X_t \mid M_t^{(s,l)}\right)\right),$$

which implies

$$E(\eta_t X_t) = E\left(\eta_t\left(b_1 + \frac{\sigma_{12}}{\sigma_2^2}(m - b_2)\right)\right) = \left(b_1 - \frac{\sigma_{12} b_2}{\sigma_2^2}\right) E(\eta_t) + \frac{\sigma_{12}}{\sigma_2^2} E\left(\eta_t M_t^{(s,l)}\right).$$

The Proposition is proved.

Proposition 2.5 *Let λ be fixed constant for the SGMA strategy η_t when $t < l$. Let $\epsilon > 0$ be the investor specified risk tolerance, then the investor's expected log-utility of wealth at the end of investment period T is*

$$E(\log W_T) = c_3 + (T - l)\left[(\mu_0 - r) E(\eta_t) - \frac{\sigma_p^2}{2} E\left(\eta_t^2\right) + \mu_1 E(\eta_t X_t)\right], \quad (2.15)$$

where c_3 is a constant depending on l, i.e.,

$$c_3 = \log W_0 + rT + l\left[\lambda(\mu_0 - r) - \frac{1}{2}\lambda^2\sigma_p^2 - \frac{\lambda\mu_1\theta_0}{\theta_1}\right].$$

By Propositions 2.1, 2.2 *and* 2.4, *Equation* 2.15 *can be rewritten as*

$$\begin{aligned} E(\log W_T) = {} & -\alpha_1^{(\epsilon)}\left(\delta_1 - \frac{\beta_1^{(\epsilon)}}{2\alpha_1^{(\epsilon)}}\right)^2 - \alpha_2^{(\epsilon)}\left(\delta_2 - \frac{\beta_2^{(\epsilon)}}{2\alpha_2^{(\epsilon)}}\right)^2 \\ & + \left(\gamma^{(\epsilon)} + \frac{\left(\beta_1^{(\epsilon)}\right)^2}{4\alpha_1^{(\epsilon)}} + \frac{\left(\beta_2^{(\epsilon)}\right)^2}{4\alpha_2^{(\epsilon)}}\right), \end{aligned} \quad (2.16)$$

where $\alpha_1^{(\epsilon)} > 0$, $\beta_1^{(\epsilon)}$, $\alpha_2^{(\epsilon)} > 0$, $\beta_2^{(\epsilon)}$ *and* $\gamma^{(\epsilon)}$ *are constants depending on* ϵ, *i.e.,*

$$\alpha_1^{(\epsilon)} = \frac{\sigma_p^2 (T-l)}{2}\left(\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right),$$

$$\beta_1^{(\epsilon)} = (T-l)\Bigg[(\mu_0 - r + \mu_1 b_1)\left(\Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Phi\left(\frac{-b_2}{\sigma_2}\right)\right)$$

$$+ \frac{\mu_1\sigma_{12}}{\sigma_2}\left(\Psi\left(\frac{\epsilon - b_2}{\sigma_2}\right) - \Psi\left(\frac{-b_2}{\sigma_2}\right)\right)\Bigg],$$

$$\alpha_2^{(\epsilon)} = \frac{\sigma_p^2 (T-l)}{2}\left(\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right),$$

$$\beta_2^{(\epsilon)} = (T-l)\Bigg[(\mu_0 - r + \mu_1 b_1)\left(\Phi\left(\frac{-b_2}{\sigma_2}\right) - \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right)$$

$$+ \frac{\mu_1\sigma_{12}}{\sigma_2}\left(\Psi\left(\frac{-b_2}{\sigma_2}\right) - \Psi\left(\frac{-\epsilon - b_2}{\sigma_2}\right)\right)\Bigg],$$

$$\gamma^{(\epsilon)} = \log W_0 + rT + \left[\lambda(\mu_0 - r) - \frac{1}{2}\lambda^2\sigma_p^2 - \frac{\lambda\mu_1\theta_0}{\theta_1}\right] l + (T-l)$$

$$\left[\left(\mu_0 - r + \mu_1 b_1 - \frac{\sigma_p^2}{2}\right)\left(1 - \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right)\right) + \frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(1 - b_2 - \sigma_2\Psi\left(\frac{\epsilon - b_2}{\sigma_2}\right)\right)\right],$$

where Φ *is standard normal cumulative distribution function, i.e.,* $\Phi(u) = \int_{-\infty}^{u}\phi(v)\,dv$ *and* $\Psi(u) = \int_{-\infty}^{u} v\phi(v)\,dv$.

Proof.

Based on Equations (2.1) and (2.2), the budget constraint follows

$$\frac{dW_t}{W_t} = \eta_t\frac{dP_t}{P_t} + (1-\eta_t)\frac{dP_t^f}{P_t^f} = \eta_t\left[(\mu_0 + \mu_1 X_t)\,dt + \sigma_p dB_t\right] + (1-\eta_t)\,rdt$$

$$= \left[r + \eta_t(\mu_0 + \mu_1 X_t - r)\right]dt + \eta_t\sigma_p dB_t,$$

Equation (2.10) is proved. Since $(du)^2 = o(du), dudB_u = o(du)$ and $(dB_u)^2 = du$,

$$\left(\frac{dW_t}{W_t}\right)^2 = [r + \eta_t (\mu_0 + \mu_1 X_t - r)]^2 (dt)^2 + 2\eta_t \sigma_p dB_t [r + \eta_t (\mu_0 + \mu_1 X_t - r)] dt$$

$$+ (\eta_t \sigma_p dB_t)^2 = \eta_t^2 \sigma_p^2 dt,$$

which implies

$$d(\log W_t) = [\log W_t]' dW_t + \frac{1}{2} [\log W_t]'' (dW_t)^2 = \frac{dW_t}{W_t} + \frac{1}{2}\left(-\frac{1}{W_t^2}\right)(dW_t)^2$$

$$= \frac{dW_t}{W_t} - \frac{1}{2}\left(\frac{dW_t}{W_t}\right)^2 = [r + \eta_t (\mu_0 + \mu_1 X_t - r)] dt + \eta_t \sigma_p dB_t - \frac{1}{2}\eta_t^2 \sigma_p^2 dt$$

$$= \left[r + \eta_t (\mu_0 + \mu_1 X_t - r) - \frac{1}{2}\eta_t^2 \sigma_p^2\right] dt + \eta_t \sigma_p dB_t.$$

Since $\log W_t \,|_0^T = \log W_T - \log W_0 = \int_0^T d(\log W_t)$ and by Equation (2.8) with $T \geq l$,

$$\log W_T = \log W_0 + \int_0^T \left[\left(r + \eta_t (\mu_0 + \mu_1 X_t - r) - \frac{1}{2}\eta_t^2 \sigma_p^2\right) dt + \eta_t \sigma_p dB_t\right]$$

$$= \log W_0 + \int_0^T r dt + \int_0^T \eta_t (\mu_0 - r) dt + \int_0^T \eta_t \mu_1 X_t dt - \frac{1}{2}\int_0^T \eta_t^2 \sigma_p^2 dt + \int_0^T \eta_t \sigma_p dB_t$$

$$= \log W_0 + rT + \int_0^l \eta_t (\mu_0 - r) dt + \int_l^T \eta_t (\mu_0 - r) dt + \int_0^l \eta_t \mu_1 X_t dt + \int_l^T \eta_t \mu_1 X_t dt$$

$$- \frac{1}{2}\int_0^l \eta_t^2 \sigma_p^2 dt - \frac{1}{2}\int_l^T \eta_t^2 \sigma_p^2 dt + \int_0^T \eta_t \sigma_p dB_t$$

$$= \log W_0 + rT + \lambda (\mu_0 - r) l + (\mu_0 - r) \int_l^T \eta_t dt + \lambda \mu_1 \int_0^l X_t dt + \mu_1 \int_l^T \eta_t X_t dt$$

$$- \frac{1}{2}\lambda^2 \sigma_p^2 l - \frac{\sigma_p^2}{2}\int_l^T \eta_t^2 dt + \int_0^T \eta_t \sigma_p dB_t,$$

which implies

$$E(\log W_T) = \log W_0 + rT + \lambda (\mu_0 - r) l + (\mu_0 - r) \int_l^T E(\eta_t) dt + \lambda \mu_1 \int_0^l E(X_t) dt$$

$$+ \mu_1 \int_l^T E(\eta_t X_t) dt - \frac{1}{2}\lambda^2 \sigma_p^2 l - \frac{\sigma_p^2}{2}\int_l^T E\left(\eta_t^2\right) dt + \int_0^T E(\eta_t) \sigma_p E(dB_t).$$

By Propositions 2.1, 2.2 and 2.4, we note that $E(\eta_t)$, $E(\eta_t^2)$ and $E(\eta_t X_t)$ are all independent of time t. Since $E(dB_t) = 0$, $E(X_t) = -\frac{\theta_0}{\theta_1}$ by Lemma 2.1, we derive

$$E(\log W_T) = c_3 + (T - l)\left[(\mu_0 - r) E(\eta_t) - \frac{\sigma_p^2}{2} E\left(\eta_t^2\right) + \mu_1 E(\eta_t X_t)\right],$$

where c_3 is a constant, i.e.,

$$c_3 = \log W_0 + rT + l\left[\lambda(\mu_0 - r) - \frac{1}{2}\lambda^2\sigma_p^2 - \frac{\lambda\mu_1\theta_0}{\theta_1}\right],$$

Equation (2.15) is proved. Let us define c_4, c_5, c_6, c_7, c_8 and c_9 as below

$$\begin{aligned} c_4 &= \Phi\left(\frac{\epsilon - b_2}{\sigma_2}\right) \quad \& \quad c_5 = \Phi\left(\frac{-b_2}{\sigma_2}\right) \quad \& \quad c_6 = \Phi\left(\frac{-\epsilon - b_2}{\sigma_2}\right) \\ c_7 &= \Psi\left(\frac{\epsilon - b_2}{\sigma_2}\right) \quad \& \quad c_8 = \Psi\left(\frac{-b_2}{\sigma_2}\right) \quad \& \quad c_9 = \Psi\left(\frac{-\epsilon - b_2}{\sigma_2}\right) \end{aligned}$$

then $E(\eta_t)$, $E(\eta_t^2)$ and $E\left(\eta_t M_t^{(s,l)}\right)$ are

$$\begin{aligned} &E(\eta_t) = 1 - c_4 + \delta_1(c_4 - c_5) + \delta_2(c_5 - c_6), \\ &E\left(\eta_t^2\right) = 1 - c_4 + \delta_1^2(c_4 - c_5) + \delta_2^2(c_5 - c_6), \\ &E\left(\eta_t M_t^{(s,l)}\right) \\ &= 1 - (b_2 c_4 + \sigma_2 c_7) + \delta_1(b_2 c_4 + \sigma_2 c_7 - b_2 c_5 - \sigma_2 c_8) + \delta_2(b_2 c_5 + \sigma_2 c_8 - b_2 c_6 - \sigma_2 c_9). \end{aligned}$$

Equation (2.15) can be re-written as

$$
\begin{aligned}
&\frac{E\left(\log W_T\right)-c_3}{T-L} \\
&=\left(\mu_0-r\right)E\left(\eta_t\right)-\frac{\sigma_p^2}{2}E\left(\eta_t^2\right)+\mu_1 E\left(\eta_t X_t\right) \\
&=\left(\mu_0-r\right)E\left(\eta_t\right)-\frac{\sigma_p^2}{2}E\left(\eta_t^2\right)+\mu_1\left(b_1-\frac{\sigma_{12}b_2}{\sigma_2^2}\right)E\left(\eta_t\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}E\left(\eta_t M_t^{(s,l)}\right) \\
&=\left(\mu_0-r+\mu_1 b_1-\frac{\mu_1\sigma_{12}b_2}{\sigma_2^2}\right)E\left(\eta_t\right)-\frac{\sigma_p^2}{2}E\left(\eta_t^2\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}E\left(\eta_t M_t^{(s,l)}\right) \\
&=\left(\mu_0-r+\mu_1 b_1-\frac{\mu_1\sigma_{12}b_2}{\sigma_2^2}\right)\left[1-c_4+\delta_1\left(c_4-c_5\right)+\delta_2\left(c_5-c_6\right)\right] \\
&-\frac{\sigma_p^2}{2}\left[1-c_4+\delta_1^2\left(c_4-c_5\right)+\delta_2^2\left(c_5-c_6\right)\right]+\frac{\mu_1\sigma_{12}}{\sigma_2^2} \\
&\left[1-\left(b_2c_4+\sigma_2c_7\right)+\delta_1\left(b_2c_4+\sigma_2c_7-b_2c_5-\sigma_2c_8\right)+\delta_2\left(b_2c_5+\sigma_2c_8-b_2c_6-\sigma_2c_9\right)\right] \\
&=-\frac{\sigma_p^2}{2}\left(c_4-c_5\right)\delta_1^2 \\
&+\left[\left(\mu_0-r+\mu_1 b_1-\frac{\mu_1\sigma_{12}b_2}{\sigma_2^2}\right)\left(c_4-c_5\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(b_2c_4+\sigma_2c_7-b_2c_5-\sigma_2c_8\right)\right]\delta_1 \\
&-\frac{\sigma_p^2}{2}\left(c_5-c_6\right)\delta_2^2 \\
&+\left[\left(\mu_0-r+\mu_1 b_1-\frac{\mu_1\sigma_{12}b_2}{\sigma_2^2}\right)\left(c_5-c_6\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(b_2c_5+\sigma_2c_8-b_2c_6-\sigma_2c_9\right)\right]\delta_2 \\
&+\left[\left(\mu_0-r+\mu_1 b_1-\frac{\mu_1\sigma_{12}b_2}{\sigma_2^2}-\frac{\sigma_p^2}{2}\right)\left(1-c_4\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(1-\left(b_2c_4+\sigma_2c_7\right)\right)\right] \\
&=-\frac{\sigma_p^2}{2}\left(c_4-c_5\right)\delta_1^2+\left[\left(\mu_0-r+\mu_1 b_1\right)\left(c_4-c_5\right)+\frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_7-c_8\right)\right]\delta_1 \\
&-\frac{\sigma_p^2}{2}\left(c_5-c_6\right)\delta_2^2+\left[\left(\mu_0-r+\mu_1 b_1\right)\left(c_5-c_6\right)+\frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_8-c_9\right)\right]\delta_2 \\
&+\left[\left(\mu_0-r+\mu_1 b_1-\frac{\sigma_p^2}{2}\right)\left(1-c_4\right)+\frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(1-b_2-\sigma_2c_7\right)\right],
\end{aligned}
$$

which implies

$$
\begin{aligned}
&E\left(\log W_T\right) \\
&= -\frac{\sigma_p^2}{2}\left(T-l\right)\left(c_4-c_5\right)\delta_1^2 + \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1\right)\left(c_4-c_5\right) + \frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_7-c_8\right)\right]\delta_1 \\
&- \frac{\sigma_p^2}{2}\left(T-l\right)\left(c_5-c_6\right)\delta_2^2 + \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1\right)\left(c_5-c_6\right) + \frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_8-c_9\right)\right]\delta_2 \\
&+ \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1 - \frac{\sigma_p^2}{2}\right)\left(1-c_4\right) + \frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(1-b_2-\sigma_2 c_7\right)\right] + c_3.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&E\left(\log W_T\right) = -\alpha_1^{(\epsilon)}\delta_1^2 + \beta_1^{(\epsilon)}\delta_1 - \alpha_2^{(\epsilon)}\delta_2^2 + \beta_2^{(\epsilon)}\delta_2 + \gamma^{(\epsilon)} \\
&= -\alpha_1^{(\epsilon)}\left(\delta_1 - \frac{\beta_1^{(\epsilon)}}{2\alpha_1^{(\epsilon)}}\right)^2 - \alpha_2^{(\epsilon)}\left(\delta_2 - \frac{\beta_2^{(\epsilon)}}{2\alpha_2^{(\epsilon)}}\right)^2 + \left(\gamma^{(\epsilon)} + \frac{\left(\beta_1^{(\epsilon)}\right)^2}{4\alpha_1^{(\epsilon)}} + \frac{\left(\beta_2^{(\epsilon)}\right)^2}{4\alpha_2^{(\epsilon)}}\right),
\end{aligned}
$$

where $\alpha_1^{(\epsilon)}$, $\beta_1^{(\epsilon)}$, $\alpha_2^{(\epsilon)}$, $\beta_2^{(\epsilon)}$ and $\gamma^{(\epsilon)}$ are all constants depending on ϵ,

$$
\begin{aligned}
\alpha_1^{(\epsilon)} &= \frac{\sigma_p^2}{2}\left(T-l\right)\left(c_4-c_5\right), \\
\beta_1^{(\epsilon)} &= \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1\right)\left(c_4-c_5\right) + \frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_7-c_8\right)\right], \\
\alpha_2^{(\epsilon)} &= \frac{\sigma_p^2}{2}\left(T-l\right)\left(c_5-c_6\right), \\
\beta_2^{(\epsilon)} &= \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1\right)\left(c_5-c_6\right) + \frac{\mu_1\sigma_{12}}{\sigma_2}\left(c_8-c_9\right)\right], \\
\gamma^{(\epsilon)} &= \left(T-l\right)\left[\left(\mu_0 - r + \mu_1 b_1 - \frac{\sigma_p^2}{2}\right)\left(1-c_4\right) + \frac{\mu_1\sigma_{12}}{\sigma_2^2}\left(1-b_2-\sigma_2 c_7\right)\right] + c_3,
\end{aligned}
$$

Equation (2.16) is proved.

Now that we have derived the equation for the expected log-utility of wealth $E\left(\log W_T\right)$, we can calculate optimal estimates of the asset allocation parameters for the SGMA strategy. In order to achieve this goal, we need to maximize $E(\log W_T)$ with respect to both asset allocation parameters δ_1 and δ_2. By Garlappi and Skoulakis (2009), we restrict $\delta_1, \delta_2 \in [0, 1]$, which means there are no-borrowing and no-short-sale constrains. By Equation (2.16), the optimal estimates of asset allocation parameters δ_1^* and δ_2^* are

$$\delta_1^* = \max\left\{0, \min\left(1, \frac{\beta_1^{(\epsilon)}}{2\alpha_1^{(\epsilon)}}\right)\right\} \ \& \ \delta_2^* = \max\left\{0, \min\left(1, \frac{\beta_2^{(\epsilon)}}{2\alpha_2^{(\epsilon)}}\right)\right\} \tag{2.17}$$

as $\alpha_1^{(\epsilon)} > 0$ and $\alpha_2^{(\epsilon)} > 0$ given $\epsilon > 0$. The optimal estimates δ_1^* and δ_2^* are functions of ϵ. These results illustrate that the SGMA is a better investment strategy compared with the MA strategy because it has higher expected utility of wealth for the investor. Theoretically speaking, the asset allocation parameters δ_1 and δ_2 can be any number. $\delta_i \geq 0$ can be interpreted as long portion of stocks and $\delta_i < 0$ can be interpreted as short portion of stocks. Without no-borrowing and no-short-sale constraints, the optimal estimates are $\delta_1^* = \frac{\beta_1^{(\epsilon)}}{2\alpha_1^{(\epsilon)}}$ and $\delta_2^* = \frac{\beta_2^{(\epsilon)}}{2\alpha_2^{(\epsilon)}}$.

2.4 An Investment Algorithm for Single-Asset Portfolio

We propose an investment algorithm for single-asset portfolio. The algorithm will be tested on simulation data and real data in sections 2.5 and 2.6 to evaluate performance of the SGMA strategy. The algorithm contains following steps:

Step 1. Set investment parameters W_0, r, T, ϵ, λ, s and l.

Step 2. Compute model parameters b_1, b_2, σ_1, σ_{12} and σ_2.

Step 3. Compute δ_1^*, δ_2^* and $E(\log W_T)$.

Step 4. Calculate Y_t, $M_t^{(s)}$, $M_t^{(l)}$ and $M_t^{(s,l)}$.

Step 5. Allocate the wealth between risk-free asset and risky asset according to δ_1^* and δ_2^* for each signal.

Step 6. The holding risky asset is sold at the end of the investment horizon T.

2.5 Simulation Studies

We present several numerical examples based on simulated single-asset portfolio. The investment algorithm is tested and compared with the MA strategy as benchmark.

2.5.1 Data Generating Process

We propose a data generating process for n stocks with q predictive variables time series in section 3.5.1. To generate a single-asset portfolio (one stock) with one predictive variable is a special case, i.e., $n = 1$ and $q = 1$.

2.5.2 Simulation Results

The simulated time series data are generated using parameters in table 2.1. The simulation runs 1,000 times. Each time series contains 97,500 observed points by our settings.

Table 2.1: Model parameters for simulating the single-asset portfolio time series

θ_0	Θ_1	σ_x	μ_0	μ_1	σ_p	ρ	λ	T	dt
0.01	-0.253	0.012	0.031	2.072	0.195	-0.073	0	1	0.00001026

The simulation studies are performed under two scenarios ($s = 5$ & $l = 30$ vs. $s = 5$ & $l = 10$). We set initial wealth $W_0 = 1,000,000$ and interest rate $r = 0$. Under each scenario, we test the SGMA strategy based on $\epsilon = 0.005, 0.01$ and 0.05 and compare with the MA strategy. The SGMA strategy performance results are provided in tables 2.2 and 2.3. We first report the expected log-utility of wealth $E(\log W_T)^*$ based on Equation (2.16) with percentage increase compared with the MA strategy. We then report numerical summaries calculate from the simulation results, including the expected log-utility of wealth $E(\log W_T)$, the expected of wealth $E(W_T)$ and the expected return on asset ratio $E(ROA\ \%)$ etc.

Notice that the SGMA strategy not only can increase the investor's expected log-utility of wealth, but also can increase the investor's expected wealth and the

Table 2.2: SGMA strategy performance summary for scenario 1 on simulated single-asset portfolio (1000 run; s = 5; l = 30)

	MA	$SGMA(\epsilon = 0.005)$	$SGMA(\epsilon = 0.01)$	$SGMA(\epsilon = 0.05)$
δ_1^*	na	1	1	1
δ_2^*	na	1	1	1
$E(\log W_T)^*$	13.862017	13.866941	13.871634	13.871634
$\Delta\%\ E(\log W_T)^*$	na	0.04%	0.07%	0.07%
$E(\log W_T)$	13.841958	13.847561	13.850860	13.862915
$\log E(W_T)$	13.851579	13.857920	13.861993	13.878943
$E(W_T)$	1,036,727	1,043,322	1,047,580	1,065,487
$E(ROA\ \%)$	**3.67%**	**4.33%**	**4.76%**	**6.55%**
$SD(W_T)$	147,583	154,083	159,853	192,194
$MAX(W_T)$	2,026,565	2,009,329	1,976,602	1,983,509
$MIN(W_T)$	731,787	717,116	703,789	575,019
$MEDIAN(W_T)$	1,013,047	1,017,706	1,025,376	1,042,961
$E(TRANS\ \#)$	14	35	34	21

expected return on asset ratio from the simulation results. Under scenario 1, the expected log-utility of wealth increases in range 0.04% to 0.07%. The expected return ratio increases from benchmark return 3.67% to 4.33%, 4.76% and 6.55% respectively. Under scenario 2, the expected log-utility of wealth increases in range 0.11% to 0.30%. The expected return ratio increases from benchmark return 4.33% to 5.40%, 6.20%

Table 2.3: SGMA strategy performance summary for scenario 2 on simulated single-asset portfolio (1000 run; s = 5; l = 10)

	MA	$SGMA(\epsilon = 0.005)$	$SGMA(\epsilon = 0.01)$	$SGMA(\epsilon = 0.05)$
δ_1^*	na	1	1	1
δ_2^*	na	1	1	1
$E(\log W_T)^*$	13.863862	13.878948	13.890889	13.90564
$\Delta\%\ E(\log W_T)^*$	na	0.11%	0.19%	0.30%
$E(\log W_T)$	13.848349	13.855505	13.860451	13.872205
$\log E(W_T)$	13.857880	13.868073	13.875630	13.889640
$E(W_T)$	1,043,280	1,053,968	1,061,964	1,076,947
$E(ROA\ \%)$	**4.33%**	**5.40%**	**6.20%**	**7.69%**
$SD(W_T)$	146,569	169,167	186,756	202,185
$MAX(W_T)$	1,689,367	1,796,767	1,949,026	1,995,635
$MIN(W_T)$	719,130	653,375	586,922	515,814
$MEDIAN(W_T)$	1,024,243	1,034,718	1,038,417	1,057,174
$E(TRANS\ \#)$	29	77	65	29

and 7.69% respectively. We observe that the expected wealth has a increase as we expected. At mean time, the median remains steady and could also have a moderate increase. We believe that the observed phenomenon demonstrates the robustness of

our strategy. The fluctuation of the maximum and minimal expected wealth reflects the impact of the introduction of the risk tolerance. The exact reason might be considered as future research.

2.6 Real Data Applications

We present several real data analysis based on high-frequency exchange traded fund (ETF) data. The investment algorithm is tested and compared with the MA strategy as benchmark.

We use PowerShares QQQ Trust Series 1 (QQQ), which is an exchange-traded fund incorporated in the USA. This ETF tracks performance of the Nasdaq 100 Index. It holds large cap U.S. stocks and tends to focus on technology and consumer sector. The holdings are weighted by market capitalization. As of October 6, 2017, there are 107 holding companies. The top 3 holding companies are Apple Inc (AAPL, 11.57%), Microsoft Corp (MSFT, 8.44%) and Amazon.com Inc (AMZN, 6.86%).

2.6.1 Case 1: SGMA Strategy on High-Frequency Exchange Traded Fund with Observed Predictive Variable

We collect both daily second-level QQQ ETF price time series and daily second-level MSFT stock price time series for this study. The collection period is daily trading time from 9:30 am to 4:00 pm (Eastern Time) to ensure high liquid market.

There are numerous ways of dividing the data of training session into training and test periods. We have indeed tried several combinations. The configurations presented below represent the best scenarios we have discovered based on extensive experience. We divide QQQ ETF time series into two data: ETF price P_t training data (9:30 am to 3:00 pm, which contains 19,800 seconds) and ETF price P_t test data (3:00 pm to 4:00 pm, which contains 3,601 seconds). We use the MSFT price time series as predictive variable X_t training data (9:30 am to 3:00 pm, which contains 19,800 seconds). We set initial wealth $W_0 = 10,000$ and interest rate $r = 0$. Suppose that the investor's risk tolerance is 0.000005. We restrict δ_1 and δ_2 in $[0, 1]$, s in $5, 10$ and l in $30, 60, 90, 120, 180, 240$. We use training data to choose model parameters with the highest return. We report both the SGMA strategy performance summary for QQQ ETF on training data and the SGMA strategy evaluation summary for QQQ ETF on test data. Our study spans five days from 10/2/2017 to 10/6/2017.

We use 10/2/2017 as an example first, then we report all results for 5 days. Second-level QQQ ETF price time series on day 1 (10/02/2017) is provided in figure 2.1. The SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017) training data is provided in table 2.4. The SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) test data is provided in table 2.5.

Figure 2.1: Case 1: Second-level QQQ ETF price time series on day 1 (10/02/2017)

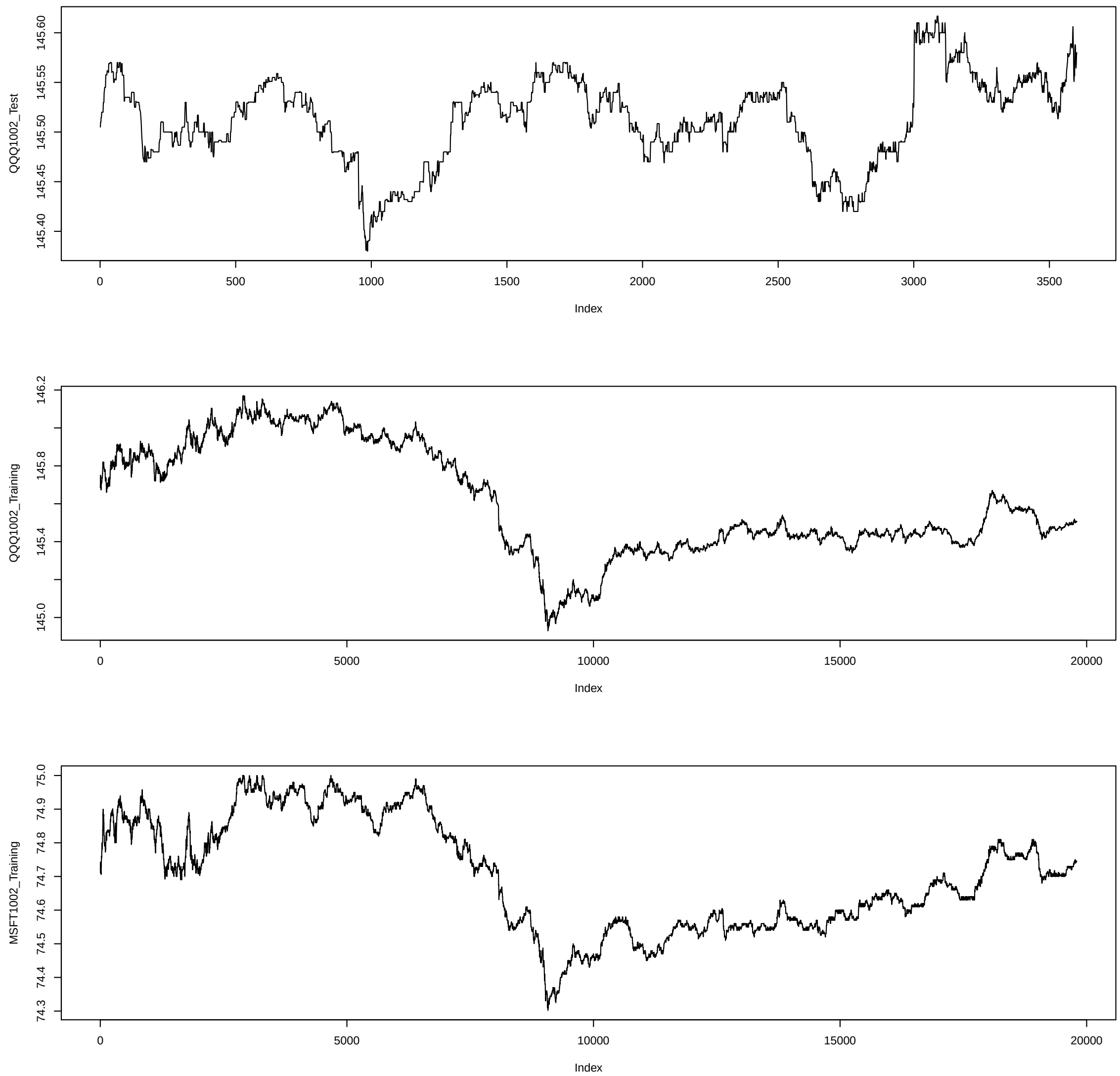


Table 2.4: Case 1: SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017) training data

training data	*day*	10/2/2017
	time	$9:30\ am - 3:00\ pm$
	T	$19,800\ seconds$
	dt	$1\ second$
	correlation	0.94
tuned parameters	s	10
	l	240
	δ_1^*	0.3808
	δ_2^*	0.9977
backward MA	$E(W_T)$	10,014.57271
	return ratio (%)	**0.14573%**
	trans num	225
backward SGMA	$E(W_T)$	10,022.21921
	return ratio (%)	**0.22219%**
	trans num	471

Notice that (1) The SGMA strategy can increase daily return ratio from 0.14573% to 0.22219% on training data, which equals to increase annual return ratio by 21.2%; (2) The SGMA strategy can increase daily return ratio from 0.04950% to 0.06757%

Table 2.5: Case 1: SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) test data

test data	*day*	10/2/2017
	time	$3:00\ pm - 4:00\ pm$
	T	$3,601\ seconds$
	dt	$1\ second$
tuned parameters	s	10
	l	240
	δ_1^*	0.3808
	δ_2^*	0.9977
forward MA	$E(W_T)$	10,004.95043
	return ratio (%)	**0.04950%**
	trans num	38
forward SGMA	$E(W_T)$	10.006.75678
	return ratio (%)	**0.06757%**
	trans num	81

on test data, which equals to increase annual return ratio by 4.7%. Both results illustrate that the SGMA strategy can outperform the MA strategy.

We repeat the study for four more days (10/03/2017 to 10/06/2017). The SG-MA strategy performance summary for QQQ ETF on day 2 (10/03/2017) to day 5

(10/06/2017) training data is provided in table 2.6. The SGMA strategy evaluation summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data is provided in table 2.7.

Table 2.6: Case 1: SGMA strategy performance summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data

training data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$
	T	$19,800$ *seconds*	$19,800$ *seconds*	$19,800$ *seconds*	$19,800$ *seconds*
	dt	1 *second*	1 *second*	1 *second*	1 *second*
	correlation	-0.07	0.76	0.93	0.16
tuned parameters	s	10	5	10	10
	l	30	240	60	90
	δ_1^*	1	0.8097	0.7776	1
	δ_2^*	1	0.6858	0.8700	1
backward MA	$E(W_T)$	9,991.60785	10,010.93778	10,024.71374	10,0002.62412
	return ratio (%)	**-0.08392%**	**0.10938%**	**0.24714%**	**0.02624%**
	trans num	783	269	513	380
backward SGMA	$E(W_T)$	9,999.72530	10,011.33440	10,030.39858	10,010.00331
	return ratio (%)	**-0.00275%**	**0.11334%**	**0.30399%**	**0.10003%**
	trans num	1,941	512	1,190	890

The SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data is provided in figure 2.2. The SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017)

Table 2.7: Case 1: SGMA strategy evaluation summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data

test data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$
	T	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*
	dt	1 *second*	1 *second*	1 *second*	1 *second*
tuned parameters	s	10	5	10	10
	l	30	240	60	90
	δ_1^*	1	0.8097	0.7776	1
	δ_2^*	1	0.6858	0.8700	1
forward MA	$E(W_T)$	9,992.41747	10,001.55533	10,006.94200	9,998.58244
	return ratio (%)	**-0.07583%**	**0.01555%**	**0.06942%**	**-0.01418%**
	trans num	148	44	93	69
forward SGMA	$E(W_T)$	9,994.12194	10,001.22150	10,007.74697	10,000.54010
	return ratio (%)	**-0.05878%**	**0.01222%**	**0.07747%**	**0.00540%**
	trans num	403	85	231	161

test data is provided in figure 2.3. Notice that the SGMA strategy in general can outperform the MA strategy for both backward investment on training data and forward investment on test data.

Figure 2.2: Case 1: SGMA strategy performance summary plot for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data

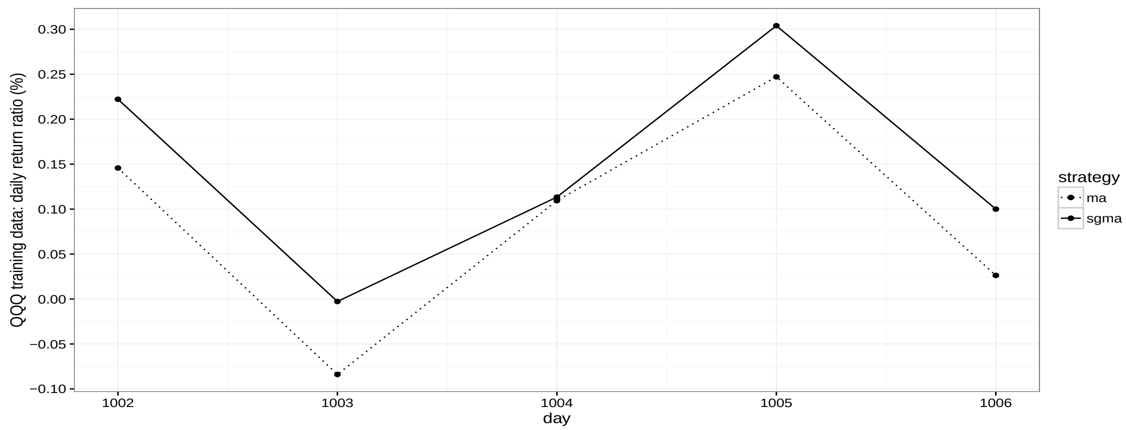
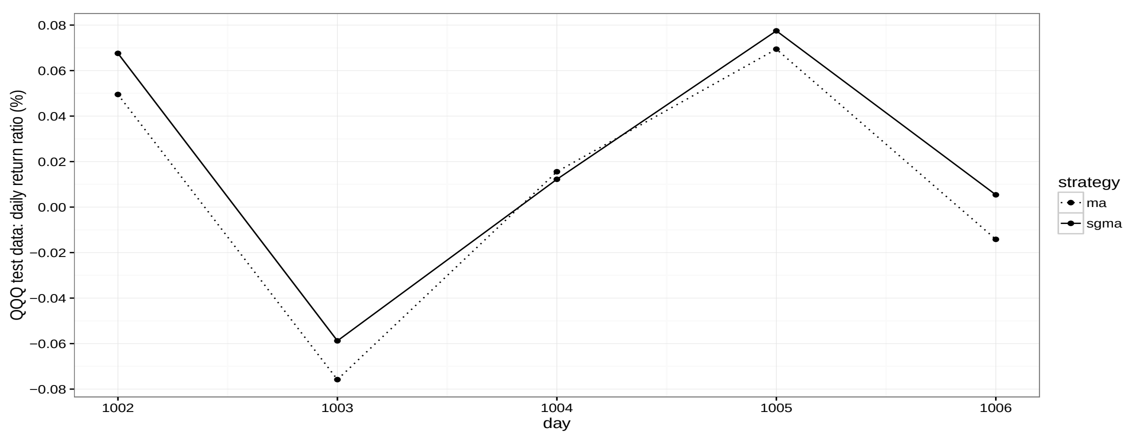


Figure 2.3: Case 1: SGMA strategy evaluation summary plot for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data



2.6.2 Case 2: SGMA Strategy on High-Frequency Exchange Traded Fund without Observed Predictive Variable

We collect daily second-level QQQ ETF price time series for this study. The collection period is daily trading time from 9:30 am to 4:00 pm (Eastern Time) to ensure high liquid market. We divide time series into three data: predictive variable X_t training data (9:30 am to 12:15 pm, which contains 9,900 seconds), ETF price P_t training data (12:15 pm to 3:00 pm, which contains 9,900 seconds) and ETF price test data (3:00 pm to 4:00 pm, which contains 3,601 seconds). We set initial wealth $W_0 = 10,000$ and interest rate $r = 0$. Suppose that the investor's risk tolerance is 0.000005. For easy illustration, we restrict δ_1 and δ_2 in $[0, 1]$, s in $5, 10$ and l in $30, 60, 90, 120, 180, 240$. We use training data to choose model parameters with the highest return. We first report the SGMA strategy performance summary for QQQ ETF on training data, then we report the SGMA strategy evaluation summary for QQQ ETF on test data. Our study spans five days from 10/2/2017 to 10/6/2017.

Let us use 10/2/2017 as an example, then we will report all investment results for 5 days. Second-level QQQ ETF price time series on day 1 (10/02/2017) is provided in figure 2.4. The SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017) training data is provided in table 2.8. The SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) test data is provided in table 2.9.

Figure 2.4: Case 2: Second-level QQQ ETF price time series on day 1 (10/02/2017)

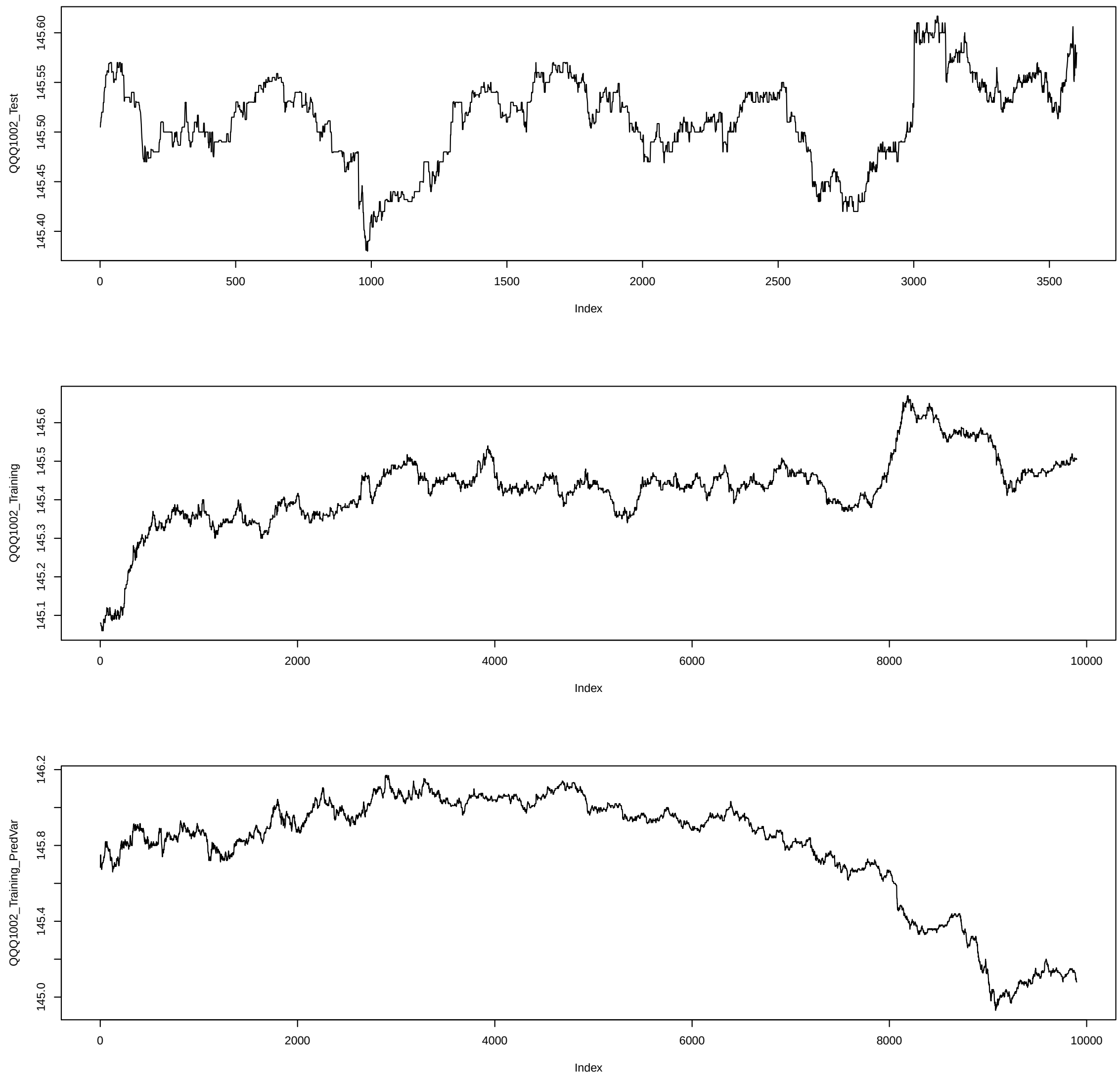


Table 2.8: Case 2: SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017) training data

training data	*day*	10/2/2017
	time	$12:15\ pm - 3:00\ pm$
	T	$9,900$ *seconds*
	dt	1 *second*
	correlation	-0.35
tuned parameters	s	5
	l	120
	δ_1^*	0.7759
	δ_2^*	0.7531
backward MA	$E(W_T)$	10,028.51099
	return ratio (%)	**0.28511%**
	trans num	202
backward SGMA	$E(W_T)$	10,028.81834
	return ratio (%)	**0.28818%**
	trans num	402

Notice that (1) The SGMA strategy can increase daily return ratio from 0.28511% to 0.28818% on training data; (2) The SGMA strategy can increase daily return ratio

Table 2.9: Case 2: SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) test data

test data	*day*	10/2/2017
	time	$3:00\ pm - 4:00\ pm$
	T	$3,601\ seconds$
	dt	$1\ second$
tuned parameters	s	5
	l	120
	δ_1^*	0.7759
	δ_2^*	0.7531
forward MA	$E(W_T)$	10,002.01877
	return ratio (%)	**0.02019%**
	trans num	70
forward SGMA	$E(W_T)$	10.003.08036
	return ratio (%)	**0.03080%**
	trans num	123

from 0.02019% to 0.03080% on test data, which equals to increase annual return ratio by 2.7%. Both results illustrate that the SGMA strategy can outperform the MA

strategy.

We repeat this study for four more days (10/03/2017 to 10/06/2017). The SGMA strategy performance summary for QQQ ETF on day 2 (10/03/2017) to day 5 10/06/2017 training data is provided in table 2.10. The SGMA strategy evaluation summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data is provided in table 2.11.

Table 2.10: Case 2: SGMA strategy performance summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data

training data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$
	T	$9,900\ seconds$	$9,900\ seconds$	$9,900\ seconds$	$9,900\ seconds$
	dt	$1\ second$	$1\ second$	$1\ second$	$1\ second$
	correlation	0.20	-0.32	0.20	0.46
tuned parameters	s	10	5	10	10
	l	30	240	180	90
	δ_1^*	1	0.8685	0.9818	0.9818
	δ_2^*	1	0.8662	0.3787	0.3787
backward MA	$E(W_T)$	9,979.87244	9,998.65146	10,019.34142	9,997.49047
	return ratio (%)	**-0.20128%**	**-0.01349%**	**0.19341%**	**-0.02510%**
	trans num	414	135	121	193
backward SGMA	$E(W_T)$	9,984.55190	9,999.20742	10,019.42782	9,998.56184
	return ratio (%)	**-0.15448%**	**-0.00793%**	**0.19428%**	**-0.01438%**
	trans num	1,103	264	290	493

The SGMA strategy performance summary for QQQ ETF on day 1 (10/02/2017)

Table 2.11: Case 2: SGMA strategy evaluation summary for QQQ ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data

test data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$
	T	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*
	dt	1 *second*	1 *second*	1 *second*	1 *second*
tuned parameters	s	10	5	10	10
	l	30	240	180	90
	δ_1^*	1	0.8685	0.9818	0.9818
	δ_2^*	1	0.8662	0.3787	0.3787
forward MA	$E(W_T)$	9,992.41747	10,001.55533	10,006.93567	9,998.58244
	return ratio (%)	**-0.07583%**	**0.01555%**	**0.06936%**	**-0.01418%**
	trans num	148	44	46	69
forward SGMA	$E(W_T)$	9,994.12194	10,001.35053	10,006.67620	9,999.33436
	return ratio (%)	**-0.05878%**	**0.01351%**	**0.06676%**	**-0.00666%**
	trans num	403	85	108	161

to day 5 (10/06/2017) training data is provided in figure 2.5. The SGMA strategy evaluation summary for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data is provided in figure 2.6. Notice that the SGMA strategy still can outperform the MA strategy for both backward investment on training data and forward investment on test data. It is expected that case 2 study shows under performance as case 1 study involves additional information from Microsoft time series.

Figure 2.5: Case 2: SGMA strategy performance summary plot for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data

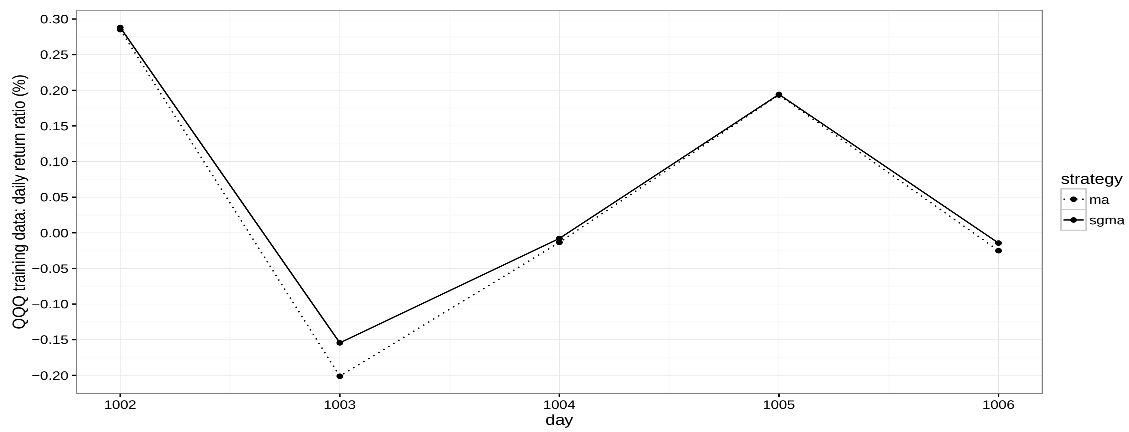
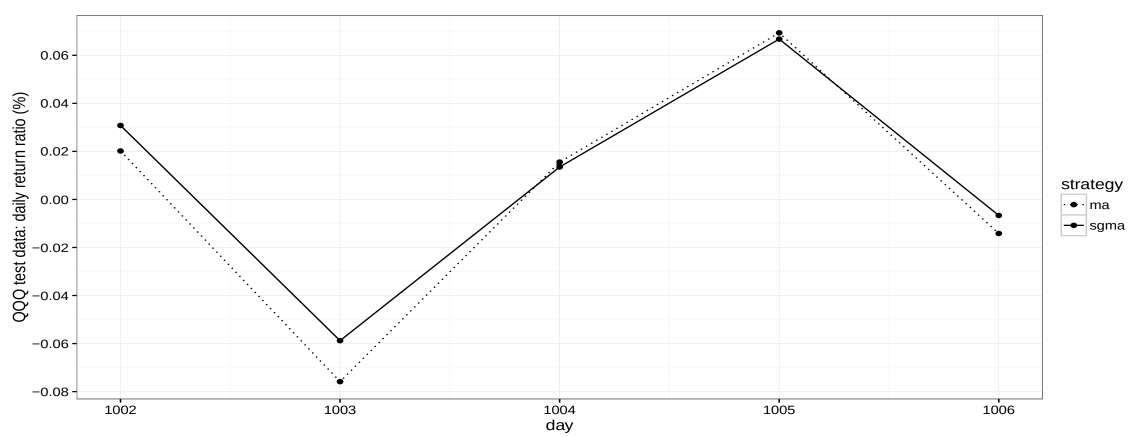


Figure 2.6: Case 2: SGMA strategy evaluation summary plot for QQQ ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data



2.7 Conclusion

Base on the simulation studies and real data analysis, we observe followings: (1) The SGMA strategy can provide more investment options when either $\delta_1^* \neq 1$ or $\delta_2^* \neq 0$, which can overcome the well-known drawback from the MA strategy; (2) The SGMA strategy can increase the investor's expected log-utility of wealth compared with the MA strategy; (3) The SGMA strategy is also able to increase the investor's expected wealth compared with the MA strategy.

Remark for (1): The SGMA strategy can provide more investment options. When the optimal estimate of δ_1^* is closed to 1, this indicates that both SGMA and MA strategies tend to long stock with the lower price. When the optimal solution of δ_1^* is closed to 0, this indicates that the SGMA strategy tends to short stock with the higher price. The MA strategy does not have this option. When the optimal estimate of δ_2^* is closed to 0, this indicates that both SGMA and MA strategies tend to short stock with the higher price. When the optimal solution of δ_2^* is closed to 1, this indicates that the SGMA strategy tends to long stock with the lower price. The MA strategy does not have this option. This also explains why the SGMA strategy is a better investment strategy.

3 Statistical Modeling and Multi-Asset Generalized Moving Average Crossover (MGMA) Strategy

In this chapter, we study statistical modeling and moving average strategy for multiple-asset portfolio. Notice that the moving average in technical analysis faces difficulty when there are more than one investment signals because it is an all-or-nothing investment strategy. We propose a multi-asset generalized moving average crossover (MGMA) strategy. The MGMA strategy is able to allocate wealth not only between risky asset and risk-free asset, but also among different risky assets with the risk tolerance specified by investor. The MGMA strategy can also increase both the investor's expected utility of wealth and the investor's expected wealth.

Chapter 3 is organized as follows. Section 3.1 introduces the general model with the MGMA strategy. Section 3.2 provides all preliminary Lemmas for analytic results. we present main analytic results in section 3.3. An investment algorithm for multi-asset portfolio is given in section 3.4. Section 3.5 provides simulation study results. Section 3.5.1 gives data generating process. Real data analysis is given in section 3.6. An algorithm to estimate model parameters is proposed in section 3.6.1. The conclusion is presented in section 3.7.

3.1 The Model and The MGMA Strategy

Suppose that there are $n+1$ assets in the market. First one is a risk-free asset. For example, cash or money market account with a constant interest rate r. Other n assets are risky assets. For example, stocks or indices representing aggregate equity market. A multi-asset portfolio contains n risky assets. Wealth can be allocated not only between the risk-free asset and the risky asset, but also among risky assets.

We follow Huang and Liu (2007) to define the general model for multi-asset portfolio with multiple predictive variables. Suppose that the price of risk-free asset P_t^f at any time t satisfies

$$dP_t^f = rP_t^f dt. \tag{3.1}$$

Suppose that there are q predictive variables can be accurately observed at continuously times, then the vector of n risky asset prices $\boldsymbol{p}_t$ at any time t satisfies

$$d\boldsymbol{p}_t = \operatorname{diag}(\boldsymbol{p}_t)\left\{ (\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t)\, dt + \mathbb{V}_p d\boldsymbol{b}_t \right\}, \tag{3.2}$$

and the dynamics of the vector of q predictive variables $\boldsymbol{x}_t$ satisfies

$$d\boldsymbol{x}_t = (\boldsymbol{\beta} + \Theta \boldsymbol{x}_t)\, dt + \mathbb{V}_x d\boldsymbol{z}_t, \tag{3.3}$$

where

$$\boldsymbol{p}_t = \begin{pmatrix} p_{1t} \\ \vdots \\ p_{nt} \end{pmatrix}, \quad \mathrm{diag}(\boldsymbol{p}_t) = \begin{pmatrix} p_{1t} & 0 & \dots & 0 \\ 0 & p_{2t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_{nt} \end{pmatrix}, \quad \boldsymbol{x}_t = \begin{pmatrix} x_{1t} \\ \vdots \\ x_{qt} \end{pmatrix},$$

and

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \mathbb{U} = \begin{pmatrix} u_{11} & \dots & u_{1q} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nq} \end{pmatrix}, \quad \mathbb{V}_p = \begin{pmatrix} v_{11}^p & \dots & v_{1n}^p \\ \vdots & \ddots & \vdots \\ v_{n1}^p & \dots & v_{nn}^p \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_q \end{pmatrix},$$

and

$$\Theta = \begin{pmatrix} \theta_{11} & \dots & \theta_{1q} \\ \vdots & \ddots & \vdots \\ \theta_{q1} & \dots & \theta_{qq} \end{pmatrix}, \quad \mathbb{V}_x = \begin{pmatrix} v_{11}^x & \dots & v_{1q}^x \\ \vdots & \ddots & \vdots \\ v_{q1}^x & \dots & v_{qq}^x \end{pmatrix}, \quad \boldsymbol{b}_t = \begin{pmatrix} b_{1t} \\ \vdots \\ b_{nt} \end{pmatrix}, \quad \boldsymbol{z}_t = \begin{pmatrix} z_{1t} \\ \vdots \\ z_{qt} \end{pmatrix}.$$

The vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and matrices $\mathbb{U}$, Θ, $\mathbb{V}_p$, $\mathbb{V}_x$ are all parameters. The vectors $\boldsymbol{b}_t$ and $\boldsymbol{z}_t$ are multi-dimensional standard Brownian motion, such that

$$Var\,(\boldsymbol{b}_t) = t\mathbb{I}_n, \quad Var\,(\boldsymbol{z}_t) = t\mathbb{I}_q, \quad Corr\,(\boldsymbol{b}_t, \boldsymbol{z}_t) = \begin{pmatrix} \rho_{11} & \dots & \rho_{1q} \\ \vdots & \ddots & \vdots \\ \rho_{n1} & \dots & \rho_{nq} \end{pmatrix} \triangleq \mathbb{V}_{\boldsymbol{bz}},$$

where $\mathbb{I}_n$ and $\mathbb{I}_q$ are identity matrices. Each predictive variable x_{it} $(i = 1, \ldots, q)$ is assumed to be stationary process for $t \geq 0$. In order to ensure x_{it} be mean-reverting process, Θ has to be symmetric negative definite, i.e., $\Theta = \Theta^T$ and $\Theta < 0$. We also assume $\Theta \mathbb{V}_x = \mathbb{V}_x \Theta$ in theoretical analysis.

Now let us define the MGMA strategy. We first define some notations for kth stock $(k = 1, \ldots, n)$ in the market. Let p_{kt} be the real stock price and y_{kt} be the log-transformed stock price, i.e., $y_{kt} = \log p_{kt}$. Given that $h > 0$ is a lag or lookback period, a continuous time version of the moving average of the log-transformed stock price at any time t is defined as

$$m_{kt}^{(h)} = \frac{1}{h} \int_{t-h}^{t} y_{ku} du, \tag{3.4}$$

i.e., the average log transformed stock price over time period $[t-h, t]$. Let $m_{kt}^{(s,l)}$ be difference between $m_{kt}^{(s)}$ and $m_{kt}^{(l)}$, where $s > 0$ is a short term lookback period and l is a long time lookback period $(l > s)$, i.e.,

$$m_{kt}^{(s,l)} = m_{kt}^{(s)} - m_{kt}^{(l)}. \tag{3.5}$$

Then the MA strategy for single-asset portfolio which contains only kth stock is defined as

$$\tau_{kt} = \begin{cases} 0, & \text{if } m_{kt}^{(s,l)} \in \tilde{\Omega}_1, \\ \\ 1, & \text{if } m_{kt}^{(s,l)} \in \tilde{\Omega}_2. \end{cases} \tag{3.6}$$

where $\tilde{\Omega}_i$ is defined as

$$\tilde{\Omega}_i = \begin{cases} (-\infty, 0), & \text{if } i = 1, \\ [0, \infty), & \text{if } i = 2. \end{cases}$$

Define Ω_i as

$$\Omega_i = \begin{cases} (-\infty, -\epsilon), & \text{if } i = 1, \\ [-\epsilon, 0), & \text{if } i = 2, \\ [0, \epsilon], & \text{if } i = 3, \\ (\epsilon, \infty), & \text{if } i = 4. \end{cases} \tag{3.7}$$

where $\epsilon > 0$ is the investor specified risk tolerance, then the SGMA strategy for single-asset portfolio which contains only kth stock is

$$\tilde{\eta}_{kt} = \begin{cases} 0, & \text{if } m_{kt}^{(s,l)} \in \Omega_1, \\ \delta_2, & \text{if } m_{kt}^{(s,l)} \in \Omega_2, \\ \delta_1, & \text{if } m_{kt}^{(s,l)} \in \Omega_3, \\ 1, & \text{if } m_{kt}^{(s,l)} \in \Omega_4. \end{cases} \tag{3.8}$$

where δ_1 and δ_2 are corresponding asset allocation parameters. Let $\boldsymbol{p}_t$ be the vector of n stock prices, $\boldsymbol{y}_t$ be the vector of n log transformed stock prices and $\boldsymbol{m}_t^{(s,l)}$ be the vector of n differences between the moving averages, i.e.,

$$\boldsymbol{p}_t = \begin{pmatrix} p_{1t} \\ \vdots \\ p_{nt} \end{pmatrix}, \quad \boldsymbol{y}_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{nt} \end{pmatrix}, \quad \boldsymbol{m}_t^{(s,l)} = \begin{pmatrix} m_{1t}^{(s,l)} \\ \vdots \\ m_{nt}^{(s,l)} \end{pmatrix}.$$

Given $\epsilon > 0$ and let $d = \{1, 2, 3, 4\}$ and $i_k \in d$, $k = 1, \ldots, n$, we define $\Omega_{(i_1,\ldots,i_n)} = \Omega_{i_1} \times \ldots \times \Omega_{i_n}$, where Ω_i is defined in Equation (3.7). Let η_{kt} be the MGMA strategy for kth risky asset in a multi-asset portfolio. Let $\delta_{k,(i_1,\ldots,i_n)}$ be asset allocation parameter for kth risky asset in the multi-asset portfolio. Suppose that $\boldsymbol{\eta}_t$ is the vector based MGMA strategy and $\boldsymbol{\delta}_{(i_1,\ldots,i_n)}$ is the vector of n asset allocation parameters, i.e.,

$$\boldsymbol{\eta}_t = \begin{pmatrix} \eta_{1t} \\ \vdots \\ \eta_{nt} \end{pmatrix}, \quad \boldsymbol{\delta}_{(i_1,\ldots,i_n)} = \begin{pmatrix} \delta_{1,(i_1,\ldots,i_n)} \\ \vdots \\ \delta_{n,(i_1,\ldots,i_n)} \end{pmatrix},$$

then for $t \geq l$, we define the MGMA strategy $\boldsymbol{\eta}_t$ as

$$\boldsymbol{\eta}_t = \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)} \; \mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}} \left(\boldsymbol{m}_t^{(s,l)} \right), \tag{3.9}$$

where $\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}} \left(\boldsymbol{m}_t^{(s,l)} \right)$ is an indicator function such that

$$\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}} \left(\boldsymbol{m}_t^{(s,l)} \right) = \begin{cases} 1, & \text{if } \boldsymbol{m}_t^{(s,l)} \in \Omega_{(i_1,\ldots,i_n)}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.10}$$

To ensure $\boldsymbol{\eta}_t$ is well defined, for $t < l$, we define $\boldsymbol{\eta}_t$ as a constant vector $\boldsymbol{\lambda}$, i.e., $\boldsymbol{\lambda}^T = (\lambda_1, \ldots, \lambda_n)$ where λ_k is a constant for $k = 1, \ldots, n$ and $\sum_{k=1}^{n} \lambda_k \leq 1$.

The MGMA strategy $\boldsymbol{\eta}_t$ is a market timing strategy that allocates wealth not only between one risk-free asset and one risky asset but also among risky assets with the risk tolerance specified by investor. Notice that the SGMA strategy is a special case of MGMA strategy when the portfolio contains only one risky asset. Theoretically speaking, the asset allocation parameter $\delta_{k,(i_1,\ldots,i_n)}$ can be any number which can be interpreted as long portion of stocks if $\delta_{k,(i_1,\ldots,i_n)} \geq 0$ and short portion of stocks if $\delta_{k,(i_1,\ldots,i_n)} < 0$. Therefore, there are $n * 4^n$ parameters for the MGMA strategy on a multi-asset portfolio which contains n risky-assets.

We give an example of the MGMA strategy for a two-asset portfolio ($n = 2$). The MGMA strategy $\boldsymbol{\eta}_t^T = (\eta_{1t}, \eta_{2t})$ for the two-asset portfolio consists of 32 parameters as shown in table 3.1 below. It is obvious that the MGMA strategy is very complex even for a two-asset portfolio. For easy illustration purpose, we use no-borrowing and no-short-sale constrains, i.e., $\delta_{k,(i_1,\ldots,i_n)} \in [0,1]$ and $\sum_{k=1}^{n} \delta_{k,(i_1,\ldots,i_n)} \leq 1$. We use these constrains to reduce parameters to 5, i.e., $(a_1, a_2, a_3, a_4, a_5 \in [0,1])$ in table 3.2 for implementation.

The MA strategy for a two-asset portfolio $\boldsymbol{\tau}_t^T = (\tau_{1t}, \tau_{2t})$ is provided in table 3.3 as benchmark strategy. We follow the common approach to assign equal weights

Table 3.1: MGMA strategy $\boldsymbol{\eta}_t^T = (\eta_{1t}, \eta_{2t})$ for a two-asset portfolio

(η_{1t}, η_{2t})	$m_{2t}^{(s,l)} \in \Omega_1$	$m_{2t}^{(s,l)} \in \Omega_2$	$m_{2t}^{(s,l)} \in \Omega_3$	$m_{2t}^{(s,l)} \in \Omega_4$
$m_{1t}^{(s,l)} \in \Omega_1$	$\big(\delta_{1,(1,1)}, \delta_{2,(1,1)}\big)$	$\big(\delta_{1,(1,2)}, \delta_{2,(1,2)}\big)$	$\big(\delta_{1,(1,3)}, \delta_{2,(1,3)}\big)$	$\big(\delta_{1,(1,4)}, \delta_{2,(1,4)}\big)$
$m_{1t}^{(s,l)} \in \Omega_2$	$\big(\delta_{1,(2,1)}, \delta_{2,(2,1)}\big)$	$\big(\delta_{1,(2,2)}, \delta_{2,(2,2)}\big)$	$\big(\delta_{1,(2,3)}, \delta_{2,(2,3)}\big)$	$\big(\delta_{1,(2,4)}, \delta_{2,(2,4)}\big)$
$m_{1t}^{(s,l)} \in \Omega_3$	$\big(\delta_{1,(3,1)}, \delta_{2,(3,1)}\big)$	$\big(\delta_{1,(3,2)}, \delta_{2,(3,2)}\big)$	$\big(\delta_{1,(3,3)}, \delta_{2,(3,3)}\big)$	$\big(\delta_{1,(3,4)}, \delta_{2,(3,4)}\big)$
$m_{1t}^{(s,l)} \in \Omega_4$	$\big(\delta_{1,(4,1)}, \delta_{2,(4,1)}\big)$	$\big(\delta_{1,(4,2)}, \delta_{2,(4,2)}\big)$	$\big(\delta_{1,(4,3)}, \delta_{2,(4,3)}\big)$	$\big(\delta_{1,(4,4)}, \delta_{2,(4,4)}\big)$

Table 3.2: Simplified MGMA strategy $\boldsymbol{\eta}_t^T = (\eta_{1t}, \eta_{2t})$ for a two-asset portfolio

(η_{1t}, η_{2t})	$m_{2t}^{(s,l)} \in \Omega_1$	$m_{2t}^{(s,l)} \in \Omega_2$	$m_{2t}^{(s,l)} \in \Omega_3$	$m_{2t}^{(s,l)} \in \Omega_4$
$m_{1t}^{(s,l)} \in \Omega_1$	$(0,0)$	$(0,a_1)$	$(0,a_2)$	$(0,1)$
$m_{1t}^{(s,l)} \in \Omega_2$	$(a_3,0)$	$(a_3[1-a_1(1-a_5)], a_1[1-a_3a_5])$	$(a_3[1-a_2(1-a_5)], a_2[1-a_3a_5])$	$(a_3a_5, 1-a_3a_5)$
$m_{1t}^{(s,l)} \in \Omega_3$	$(a_4,0)$	$(a_4[1-a_1(1-a_5)], a_1[1-a_4a_5])$	$(a_4[1-a_2(1-a_5)], a_2[1-a_4a_5])$	$(a_4a_5, 1-a_4a_5)$
$m_{1t}^{(s,l)} \in \Omega_4$	$(1,0)$	$(1-a_1(1-a_5), a_1(1-a_5))$	$(1-a_2(1-a_5), a_2(1-a_5))$	$(a_5, 1-a_5)$

when there are more than one investment signals. We will use table 3.2 and table 3.3 for our simulation studies and real data analysis in section 3.5 and section 3.6.

Table 3.3: MA strategy $\boldsymbol{\tau}_t^T = (\tau_{1t}, \tau_{2t})$ for a two-asset portfolio

(τ_{1t}, τ_{2t})	$m_{2t}^{(s,l)} \in \tilde{\Omega}_1$	$m_{2t}^{(s,l)} \in \tilde{\Omega}_2$
$m_{1t}^{(s,l)} \in \tilde{\Omega}_1$	$(0,0)$	$(0,1)$
$m_{1t}^{(s,l)} \in \tilde{\Omega}_2$	$(1,0)$	$(0.5,0.5)$

The MGMA strategy from an asset allocation perspective now becomes finding optimal $\boldsymbol{\eta}_t$ that maximizes the investor's expected log-utility of wealth

$$\max_{\boldsymbol{\eta}_t} E\left(\log w_T\right), \tag{3.11}$$

subject to a budget constraint

$$\frac{dw_t}{w_t} = r dt + \boldsymbol{\eta}_t^T \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - r\mathbb{1}_n\right) dt + \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t, \tag{3.12}$$

given an initial wealth w_0 for multi-asset portfolio, a constant rate of interest r, an investment horizon T and $\mathbb{1}_n^T = (1, \dots, 1)$.

3.2 Preliminaries

All preliminary Lemmas in this section are used to derive analytical results in next section. We use matrix exponential properties to prove all Lemmas. We also provide detailed proofs. The matrix exponential definition and properties are presented in Appendix A.

Lemma 3.1 *Let Θ be symmetric negative definite. If Θ and $\mathbb{V}_x$ are exchangeable, i.e., $\Theta\mathbb{V}_x = \mathbb{V}_x\Theta$, then Θ and $e^{t\Theta}$, $e^{t\Theta}$ and $\mathbb{V}_x$, $\mathbb{V}_x^T$ and $e^{t\Theta}$ are also exchangeable.*

Proof:

Since Θ is symmetric negative definite and by definition of matrix exponential,

$$\Theta e^{t\Theta} = \Theta \sum_{k=0}^{\infty} \frac{1}{k!}(t\Theta)^k = \sum_{k=0}^{\infty} \frac{1}{k!}(t\Theta)^k \Theta = e^{t\Theta}\Theta.$$

Since $\Theta\mathbb{V}_x = \mathbb{V}_x\Theta$ and Θ is symmetric negative definite, then $\Theta^k\mathbb{V}_x = \mathbb{V}_x\Theta^k$, which implies

$$e^{t\Theta}\mathbb{V}_x = \sum_{k=0}^{\infty}\frac{1}{k!}(t\Theta)^k\mathbb{V}_x = \mathbb{V}_x\sum_{k=0}^{\infty}\frac{1}{k!}(t\Theta)^k = \mathbb{V}_x e^{t\Theta}.$$

Therefore

$$\mathbb{V}_x^T e^{t\Theta} = \mathbb{V}_x^T\left(e^{t\Theta}\right)^T = \left(e^{t\Theta}\mathbb{V}_x\right)^T = \left(\mathbb{V}_x e^{t\Theta}\right)^T = \left(e^{t\Theta}\right)^T\mathbb{V}_x^T = e^{t\Theta^T}\mathbb{V}_x^T = e^{t\Theta}\mathbb{V}_x^T.$$

The Lemma is proved.

Lemma 3.2 *Let $\boldsymbol{x}_t$ be the vector of predictive variables in the market. Let $\boldsymbol{\mu_x}$ be the vector of expectation of $\boldsymbol{x}_t$. Let $\Sigma_{\boldsymbol{x}}$ be the variance-covariance matrix of $\boldsymbol{x}_t$, then $\boldsymbol{x}_t$ follows multi-normal distribution where*

$$\boldsymbol{x}_t = e^{t\Theta}\boldsymbol{x}_0 - \left(\mathbb{I}_q - e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta} + \int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u,$$

and

$$\begin{aligned}
\boldsymbol{\mu_x} &= -\Theta^{-1}\boldsymbol{\beta},\\
\Sigma_{\boldsymbol{x}} &= -\frac{1}{2}\mathbb{V}_x\Theta^{-1}\mathbb{V}_x^T,\\
Cov\left(\boldsymbol{x}_t, \boldsymbol{x}_s\right) &= -\frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{|t-s|\Theta}\mathbb{V}_x^T.
\end{aligned}$$

Proof:

Suppose that $\mathbb{I}_q$ is identity matrix and $\mathbb{0}_n^T = (0, \ldots, 0)$, then

$$d\left(e^{-u\Theta}\boldsymbol{x}_u\right) = d\left(e^{-u\Theta}\right)\boldsymbol{x}_u + e^{-u\Theta}d\boldsymbol{x}_u = -\Theta e^{-u\Theta}\boldsymbol{x}_u du + e^{-u\Theta}d\boldsymbol{x}_u$$

$$= -\Theta e^{-u\Theta}\boldsymbol{x}_u du + e^{-u\Theta}\left(\left(\boldsymbol{\beta} + \Theta\boldsymbol{x}_u\right)du + \mathbb{V}_x d\boldsymbol{z}_u\right) = e^{-u\Theta}\boldsymbol{\beta}du + e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u,$$

which implies

$$e^{-u\Theta}\boldsymbol{x}_u\mid_0^t= e^{-t\Theta}\boldsymbol{x}_t-\boldsymbol{x}_0=\int_0^t d\left(e^{-u\Theta}\boldsymbol{x}_u\right)=\int_0^t e^{-u\Theta}\boldsymbol{\beta}du+\int_0^t e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u$$
$$=-\Theta^{-1}e^{-u\Theta}\mid_0^t\boldsymbol{\beta}+\int_0^t e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u=\left(-\Theta^{-1}e^{-t\Theta}+\Theta^{-1}\right)\boldsymbol{\beta}+\int_0^t e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u$$
$$=-\left(e^{-t\Theta}-\mathbb{I}_q\right)\Theta^{-1}\boldsymbol{\beta}+\int_0^t e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u.$$

Therefore

$$e^{-t\Theta}\boldsymbol{x}_t=\boldsymbol{x}_0-\left(e^{-t\Theta}-\mathbb{I}_q\right)\Theta^{-1}\boldsymbol{\beta}+\int_0^t e^{-u\Theta}\mathbb{V}_x d\boldsymbol{z}_u,$$

and

$$\boldsymbol{x}_t=e^{t\Theta}\boldsymbol{x}_0-\left(\mathbb{I}_q-e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta}+\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u.$$

As a result, $\boldsymbol{x}_t$ follows multi-normal distribution.

Given $E\left(d\boldsymbol{z}_u\right)=\mathbb{0}_n$,

$$\boldsymbol{\mu}_x=E\left(\boldsymbol{x}_t\right)=e^{t\Theta}E\left(\boldsymbol{x}_0\right)-\left(\mathbb{I}_q-e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta}+\int_0^t e^{(t-u)\Theta}\mathbb{V}_x E\left(d\boldsymbol{z}_u\right)$$
$$=e^{t\Theta}E\left(\boldsymbol{x}_0\right)-\left(\mathbb{I}_q-e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta},$$

since $\boldsymbol{x}_t$ is stationary process, i.e., $E\left(\boldsymbol{x}_t\right)=E\left(\boldsymbol{x}_0\right)$, then it can easy to check $\boldsymbol{\mu}_x=-\Theta^{-1}\boldsymbol{\beta}$ is a solution.

Given $E\left(d\boldsymbol{z}_u d\boldsymbol{z}_u^T\right)=du\mathbb{I}_q$, $\mathbb{V}_x$ and $e^{t\Theta}$ are exchangeable and Θ is symmetric, then the variance-covariance matrix $\Sigma_{\boldsymbol{x}}$ is

$$\Sigma_{\boldsymbol{x}}=Var\left(\boldsymbol{x}_t\right)=Var\left(e^{t\Theta}\boldsymbol{x}_0-\left(\mathbb{I}_q-e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta}+\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right)$$
$$=e^{t\Theta}Var\left(\boldsymbol{x}_0\right)e^{t\Theta^T}+Var\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right),$$

where

$$\begin{aligned}
&Var\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right)\\
&= E\left[\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u \int_0^t d\boldsymbol{z}_u^T \mathbb{V}_x^T e^{(t-u)\Theta^T}\right] - E\left[\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right] E\left[\int_0^t d\boldsymbol{z}_u^T \mathbb{V}_x^T e^{(t-u)\Theta^T}\right]\\
&= \int_0^t e^{(t-u)\Theta}\mathbb{V}_x \mathbb{V}_x^T e^{(t-u)\Theta^T} du = \int_0^t \mathbb{V}_x e^{(t-u)\Theta} e^{(t-u)\Theta^T} du \mathbb{V}_x^T = \mathbb{V}_x \int_0^t e^{2(t-u)\Theta} du \mathbb{V}_x^T\\
&= \mathbb{V}_x \left[-\frac{1}{2}\Theta^{-1} e^{2(t-u)\Theta} \mid_0^t\right] \mathbb{V}_x^T = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T + \frac{1}{2}\mathbb{V}_x \Theta^{-1} e^{t\Theta} e^{t\Theta^T}\mathbb{V}_x^T\\
&= -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T + \frac{1}{2} e^{t\Theta}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T e^{t\Theta^T},
\end{aligned}$$

which implies

$$\Sigma_{\boldsymbol{x}} = Var\left(\boldsymbol{x}_t\right) = e^{t\Theta} Var\left(\boldsymbol{x}_0\right) e^{t\Theta^T} - \frac{1}{2}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T + \frac{1}{2} e^{t\Theta}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T e^{t\Theta^T},$$

since $\boldsymbol{x}_t$ is stationary process, i.e., $Var\left(\boldsymbol{x}_t\right) = Var\left(\boldsymbol{x}_0\right)$, then it can easy to check $\Sigma_x = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T$ is a solution. Moreover,

$$\begin{aligned}
&Cov\left(\boldsymbol{x}_t, \boldsymbol{x}_s\right)\\
&= Cov(e^{t\Theta}\boldsymbol{x}_0 - \left(\mathbb{I}_q - e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta} + \int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u,\\
&e^{s\Theta}\boldsymbol{x}_0 - \left(\mathbb{I}_q - e^{s\Theta}\right)\Theta^{-1}\boldsymbol{\beta} + \int_0^s e^{(s-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u)\\
&= Cov\left(e^{t\Theta}\boldsymbol{x}_0, e^{s\Theta}\boldsymbol{x}_0\right) + Cov\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \int_0^s e^{(s-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right),
\end{aligned}$$

where

$$Cov\left(e^{t\Theta}\boldsymbol{x}_0, e^{s\Theta}\boldsymbol{x}_0\right) = e^{t\Theta} Var\left(\boldsymbol{x}_0\right) e^{s\Theta} = -\frac{1}{2} e^{t\Theta}\mathbb{V}_x \Theta^{-1}\mathbb{V}_x^T e^{s\Theta},$$

and given $E\left(d\boldsymbol{z}_a d\boldsymbol{z}_b^T\right) = 0$ if $a \neq b$,

$$
\begin{aligned}
& Cov\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \int_0^s e^{(s-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u\right) \\
& = E\left[\int_0^{\min(t,s)} e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u \int_0^{\min(t,s)} d\boldsymbol{z}_u^T \mathbb{V}_x^T e^{(s-u)\Theta^T}\right] \\
& = \int_0^{\min(t,s)} e^{(t-u)\Theta}\mathbb{V}_x\mathbb{V}_x^T e^{(s-u)\Theta^T} du = \mathbb{V}_x \int_0^{\min(t,s)} e^{(t-u)\Theta} e^{(s-u)\Theta^T} du \mathbb{V}_x^T \\
& = \mathbb{V}_x \int_0^{\min(t,s)} e^{(t+s-2u)\Theta} du \mathbb{V}_x^T \\
& = \mathbb{V}_x \left[-\frac{1}{2}\Theta^{-1}\left(e^{(t+s-2\min(t,s))\Theta} - e^{(t+s)\Theta}\right)\right]\mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{|t-s|\Theta}\mathbb{V}_x^T + \frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{(t+s)\Theta}\mathbb{V}_x^T.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Cov\left(\boldsymbol{x}_t, \boldsymbol{x}_s\right) &= -\frac{1}{2}e^{t\Theta}\mathbb{V}_x\Theta^{-1}\mathbb{V}_x^T e^{s\Theta} - \frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{|t-s|\Theta}\mathbb{V}_x^T + \frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{(t+s)\Theta}\mathbb{V}_x^T \\
&= -\frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{(t+s)\Theta}\mathbb{V}_x^T - \frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{|t-s|\Theta}\mathbb{V}_x^T + \frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{(t+s)\Theta}\mathbb{V}_x^T \\
&= -\frac{1}{2}\mathbb{V}_x\Theta^{-1}e^{|t-s|\Theta}\mathbb{V}_x^T.
\end{aligned}
$$

The Lemma is proved.

Lemma 3.3 *Let $\boldsymbol{y}_t$ be the vector of log transformed stock prices. Let $\boldsymbol{\mu_y}$ be the vector of expectation of $\boldsymbol{y}_t$, then $\boldsymbol{y}_t$ follows multi-normal distribution, i.e.,*

$$
\boldsymbol{y}_t = \boldsymbol{y}_0 + \int_0^t \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_u - \frac{1}{2}\boldsymbol{\gamma}\right) du + \mathbb{V}_p \boldsymbol{b}_t,
$$

where $\boldsymbol{\gamma}$ *is a column vector of diagonal of matrix* $\mathbb{V}_p\mathbb{V}_p^T$*, i.e.,*

$$\boldsymbol{\gamma} = \begin{pmatrix} \sum_{i=1}^n \left(v_{1i}^p\right)^2 \\ \vdots \\ \sum_{i=1}^n \left(v_{ni}^p\right)^2 \end{pmatrix},$$

and

$$\boldsymbol{\mu_y} = \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)t.$$

Proof:

Since for kth stock,

$$\frac{dp_{kt}}{p_{kt}} = \left(\alpha_k + \sum_{j=1}^q u_{kj}x_{jt}\right)dt + \sum_{i=1}^n v_{ki}^p db_{it},$$

and $(dt)^2 = o(dt), dtdb_{it} = o(dt)$, $db_{it}db_{it} = dt$ and $db_{it}db_{jt} = 0$ if $i \neq j$, then

$$\left(\frac{dp_{kt}}{p_{kt}}\right)^2 = \left(\sum_{i=1}^n v_{ki}^p db_{it}\right)^2 = \sum_{i=1}^n \left(v_{ki}^p\right)^2 dt.$$

Therefore

$$\begin{aligned} d(y_{kt}) = d(\log p_{kt}) &= (\log p_{kt})' \, dp_{kt} + \frac{1}{2}\left(\log p_{kt}\right)'' (dp_{kt})^2 = \frac{dp_{kt}}{p_{kt}} - \frac{1}{2}\left(\frac{dp_{kt}}{p_{kt}}\right)^2 \\ &= \left(\alpha_k + \sum_{j=1}^q u_{kj}x_{jt}\right)dt + \sum_{i=1}^n v_{ki}^p db_{it} - \frac{1}{2}\sum_{i=1}^n \left(v_{ki}^p\right)^2 dt \\ &= \left(\alpha_k + \sum_{j=1}^q u_{kj}x_{jt} - \frac{1}{2}\sum_{i=1}^n \left(v_{ki}^p\right)^2\right)dt + \sum_{i=1}^n v_{ki}^p db_{it}. \end{aligned}$$

As a result, we can derive

$$d(\boldsymbol{y}_t) = \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - \frac{1}{2}\boldsymbol{\gamma}\right) dt + \mathbb{V}_p d\boldsymbol{b}_t, \tag{3.13}$$

where

$$\boldsymbol{\gamma} = \begin{pmatrix} \sum_{i=1}^{n} \left(v_{1i}^{p}\right)^2 \\ \vdots \\ \sum_{i=1}^{n} \left(v_{ni}^{p}\right)^2 \end{pmatrix}.$$

Therefore

$$\boldsymbol{y}_t - \boldsymbol{y}_0 = \int_0^t d(\boldsymbol{y}_u) = \int_0^t \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_u - \frac{1}{2}\boldsymbol{\gamma}\right) du + \int_0^t \mathbb{V}_p d\boldsymbol{b}_u,$$

and

$$\boldsymbol{y}_t = \boldsymbol{y}_0 + \int_0^t \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_u - \frac{1}{2}\boldsymbol{\gamma}\right) du + \mathbb{V}_p \boldsymbol{b}_t.$$

Then

$$\begin{aligned} \boldsymbol{\mu}_y = E\left(\boldsymbol{y}_t\right) &= E\left(\boldsymbol{y}_0\right) + \int_0^t \left(\boldsymbol{\alpha} + \mathbb{U}E\left(\boldsymbol{x}_u\right) - \frac{1}{2}\boldsymbol{\gamma}\right) du + \mathbb{V}_p E\left(\boldsymbol{b}_t\right) \\ &= \boldsymbol{y}_0 + \int_0^t \left(\boldsymbol{\alpha} + \mathbb{U}\left(-\Theta^{-1}\boldsymbol{\beta}\right) - \frac{1}{2}\boldsymbol{\gamma}\right) du + 0 \\ &= \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right) t. \end{aligned}$$

The Lemma is proved.

Lemma 3.4 *Let $\boldsymbol{m}_t^{(h)}$ be the vector of moving averages based on lookback period h. Let $\boldsymbol{m}_t^{(s,l)}$ be the vector of differences between the moving averages based on lookback* *period s and l, then $\boldsymbol{m}_t^{(h)}$ follows multi-normal distribution with mean $E\left(\boldsymbol{m}_t^{(h)}\right)$ and $\boldsymbol{m}_t^{(s,l)}$ follows multi-normal distribution with mean $E\left(\boldsymbol{m}_t^{(s,l)}\right)$, i.e,*

$$E\left(\boldsymbol{m}_t^{(h)}\right) = \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)\left(t - \frac{h}{2}\right),$$

and

$$E\left(\boldsymbol{m}_t^{(s,l)}\right) = \frac{1}{2}(l-s)\left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right),$$

where

$$\boldsymbol{\gamma} = \begin{pmatrix} \sum_{i=1}^{n}\left(v_{1i}^{p}\right)^2 \\ \vdots \\ \sum_{i=1}^{n}\left(v_{ni}^{p}\right)^2 \end{pmatrix}.$$

Proof:

Based on definition of $\boldsymbol{m}_t^{(h)}$,

$$\begin{aligned} E\left(\boldsymbol{m}_t^{(h)}\right) &= E\left(\frac{1}{h}\int_{t-h}^{t}\boldsymbol{y}_u du\right) = \frac{1}{h}\int_{t-h}^{t} E\left(\boldsymbol{y}_u\right) du \\ &= \frac{1}{h}\int_{t-h}^{t}\boldsymbol{y}_0 du + \frac{1}{h}\int_{t-h}^{t}\left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right) u du \\ &= \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)\frac{1}{2h}\left(t^2 - (t-h)^2\right) \\ &= \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)\left(t - \frac{h}{2}\right), \end{aligned}$$

and

$$
\begin{aligned}
E\left(\boldsymbol{m}_t^{(s,l)}\right) &= E\left(\boldsymbol{m}_t^{(s)} - \boldsymbol{m}_t^{(l)}\right) = E\left(\boldsymbol{m}_t^{(s)}\right) - E\left(\boldsymbol{m}_t^{(l)}\right) \\
&= \boldsymbol{y}_0 + \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)\left(t - \frac{s}{2}\right) - \boldsymbol{y}_0 - \left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right)\left(t - \frac{l}{2}\right) \\
&= \frac{1}{2}(l-s)\left(\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right).
\end{aligned}
$$

The Lemma is proved.

Lemma 3.5 *Let $\boldsymbol{z}_u$ and $\boldsymbol{b}_v$ be multi-dimensional standard Brownian motion and $Corr(\boldsymbol{b}_v, \boldsymbol{z}_u) = \mathbb{V}_{\boldsymbol{bz}}$, then*

$$
Cov\left(\boldsymbol{z}_u, \boldsymbol{b}_v\right) = \min(u,v)\mathbb{V}_{\boldsymbol{bz}}^T,
$$

and

$$
Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) = \begin{cases} \mathbb{V}_{\boldsymbol{bz}}^T du, & \textit{if } u < v, \\ 0, & \textit{if } u \geq v. \end{cases}
$$

Proof:

Given that z_{iu} and b_{jv} are two-dimensional standard Brownian motion with correlation coefficient ρ_{ji}, then z_{iu} can be represented by $\rho_{ji}b_{ju} + \sqrt{1-\rho_{ji}^2}b'_{ju}$, where b_{ju} and b'_{ju} are independent, then

$$
\begin{aligned}
Cov\left(z_{iu}, b_{jv}\right) &= Cov\left(\rho_{ji}b_{ju} + \sqrt{1-\rho_{ji}^2}b'_{ju}, b_{jv}\right) \\
&= \rho_{ji}Cov\left(b_{ju}, b_{jv}\right) + \sqrt{1-\rho_{ji}^2}Cov\left(b'_{ju}, b_{jv}\right) \\
&= \rho_{ji}\min(u,v),
\end{aligned}
$$

which implies

$$Cov\left(\boldsymbol{z}_u, \boldsymbol{b}_v\right) = \min(u, v)\mathbb{V}_{\boldsymbol{bz}}^T,$$

and

$$\begin{aligned} Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) &= Cov\left(\boldsymbol{z}_{u+du} - \boldsymbol{z}_u, \boldsymbol{b}_v\right) = Cov\left(\boldsymbol{z}_{u+du}, \boldsymbol{b}_v\right) - Cov\left(\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\ &= \min(u + du, v)\mathbb{V}_{\boldsymbol{bz}}^T - \min(u, v)\mathbb{V}_{\boldsymbol{bz}}^T \\ &= \begin{cases} \mathbb{V}_{\boldsymbol{bz}}^T du, & \text{if } u < v, \\ 0, & \text{if } u \geq v. \end{cases} \end{aligned}$$

The Lemma is proved.

Lemma 3.6 *Let $\boldsymbol{x}_t$ be the vector of predictive variables in the market. Let $\boldsymbol{b}_v$ be multi-dimensional standard Brownian motion, then*

$$Cov\left(\boldsymbol{x}_t, \boldsymbol{b}_v\right) = \begin{cases} -\Theta^{-1}\left(e^{(t-v)\Theta} - e^{t\Theta}\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T, & \textit{if } t \geq v, \\ -\Theta^{-1}\left(\mathbb{I}_q - e^{t\Theta}\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T, & \textit{if } t < v. \end{cases}$$

Proof:

If $t \geq v$, based on Lemmas 3.2 and 3.5,

$$
\begin{aligned}
Cov\left(\boldsymbol{x}_t, \boldsymbol{b}_v\right) &= Cov\left(e^{t\Theta}\boldsymbol{x}_0 - \left(\mathbb{I}_q - e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta} + \int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\
&= Cov\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \boldsymbol{b}_v\right) = \int_0^t e^{(t-u)\Theta}\mathbb{V}_x Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\
&= \int_0^v e^{(t-u)\Theta}\mathbb{V}_x Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) + \int_v^t e^{(t-u)\Theta}\mathbb{V}_x Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\
&= \int_0^v e^{(t-u)\Theta}\mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T du + 0 \\
&= -\Theta^{-1}\left(e^{(t-v)\Theta} - e^{t\Theta}\right)\mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T.
\end{aligned}
$$

If $t < v$, based on Lemmas 3.2 and 3.5,

$$
\begin{aligned}
Cov\left(\boldsymbol{x}_t, \boldsymbol{b}_v\right) &= Cov\left(e^{t\Theta}\boldsymbol{x}_0 - \left(\mathbb{I}_q - e^{t\Theta}\right)\Theta^{-1}\boldsymbol{\beta} + \int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\
&= Cov\left(\int_0^t e^{(t-u)\Theta}\mathbb{V}_x d\boldsymbol{z}_u, \boldsymbol{b}_v\right) = \int_0^t e^{(t-u)\Theta}\mathbb{V}_x Cov\left(d\boldsymbol{z}_u, \boldsymbol{b}_v\right) \\
&= \int_0^t e^{(t-u)\Theta}\mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T du \\
&= -\Theta^{-1}\left(\mathbb{I}_q - e^{t\Theta}\right)\mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T.
\end{aligned}
$$

The Lemma is proved.

Lemma 3.7 *Let $\boldsymbol{x}_t$ be the vector of predictive variables in the market. Let $\boldsymbol{y}_u$ be the vector of log transformed stock prices, then for $t \geq u$,*

$$
Cov\left(\boldsymbol{x}_t, \boldsymbol{y}_u\right) = \Theta^{-1}e^{t\Theta}\left(e^{-u\Theta} - \mathbb{I}_q\right)\mathbb{V}_x\left(\frac{1}{2}\Theta^{-1}\mathbb{V}_x^T\mathbb{U}^T - \mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T\right).
$$

Proof:

If $t \geq u$, based on Lemmas 3.2, 3.3 and 3.6,

$$
\begin{aligned}
&Cov\left(\boldsymbol{x}_t, \boldsymbol{y}_u\right) = Cov\left(\boldsymbol{x}_t, \boldsymbol{y}_0 + \int_0^u \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_v - \frac{1}{2}\boldsymbol{\gamma}\right) dv + \mathbb{V}_p \boldsymbol{b}_u\right) \\
&= \int_0^u Cov\left(\boldsymbol{x}_t, \boldsymbol{x}_v\right) \mathbb{U}^T dv + Cov\left(\boldsymbol{x}_t, \boldsymbol{b}_u\right) \mathbb{V}_p^T \\
&= \int_0^u -\frac{1}{2}\mathbb{V}_x \Theta^{-1} e^{|t-v|\Theta} \mathbb{V}_x^T \mathbb{U}^T dv - \Theta^{-1}\left(e^{(t-u)\Theta} - e^{t\Theta}\right) \mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T \mathbb{V}_p^T \\
&= \int_0^u -\frac{1}{2}\mathbb{V}_x \Theta^{-1} e^{(t-v)\Theta} \mathbb{V}_x^T \mathbb{U}^T dv - \Theta^{-1}\left(e^{(t-u)\Theta} - e^{t\Theta}\right) \mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T \mathbb{V}_p^T \\
&= -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\left(-\Theta^{-1}\left[e^{(t-u)\Theta} - e^{t\Theta}\right]\right) \mathbb{V}_x^T \mathbb{U}^T - \Theta^{-1}\left(e^{(t-u)\Theta} - e^{t\Theta}\right) \mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T \mathbb{V}_p^T \\
&= \Theta^{-1} e^{t\Theta}\left(e^{-u\Theta} - \mathbb{I}_q\right) \mathbb{V}_x \left(\frac{1}{2}\Theta^{-1}\mathbb{V}_x^T \mathbb{U}^T - \mathbb{V}_{\boldsymbol{bz}}^T \mathbb{V}_p^T\right).
\end{aligned}
$$

The Lemma is proved.

Lemma 3.8 *Let $\boldsymbol{x}_t$ be the vector of predictive variables in the market. Let $\boldsymbol{m}_t^{(h)}$ be the vector of moving averages based on lookback period h, then*

$$
Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(h)}\right) = \left(\Theta e^{t\Theta} + \frac{1}{h}\left(\mathbb{I}_q - e^{h\Theta}\right)\right) \mathbb{Q}_3^T,
$$

where

$$
\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \qquad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x \mathbb{V}_{\boldsymbol{bz}}^T \mathbb{V}_p^T, \qquad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x \mathbb{V}_x^T \Theta^{-3}.
$$

Proof:

Based on definition of $\boldsymbol{m}_t^{(h)}$ and Lemma 3.7,

$$
\begin{aligned}
&Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(h)}\right) = Cov\left(\boldsymbol{x}_t, \frac{1}{h}\int_{t-h}^{t} \boldsymbol{y}_u du\right) = \frac{1}{h}\int_{t-h}^{t} Cov\left(\boldsymbol{x}_t, \boldsymbol{y}_u\right) du \\
&= \frac{1}{h}\int_{t-h}^{t} \Theta^{-1} e^{t\Theta}\left(e^{-u\Theta} - \mathbb{I}_q\right)\mathbb{V}_x\left(\frac{1}{2}\Theta^{-1}\mathbb{V}_x^T\mathbb{U}^T - \mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T\right) du \\
&= \frac{1}{h}\Theta^{-1} e^{t\Theta}\left(-\Theta^{-1}\left(e^{-t\Theta} - e^{-(t-h)\Theta}\right) - h\mathbb{I}_q\right)\mathbb{V}_x\left(\frac{1}{2}\Theta^{-1}\mathbb{V}_x^T\mathbb{U}^T - \mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T\right) \\
&= \left(-\frac{1}{h}\Theta^{-2}\left(\mathbb{I}_q - e^{h\Theta}\right) - \Theta^{-1}e^{t\Theta}\right)\mathbb{V}_x\left(\frac{1}{2}\Theta^{-1}\mathbb{V}_x^T\mathbb{U}^T - \mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T\right) \\
&= \left(\Theta e^{t\Theta} + \frac{1}{h}\left(\mathbb{I}_q - e^{h\Theta}\right)\right)\left(\Theta^{-2}\mathbb{V}_x\left(\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T - \frac{1}{2}\Theta^{-1}\mathbb{V}_x^T\mathbb{U}^T\right)\right) \\
&= \left(\Theta e^{t\Theta} + \frac{1}{h}\left(\mathbb{I}_q - e^{h\Theta}\right)\right)\mathbb{Q}_3^T,
\end{aligned}
$$

where

$$
\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \quad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \quad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.
$$

The Lemma is proved.

Lemma 3.9 *Let $\boldsymbol{x}_t$ be the vector of predictive variables in the market. Let $\boldsymbol{m}_t^{(s,l)}$ be the vector of moving averages differences based on lookback period s and l ($l > s$), then*

$$
Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right) = \left(\frac{1}{s}\left(\mathbb{I}_q - e^{s\Theta}\right) - \frac{1}{l}\left(\mathbb{I}_q - e^{l\Theta}\right)\right)\mathbb{Q}_3^T,
$$

where

$$
\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \quad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \quad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.
$$

Proof:

Based on definition of $\boldsymbol{m}_t^{(s,l)}$ and Lemma 3.8,

$$\begin{aligned}
Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right) &= Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s)} - \boldsymbol{m}_t^{(l)}\right) \\
&= Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s)}\right) - Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(l)}\right) \\
&= \left(\Theta e^{t\Theta} + \frac{1}{s}\left(\mathbb{I}_q - e^{s\Theta}\right)\right)\mathbb{Q}_3^T - \left(\Theta e^{t\Theta} + \frac{1}{l}\left(\mathbb{I}_q - e^{l\Theta}\right)\right)\mathbb{Q}_3^T \\
&= \left(\frac{1}{s}\left(\mathbb{I}_q - e^{s\Theta}\right) - \frac{1}{l}\left(\mathbb{I}_q - e^{l\Theta}\right)\right)\mathbb{Q}_3^T,
\end{aligned}$$

where

$$\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \qquad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \qquad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.$$

The Lemma is proved. Notice that $Cov\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right)$ is independent of time t.

Lemma 3.10 *Let $\boldsymbol{y}_t$ be the vector of log transformed stock prices, then*

$$Cov\left(\boldsymbol{y}_u, \boldsymbol{y}_v\right) = \begin{cases} \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_3^T + \mathbb{Q}_3\mathbb{K}_2(v)\mathbb{U}^T + v\mathbb{V}_p\mathbb{V}_p^T, & \text{if } u \geq v, \\ \\ \mathbb{Q}_3\mathbb{K}_1(v,u)\mathbb{U}^T + \mathbb{U}\mathbb{K}_2(u)\mathbb{Q}_3^T + u\mathbb{V}_p\mathbb{V}_p^T, & \text{if } u < v. \end{cases}$$

where

$$\mathbb{K}_1(u,v) = -e^{(u-v)\Theta} + e^{u\Theta} - v\Theta, \qquad \mathbb{K}_2(v) = e^{v\Theta} - v\Theta - \mathbb{I}_q,$$

$$\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \qquad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \qquad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.$$

Proof:

For $u \geq v$, based on Lemmas 3.2, 3.3 and 3.6,

$$
\begin{aligned}
& Cov\left(\boldsymbol{y}_u, \boldsymbol{y}_v\right) \\
& = Cov\left(\boldsymbol{y}_0 + \int_0^u \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_a - \frac{1}{2}\boldsymbol{\gamma}\right) da + \mathbb{V}_p \boldsymbol{b}_u, \boldsymbol{y}_0 + \int_0^v \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_j - \frac{1}{2}\boldsymbol{\gamma}\right) dj + \mathbb{V}_p \boldsymbol{b}_v\right) \\
& = Cov\left(\mathbb{U}\int_0^u \boldsymbol{x}_a da + \mathbb{V}_p \boldsymbol{b}_u, \mathbb{U}\int_0^v \boldsymbol{x}_j dj + \mathbb{V}_p \boldsymbol{b}_v\right) \\
& = \mathbb{U}\int_0^u da \int_0^v Cov\left(\boldsymbol{x}_a, \boldsymbol{x}_j\right) dj \mathbb{U}^T + \mathbb{U}\int_0^u Cov\left(\boldsymbol{x}_a, \boldsymbol{b}_v\right) da \mathbb{V}_p^T \\
& + \mathbb{V}_p \int_0^v Cov\left(\boldsymbol{b}_u, \boldsymbol{x}_j\right) dj \mathbb{U}^T + \mathbb{V}_p Cov\left(\boldsymbol{b}_u, \boldsymbol{b}_v\right) \mathbb{V}_p^T,
\end{aligned}
$$

where

$$
\begin{aligned}
& \int_0^u da \int_0^v Cov\left(\boldsymbol{x}_a, \boldsymbol{x}_j\right) dj \\
& = \int_0^u da \int_0^v -\frac{1}{2}\mathbb{V}_x \Theta^{-1} e^{|a-j|\Theta} \mathbb{V}_x^T dj \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\left[\int_0^v dj \int_0^u e^{|a-j|\Theta} da\right] \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\left[\int_0^v dj \left(\int_0^j e^{|a-j|\Theta} da + \int_j^u e^{|a-j|\Theta} da\right)\right] \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\left\{\int_0^v \left[\int_0^j e^{(j-a)\Theta} da + \int_j^u e^{(a-j)\Theta} da\right] dj\right\} \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\left\{\int_0^v \left[-\Theta^{-1}\left(\mathbb{I}_q - e^{j\Theta}\right) + \Theta^{-1}\left(e^{(u-j)\Theta} - \mathbb{I}_q\right)\right] dj\right\} \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\Theta^{-1}\left[\int_0^v e^{j\Theta} dj - 2\int_0^v \mathbb{I}_q dj + \int_0^v e^{(u-j)\Theta} dj\right] \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-1}\Theta^{-1}\left[\Theta^{-1}\left(e^{v\Theta} - \mathbb{I}_q\right) - 2v\mathbb{I}_q - \Theta^{-1}\left(e^{(u-v)\Theta} - e^{u\Theta}\right)\right] \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \Theta^{-3}\left(e^{u\Theta} + e^{v\Theta} - e^{(u-v)\Theta} - 2v\Theta - \mathbb{I}_q\right) \mathbb{V}_x^T \\
& = -\frac{1}{2}\mathbb{V}_x \mathbb{V}_x^T \Theta^{-3}\left(e^{u\Theta} + e^{v\Theta} - e^{(u-v)\Theta} - 2v\Theta - \mathbb{I}_q\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_0^u Cov\left(\boldsymbol{x}_a, \boldsymbol{b}_v\right) da \\
&= \int_0^v Cov\left(\boldsymbol{x}_a, \boldsymbol{b}_v\right) da + \int_v^u Cov\left(\boldsymbol{x}_a, \boldsymbol{b}_v\right) da \\
&= \int_0^v -\Theta^{-1}\left(\mathbb{I}_q - e^{a\Theta}\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T da + \int_v^u -\Theta^{-1}\left(e^{(a-v)\Theta} - e^{a\Theta}\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T da \\
&= -\Theta^{-1}\left(\int_0^v \left(\mathbb{I}_q - e^{a\Theta}\right) da\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T - \Theta^{-1}\left(\int_v^u \left(e^{(a-v)\Theta} - e^{a\Theta}\right) da\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T \\
&= -\Theta^{-1}\left(v\mathbb{I}_q - \Theta^{-1}\left(e^{v\Theta} - \mathbb{I}_q\right)\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T - \Theta^{-2}\left(e^{(u-v)\Theta} - \mathbb{I}_q - e^{u\Theta} + e^{v\Theta}\right)\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T \\
&= \left(-v\Theta + e^{v\Theta} - \mathbb{I}_q - e^{(u-v)\Theta} + \mathbb{I}_q + e^{u\Theta} - e^{v\Theta}\right)\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T \\
&= \left(e^{u\Theta} - e^{(u-v)\Theta} - v\Theta\right)\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T,
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_0^v Cov\left(\boldsymbol{b}_u, \boldsymbol{x}_j\right) dj = \int_0^v Cov\left(\boldsymbol{x}_j, \boldsymbol{b}_u\right)^T dj = \int_0^v -\mathbb{V}_{\boldsymbol{bz}}\mathbb{V}_x^T\left(\mathbb{I}_q - e^{j\Theta}\right)\Theta^{-1} dj \\
&= -\mathbb{V}_{\boldsymbol{bz}}\mathbb{V}_x^T\left(\int_0^v \left(\mathbb{I}_q - e^{j\Theta}\right) dj\right)\Theta^{-1} = -\mathbb{V}_{\boldsymbol{bz}}\mathbb{V}_x^T\left(v\mathbb{I}_q - \Theta^{-1}\left(e^{v\Theta} - \mathbb{I}_q\right)\right)\Theta^{-1} \\
&= \mathbb{V}_{\boldsymbol{bz}}\mathbb{V}_x^T\Theta^{-2}\left(e^{v\Theta} - \mathbb{I}_q - v\Theta\right) = \left(\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\right)^T\left(e^{v\Theta} - v\Theta - \mathbb{I}_q\right),
\end{aligned}
$$

and

$$
Cov\left(\boldsymbol{b}_u, \boldsymbol{b}_v\right) = \min(u, v)\mathbb{I}_q.
$$

Therefore, for $u \geq v$,

$$
\begin{aligned}
Cov\left(\boldsymbol{y}_u, \boldsymbol{y}_v\right) = \mathbb{U}&\left[-\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}\left(e^{u\Theta} + e^{v\Theta} - e^{(u-v)\Theta} - 2v\Theta - \mathbb{I}_q\right)\right]\mathbb{U}^T \\
&+ \mathbb{U}\left(e^{u\Theta} - e^{(u-v)\Theta} - v\Theta\right)\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T \\
&+ \mathbb{V}_p\left(\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\right)^T\left(e^{v\Theta} - v\Theta - \mathbb{I}_q\right)\mathbb{U}^T + \mathbb{V}_p\min(u, v)\mathbb{V}_p^T.
\end{aligned}
$$

Let

$$\mathbb{K}_1(u,v) = -e^{(u-v)\Theta} + e^{u\Theta} - v\Theta, \quad \mathbb{K}_2(v) = e^{v\Theta} - v\Theta - \mathbb{I}_q,$$

$$\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \quad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \quad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3},$$

then $\mathbb{K}_1(u,v), \mathbb{K}_2(v)$ and $\mathbb{Q}_1$ are all symmetric, and $\mathbb{Q}_1\mathbb{K}_1(u,v) = \mathbb{K}_1(u,v)\mathbb{Q}_1$, and $\mathbb{Q}_1, \mathbb{Q}_2, \mathbb{Q}_3$ are all independent of time t.

Therefore, for $u \geq v$,

$$\begin{aligned}
Cov\left(\boldsymbol{y}_u, \boldsymbol{y}_v\right) &= \mathbb{U}\mathbb{Q}_1\left(\mathbb{K}_1(u,v) + \mathbb{K}_2(v)\right)\mathbb{U}^T + \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_2 + \mathbb{Q}_2^T\mathbb{K}_2(v)\mathbb{U}^T + v\mathbb{V}_p\mathbb{V}_p^T \\
&= \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_1\mathbb{U}^T + \mathbb{U}\mathbb{Q}_1\mathbb{K}_2(v)\mathbb{U}^T + \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_2 + \mathbb{Q}_2^T\mathbb{K}_2(v)\mathbb{U}^T + v\mathbb{V}_p\mathbb{V}_p^T \\
&= \mathbb{U}\mathbb{K}_1(u,v)\left(\mathbb{Q}_1\mathbb{U}^T + \mathbb{Q}_2\right) + \left(\mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T\right)\mathbb{K}_2(v)\mathbb{U}^T + v\mathbb{V}_p\mathbb{V}_p^T \\
&= \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_3^T + \mathbb{Q}_3\mathbb{K}_2(v)\mathbb{U}^T + v\mathbb{V}_p\mathbb{V}_p^T.
\end{aligned}$$

Similarly, for $u < v$,

$$\begin{aligned}
Cov\left(\boldsymbol{y}_u, \boldsymbol{y}_v\right) = \left(Cov\left(\boldsymbol{y}_v, \boldsymbol{y}_u\right)\right)^T &= \left(\mathbb{U}\mathbb{K}_1(v,u)\mathbb{Q}_3^T + \mathbb{Q}_3\mathbb{K}_2(u)\mathbb{U}^T + u\mathbb{V}_p\mathbb{V}_p^T\right)^T \\
&= \mathbb{Q}_3\mathbb{K}_1(v,u)\mathbb{U}^T + \mathbb{U}\mathbb{K}_2(u)\mathbb{Q}_3^T + u\mathbb{V}_p\mathbb{V}_p^T.
\end{aligned}$$

The Lemma is proved.

Lemma 3.11 *Let $\boldsymbol{m}_t^{(s)}$ and $\boldsymbol{m}_t^{(l)}$ be the vector of moving averages based on lookback period s and l $(l > s)$, then*

$$Cov\left(\boldsymbol{m}_t^{(s)}, \boldsymbol{m}_t^{(l)}\right) = \mathbb{J}\left(t; s, l\right) + \mathbb{Q}_4\left(s, l\right),$$

where

$$\mathbb{J}\left(t;s,l\right)=\frac{1}{s}\mathbb{U}\Theta^{-1}\left(e^{t\Theta}-e^{(t-s)\Theta}\right)\mathbb{Q}_3^T-t\mathbb{U}\Theta\mathbb{Q}_3^T$$
$$+\frac{1}{l}\mathbb{Q}_3\Theta^{-1}\left(e^{t\Theta}-e^{(t-l)\Theta}\right)\mathbb{U}^T-t\mathbb{Q}_3\Theta\mathbb{U}^T+t\mathbb{V}_p\mathbb{V}_p^T,$$

and

$$\mathbb{Q}_4\left(s,l\right)=\frac{1}{sl}\mathbb{U}\left\{\Theta^{-1}\left[s\mathbb{I}_q-\Theta^{-1}\left(e^{l\Theta}-e^{(l-s)\Theta}\right)\right]+\frac{1}{6}\left(3l^2s+s^3\right)\Theta-\frac{1}{2}s^2\mathbb{I}_q\right\}\mathbb{Q}_3^T$$
$$+\frac{1}{sl}\mathbb{Q}_3\left\{\Theta^{-1}\left[s\mathbb{I}_q+\Theta^{-1}\left(\mathbb{I}_q-e^{s\Theta}\right)\right]+\frac{1}{6}\left(3l^2s+s^3\right)\Theta-\frac{1}{2}\left(2ls-s^2\right)\mathbb{I}_q\right\}\mathbb{U}^T$$
$$-\frac{1}{sl}\left[\frac{1}{6}\left(3l^2s+s^3\right)\right]\mathbb{V}_p\mathbb{V}_p^T,$$

and

$$\mathbb{Q}_3=\mathbb{U}\mathbb{Q}_1+\mathbb{Q}_2^T,\qquad \mathbb{Q}_2=\Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T,\qquad \mathbb{Q}_1=-\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.$$

Proof:

Based on definition of $\boldsymbol{m}_t^{(s)}$ and $\boldsymbol{m}_t^{(l)}$, and based on Lemma 3.10,

$$Cov\left(\boldsymbol{m}_t^{(s)},\boldsymbol{m}_t^{(l)}\right)$$
$$=Cov\left(\frac{1}{s}\int_{t-s}^{t}\boldsymbol{y}_u du,\frac{1}{l}\int_{t-l}^{t}\boldsymbol{y}_v dv\right)=\frac{1}{sl}\int_{t-s}^{t}du\int_{t-l}^{t}Cov\left(\boldsymbol{y}_u,\boldsymbol{y}_v\right)dv$$
$$=\frac{1}{sl}\int_{t-s}^{t}du\left(\int_{t-l}^{u}Cov\left(\boldsymbol{y}_u,\boldsymbol{y}_v\right)dv+\int_{u}^{t}Cov\left(\boldsymbol{y}_u,\boldsymbol{y}_v\right)dv\right)$$
$$=\frac{1}{sl}\int_{t-s}^{t}du\bigg\{\int_{t-l}^{u}\left[\mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_3^T+\mathbb{Q}_3\mathbb{K}_2(v)\mathbb{U}^T+v\mathbb{V}_p\mathbb{V}_p^T\right]dv$$
$$+\int_{u}^{t}\left[\mathbb{Q}_3\mathbb{K}_1(v,u)\mathbb{U}^T+\mathbb{U}\mathbb{K}_2(u)\mathbb{Q}_3^T+u\mathbb{V}_p\mathbb{V}_p^T\right]dv\bigg\}$$
$$=\frac{1}{sl}\int_{t-s}^{t}du\bigg\{\int_{t-l}^{u}\mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_3^Tdv+\int_{t-l}^{u}\mathbb{Q}_3\mathbb{K}_2(v)\mathbb{U}^Tdv+\int_{t-l}^{u}v\mathbb{V}_p\mathbb{V}_p^Tdv$$
$$+\int_{u}^{t}\mathbb{Q}_3\mathbb{K}_1(v,u)\mathbb{U}^Tdv+\int_{u}^{t}\mathbb{U}\mathbb{K}_2(u)\mathbb{Q}_3^Tdv+\int_{u}^{t}u\mathbb{V}_p\mathbb{V}_p^Tdv\bigg\},$$

where

$$(1). \int_{t-l}^{u} \mathbb{U}\mathbb{K}_1(u,v)\mathbb{Q}_3^T dv = \int_{t-l}^{u} \mathbb{U}\left(-e^{(u-v)\Theta} + e^{u\Theta} - v\Theta\right)\mathbb{Q}_3^T dv$$

$$= \mathbb{U}\left\{\Theta^{-1}\left(\mathbb{I}_q - e^{(u-t+l)\Theta}\right) + e^{u\Theta}\left(u-t+l\right) - \frac{1}{2}\left(u^2-(t-l)^2\right)\Theta\right\}\mathbb{Q}_3^T,$$

$$(2). \int_{t-l}^{u} \mathbb{Q}_3\mathbb{K}_2(v)\mathbb{U}^T dv = \int_{t-l}^{u} \mathbb{Q}_3\left(e^{v\Theta} - v\Theta - \mathbb{I}_q\right)\mathbb{U}^T dv$$

$$= \mathbb{Q}_3\left\{\Theta^{-1}\left(e^{u\Theta} - e^{(t-l)\Theta}\right) - \frac{1}{2}\left(u^2-(t-l)^2\right)\Theta - \left(u-t+l\right)\mathbb{I}_q\right\}\mathbb{U}^T,$$

$$(3). \int_{t-l}^{u} v\mathbb{V}_p\mathbb{V}_p^T dv = \frac{1}{2}\left(u^2-(t-l)^2\right)\mathbb{V}_p\mathbb{V}_p^T,$$

$$(4). \int_{u}^{t} \mathbb{Q}_3\mathbb{K}_1(v,u)\mathbb{U}^T dv = \mathbb{Q}_3\int_{u}^{t}\left(-e^{(v-u)\Theta} + e^{v\Theta} - u\Theta\right)dv\mathbb{U}^T$$

$$= \mathbb{Q}_3\left(-\Theta^{-1}\left(e^{(t-u)\Theta} - \mathbb{I}_q\right) + \Theta^{-1}\left(e^{t\Theta} - e^{u\Theta}\right) - u(t-u)\Theta\right)\mathbb{U}^T,$$

$$(5). \int_{u}^{t} \mathbb{U}\mathbb{K}_2(u)\mathbb{Q}_3^T dv = \mathbb{U}\int_{u}^{t}\left(e^{u\Theta} - u\Theta - \mathbb{I}_q\right)dv\mathbb{Q}_3^T = \mathbb{U}\left[(t-u)\left(e^{u\Theta} - u\Theta - \mathbb{I}_q\right)\right]\mathbb{Q}_3^T,$$

$$(6). \int_{u}^{t} u\mathbb{V}_p\mathbb{V}_p^T dv = u(t-u)\mathbb{V}_p\mathbb{V}_p^T.$$

Therefore

$$(1)+(5) = \mathbb{U}\left\{\Theta^{-1}\left(\mathbb{I}_q - e^{(u-t+l)\Theta}\right) + e^{u\Theta}l + \left(\frac{1}{2}(t-u-l)^2 - ul\right)\Theta - (t-u)\mathbb{I}_q\right\}\mathbb{Q}_3^T,$$

$$(2)+(4)$$

$$= \mathbb{Q}_3\left\{\Theta^{-1}\left(e^{t\Theta} + \mathbb{I}_q - e^{(t-u)\Theta} - e^{(t-l)\Theta}\right) + \left(\frac{1}{2}(t-u-l)^2 - ul\right)\Theta - (u-t+l)\mathbb{I}_q\right\}\mathbb{U}^T,$$

$$(3)+(6) = -\left(\frac{1}{2}(t-u-l)^2 - ul\right)\mathbb{V}_p\mathbb{V}_p^T,$$

which implies

$$\begin{aligned}
\int_{t-s}^{t} [(1)+(5)]\,du &= \mathbb{U}\bigg\{\int_{t-s}^{t} \Theta^{-1}\left(\mathbb{I}_q - e^{(u-t+l)\Theta}\right)du + \int_{t-s}^{t} e^{u\Theta}l\,du \\
&\quad + \int_{t-s}^{t}\left(\frac{1}{2}(t-u-l)^2 - ul\right)\Theta du - \int_{t-s}^{t}(t-u)\mathbb{I}_q du\bigg\}\mathbb{Q}_3^T \\
&= \mathbb{U}\bigg\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \Theta^{-1}\left(e^{t\Theta} - e^{(t-s)\Theta}\right)l \\
&\quad + \left(\frac{1}{6}\left(l^3 - (l-s)^3\right) + \frac{1}{2}s^2 l - tsl\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\bigg\}\mathbb{Q}_3^T,
\end{aligned}$$

and

$$\begin{aligned}
\int_{t-s}^{t} [(2)+(4)]\,du &= \mathbb{Q}_3\bigg\{\int_{t-s}^{t} \Theta^{-1}\left(e^{t\Theta} + \mathbb{I}_q - e^{(t-u)\Theta} - e^{(t-l)\Theta}\right)du \\
&\quad + \int_{t-s}^{t}\left(\frac{1}{2}(t-u-l)^2 - ul\right)\Theta du - \int_{t-s}^{t}(u-t+l)\mathbb{I}_q du\bigg\}\mathbb{U}^T \\
&= \mathbb{Q}_3\bigg\{\Theta^{-1}\left(se^{t\Theta} + s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right) - se^{(t-l)\Theta}\right) \\
&\quad + \left(\frac{1}{6}\left(l^3 - (l-s)^3\right) + \frac{1}{2}s^2 l - tsl\right)\Theta - \frac{1}{2}\left(l^2 - (l-s)^2\right)\mathbb{I}_q\bigg\}\mathbb{U}^T,
\end{aligned}$$

and

$$\begin{aligned}
\int_{t-s}^{t} [(3)+(6)]\,du &= \int_{t-s}^{t} -\left(\frac{1}{2}(t-u-l)^2 - ul\right)\mathbb{V}_p\mathbb{V}_p^T du \\
&= -\left(\frac{1}{6}\left(l^3 - (l-s)^3\right) + \frac{1}{2}s^2 l - tsl\right)\mathbb{V}_p\mathbb{V}_p^T.
\end{aligned}$$

Therefore

$$
\begin{aligned}
&Cov\left(\boldsymbol{m}_t^{(s)}, \boldsymbol{m}_t^{(l)}\right) \\
&= \frac{1}{sl}\mathbb{U}\Big\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \Theta^{-1}\left(e^{t\Theta} - e^{(t-s)\Theta}\right) l \\
&+ \left(\frac{1}{6}\left(l^3 - (l-s)^3\right) + \frac{1}{2}s^2 l - tsl\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\Big\}\mathbb{Q}_3^T \\
&+ \frac{1}{sl}\mathbb{Q}_3\Big\{\Theta^{-1}\left(se^{t\Theta} + s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right) - se^{(t-l)\Theta}\right) \\
&+ \left(\frac{1}{6}\left(l^3 - (l-s)^3\right) + \frac{1}{2}s^2 l - tsl\right)\Theta - \frac{1}{2}\left(l^2 - (l-s)^2\right)\mathbb{I}_q\Big\}\mathbb{U}^T \\
&+ \frac{1}{sl}\left(-\frac{1}{6}\left(l^3 - (l-s)^3\right) - \frac{1}{2}s^2 l + tsl\right)\mathbb{V}_p\mathbb{V}_p^T \\
&= \frac{1}{s}\mathbb{U}\Theta^{-1}\left(e^{t\Theta} - e^{(t-s)\Theta}\right)\mathbb{Q}_3^T - t\mathbb{U}\Theta\mathbb{Q}_3^T + \frac{1}{l}\mathbb{Q}_3\Theta^{-1}\left(e^{t\Theta} - e^{(t-l)\Theta}\right)\mathbb{U}^T - t\mathbb{Q}_3\Theta\mathbb{U}^T \\
&+ t\mathbb{V}_p\mathbb{V}_p^T + \frac{1}{sl}\mathbb{U}\left\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\right\}\mathbb{Q}_3^T \\
&+ \frac{1}{sl}\mathbb{Q}_3\left\{\Theta^{-1}\left[s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}\left(2ls - s^2\right)\mathbb{I}_q\right\}\mathbb{U}^T \\
&- \frac{1}{sl}\left[\frac{1}{6}\left(3l^2 s + s^3\right)\right]\mathbb{V}_p\mathbb{V}_p^T \\
&= \mathbb{J}\left(t; s, l\right) + \mathbb{Q}_4\left(s, l\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{J}\left(t; s, l\right) = &\frac{1}{s}\mathbb{U}\Theta^{-1}\left(e^{t\Theta} - e^{(t-s)\Theta}\right)\mathbb{Q}_3^T - t\mathbb{U}\Theta\mathbb{Q}_3^T \\
&+ \frac{1}{l}\mathbb{Q}_3\Theta^{-1}\left(e^{t\Theta} - e^{(t-l)\Theta}\right)\mathbb{U}^T - t\mathbb{Q}_3\Theta\mathbb{U}^T + t\mathbb{V}_p\mathbb{V}_p^T,
\end{aligned}
$$

and

$$\mathbb{Q}_4(s,l) = \frac{1}{sl}\mathbb{U}\left\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\right\}\mathbb{Q}_3^T$$
$$+ \frac{1}{sl}\mathbb{Q}_3\left\{\Theta^{-1}\left[s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}\left(2ls - s^2\right)\mathbb{I}_q\right\}\mathbb{U}^T$$
$$- \frac{1}{sl}\left[\frac{1}{6}\left(3l^2 s + s^3\right)\right]\mathbb{V}_p\mathbb{V}_p^T,$$

and

$$\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \quad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \quad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.$$

The Lemma is proved. Notice that $\mathbb{Q}_1$ is symmetric, $\mathbb{Q}_1, \mathbb{Q}_2$ and $\mathbb{Q}_3$ are independent of s, l and t, $\mathbb{Q}_4$ is independent of t but is dependent of s and l, and $\mathbb{J}$ is dependent of s, l and t.

Lemma 3.12 *Let $\boldsymbol{m}_t^{(s,l)}$ be the vector of moving average differences based on lookback period s and l $(l > s)$, then $Var\left(\boldsymbol{m}_t^{(s,l)}\right)$ is independent of time t, i.e.,*

$$Var\left(\boldsymbol{m}_t^{(s,l)}\right) = \mathbb{Q}_4(s,s) - \mathbb{Q}_4(s,l) - \mathbb{Q}_4^T(s,l) + \mathbb{Q}_4(l,l),$$

where

$$\mathbb{Q}_4(s,l) = \frac{1}{sl}\mathbb{U}\left\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\right\}\mathbb{Q}_3^T$$
$$+ \frac{1}{sl}\mathbb{Q}_3\left\{\Theta^{-1}\left[s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right)\right] + \frac{1}{6}\left(3l^2 s + s^3\right)\Theta - \frac{1}{2}\left(2ls - s^2\right)\mathbb{I}_q\right\}\mathbb{U}^T$$
$$- \frac{1}{sl}\left[\frac{1}{6}\left(3l^2 s + s^3\right)\right]\mathbb{V}_p\mathbb{V}_p^T,$$

and

$$\mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T, \quad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \quad \mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}.$$

Proof:

Based on definition of $\boldsymbol{m}_t^{(s,l)}$ and Lemma 3.11,

$$
\begin{aligned}
& Var\left(\boldsymbol{m}_t^{(s,l)}\right) \\
& = Var\left(\boldsymbol{m}_t^{(s)} - \boldsymbol{m}_t^{(l)}\right) \\
& = Cov\left(\boldsymbol{m}_t^{(s)}, \boldsymbol{m}_t^{(s)}\right) - Cov\left(\boldsymbol{m}_t^{(s)}, \boldsymbol{m}_t^{(l)}\right) - Cov\left(\boldsymbol{m}_t^{(l)}, \boldsymbol{m}_t^{(s)}\right) + Cov\left(\boldsymbol{m}_t^{(l)}, \boldsymbol{m}_t^{(l)}\right) \\
& = \left[\mathbb{J}(t; s, s) - \mathbb{J}(t; s, l) - \mathbb{J}^T(t; s, l) + \mathbb{J}(t; l, l)\right] \\
& + \left[\mathbb{Q}_4(s, s) - \mathbb{Q}_4(s, l) - \mathbb{Q}_4^T(s, l) + \mathbb{Q}_4(l, l)\right] \\
& = \mathbb{Q}_4(s, s) - \mathbb{Q}_4(s, l) - \mathbb{Q}_4^T(s, l) + \mathbb{Q}_4(l, l),
\end{aligned}
$$

as

$$\mathbb{J}(t; s, s) - \mathbb{J}(t; s, l) - \mathbb{J}^T(t; s, l) + \mathbb{J}(t; l, l) = 0.$$

The Lemma is proved.

Lemma 3.13 *Let $\boldsymbol{b}_t^{(0)}$ be n-dimensional standard Brownian motion and $\boldsymbol{z}_t^{(0)}$ be q-dimensional standard Brownian motion. $\boldsymbol{b}_t^{(0)}$ and $\boldsymbol{z}_t^{(0)}$ are independent. If there is a symmetric matrix Γ such that $\Gamma\Gamma^T = \begin{pmatrix} \mathbb{I}_n & \mathbb{V}_{\boldsymbol{bz}} \\ \mathbb{V}_{\boldsymbol{bz}}^T & \mathbb{I}_q \end{pmatrix}$, where $\mathbb{I}_n$ is n-dimensional identity matrix and $\mathbb{I}_q$ is q-dimensional identity matrix, then $\boldsymbol{b}_t$ and $\boldsymbol{z}_t$, i.e., $(\boldsymbol{b}_t\ \boldsymbol{z}_t)^T = \Gamma\ \left(\boldsymbol{b}_t^{(0)}\ \boldsymbol{z}_t^{(0)}\right)^T$ are multi-dimensional standard Brownian motion with correlation matrix $\mathbb{V}_{\boldsymbol{bz}}$.*

Proof:

Since $\boldsymbol{b}_t^{(0)}$ and $\boldsymbol{z}_t^{(0)}$ are independent standard Brownian motion with dimensions n and q respective, $Var\begin{pmatrix}\boldsymbol{b}_t^{(0)}\\ \boldsymbol{z}_t^{(0)}\end{pmatrix} = t\mathbb{I}_{(n+q)}$. Let $\begin{pmatrix}\boldsymbol{b}_t\\ \boldsymbol{z}_t\end{pmatrix} = \Gamma\begin{pmatrix}\boldsymbol{b}_t^{(0)}\\ \boldsymbol{z}_t^{(0)}\end{pmatrix}$, then

$$Var\begin{pmatrix}\boldsymbol{b}_t\\ \boldsymbol{z}_t\end{pmatrix} = \Gamma Var\begin{pmatrix}\boldsymbol{b}_t^{(0)}\\ \boldsymbol{z}_t^{(0)}\end{pmatrix}\Gamma^T = \Gamma t\mathbb{I}_{(n+q)}\Gamma^T = t\Gamma\Gamma^T = t\begin{pmatrix}\mathbb{I}_n & \mathbb{V}_{\boldsymbol{bz}}\\ \mathbb{V}_{\boldsymbol{bz}}^T & \mathbb{I}_q,\end{pmatrix}$$

which implies

$$Var\left(\boldsymbol{b}_t\right) = t\mathbb{I}_n, \quad Var\left(\boldsymbol{z}_t\right) = t\mathbb{I}_q, \quad Cov\left(\boldsymbol{b}_t, \boldsymbol{z}_t\right) = t\mathbb{V}_{\boldsymbol{bz}},$$

then the correlation matrix for $\boldsymbol{b}_t$ and $\boldsymbol{z}_t$ is $\mathbb{V}_{\boldsymbol{bz}}$. In addition,

$$Var\begin{pmatrix}\boldsymbol{b}_{t+dt} - \boldsymbol{b}_t\\ \boldsymbol{z}_{t+dt} - \boldsymbol{z}_t\end{pmatrix} = Var\left[\Gamma\begin{pmatrix}\boldsymbol{b}_{t+dt}^{(0)} - \boldsymbol{b}_t^{(0)}\\ \boldsymbol{z}_{t+dt}^{(0)} - \boldsymbol{z}_t^{(0)}\end{pmatrix}\right] = \Gamma Var\left[\begin{pmatrix}\boldsymbol{b}_{t+dt}^{(0)} - \boldsymbol{b}_t^{(0)}\\ \boldsymbol{z}_{t+dt}^{(0)} - \boldsymbol{z}_t^{(0)}\end{pmatrix}\right]\Gamma^T$$

$$= \Gamma dt\mathbb{I}_{(n+q)}\Gamma^T = dt\Gamma\Gamma^T = dt\begin{pmatrix}\mathbb{I}_n & \mathbb{V}_{\boldsymbol{bz}}\\ \mathbb{V}_{\boldsymbol{bz}}^T & \mathbb{I}_q,\end{pmatrix}$$

which implies

$$Var\left(\boldsymbol{b}_{t+dt} - \boldsymbol{b}_t\right) = dt\mathbb{I}_n, \quad Var\left(\boldsymbol{z}_{t+dt} - \boldsymbol{z}_t\right) = dt\mathbb{I}_q,$$

then $\boldsymbol{b}_t$ is standard Brownian motion and $\boldsymbol{z}_t$ is standard Brownian motion. The Lemma is proved.

3.3 The Analytic Results

In order to find optimal $\boldsymbol{\eta}_t$, we need derive the investor's expected log-utility of wealth $E(\log w_T)$. To derive $E(\log w_T)$, we need know joint distribution of $\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right)^T$. Let $\boldsymbol{\gamma}$ be a column vector of diagonal of matrix $\mathbb{V}_p\mathbb{V}_p^T$, i.e.,

$$\boldsymbol{\gamma} = \begin{pmatrix} \sum_{i=1}^{n}\left(v_{1i}^{p}\right)^2 \\ \vdots \\ \sum_{i=1}^{n}\left(v_{ni}^{p}\right)^2 \end{pmatrix}.$$

Let $\boldsymbol{\mu_x}$ be the vector of expectation of $\boldsymbol{x}_t$, $\boldsymbol{\mu_m}$ be the vector of expectation of $\boldsymbol{m}_t^{(s,l)}$, $\Sigma_{\boldsymbol{x}}$ be the variance-covariance matrix of $\boldsymbol{x}_t$, $\Sigma_{\boldsymbol{m}}$ be the variance-covariance of $\boldsymbol{m}_t^{(s,l)}$, and $\Delta_{\boldsymbol{xm}}$ be the covariance matrix between $\boldsymbol{x}_t$ and $\boldsymbol{m}_t^{(s,l)}$. Based on Lemmas 3.2, 3.4, 3.9 and 3.12, it is derived that $\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right)^T$ are multi-normal distribution, i.e.,

$$\begin{pmatrix} \boldsymbol{x}_t \\ \boldsymbol{m}_t^{(s,l)} \end{pmatrix} \sim \mathrm{MN}\left[\begin{pmatrix} \boldsymbol{\mu_x} \\ \boldsymbol{\mu_m} \end{pmatrix}, \begin{pmatrix} \Sigma_{\boldsymbol{x}} & \Delta_{\boldsymbol{xm}} \\ \Delta_{\boldsymbol{xm}}^T & \Sigma_{\boldsymbol{m}} \end{pmatrix}\right], \tag{3.14}$$

and

$$
\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{x}} &= -\Theta^{-1}\boldsymbol{\beta}, \\
\boldsymbol{\mu}_{\boldsymbol{m}} &= \frac{1}{2}(l-s)\left[\boldsymbol{\alpha} - \mathbb{U}\Theta^{-1}\boldsymbol{\beta} - \frac{1}{2}\boldsymbol{\gamma}\right], \\
\Sigma_{\boldsymbol{x}} &= -\frac{1}{2}\mathbb{V}_x\Theta^{-1}\mathbb{V}_x^T, \\
\Delta_{\boldsymbol{xm}} &= \left[\frac{1}{s}\left(\mathbb{I}_q - e^{s\Theta}\right) - \frac{1}{l}\left(\mathbb{I}_q - e^{l\Theta}\right)\right]\mathbb{Q}_3^T, \\
\Sigma_{\boldsymbol{m}} &= \mathbb{Q}_4(s,s) - \mathbb{Q}_4(s,l) - \mathbb{Q}_4^T(s,l) + \mathbb{Q}_4(l,l), \qquad (3.15)
\end{aligned}
$$

where

$$
\mathbb{Q}_1 = -\frac{1}{2}\mathbb{V}_x\mathbb{V}_x^T\Theta^{-3}, \qquad \mathbb{Q}_2 = \Theta^{-2}\mathbb{V}_x\mathbb{V}_{\boldsymbol{bz}}^T\mathbb{V}_p^T, \qquad \mathbb{Q}_3 = \mathbb{U}\mathbb{Q}_1 + \mathbb{Q}_2^T,
$$

and

$$
\begin{aligned}
\mathbb{Q}_4(s,l) = {} & \frac{1}{sl}\mathbb{U}\left\{\Theta^{-1}\left[s\mathbb{I}_q - \Theta^{-1}\left(e^{l\Theta} - e^{(l-s)\Theta}\right)\right] + \frac{1}{6}\left(3l^2s + s^3\right)\Theta - \frac{1}{2}s^2\mathbb{I}_q\right\}\mathbb{Q}_3^T \\
& + \frac{1}{sl}\mathbb{Q}_3\left\{\Theta^{-1}\left[s\mathbb{I}_q + \Theta^{-1}\left(\mathbb{I}_q - e^{s\Theta}\right)\right] + \frac{1}{6}\left(3l^2s + s^3\right)\Theta - \frac{1}{2}\left(2ls - s^2\right)\mathbb{I}_q\right\}\mathbb{U}^T \\
& - \frac{1}{sl}\left[\frac{1}{6}\left(3l^2s + s^3\right)\right]\mathbb{V}_p\mathbb{V}_p^T.
\end{aligned}
$$

Notice that the multi-normal distribution of $\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right)^T$ is independent of time t. From the multi-normal distribution of $\left(\boldsymbol{x}_t, \boldsymbol{m}_t^{(s,l)}\right)^T$, we have

$$
E\left(\boldsymbol{x}_t \mid \boldsymbol{m}_t^{(s,l)}\right) = \boldsymbol{\mu}_{\boldsymbol{x}} + \Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1}\left(\boldsymbol{m}_t^{(s,l)} - \boldsymbol{\mu}_{\boldsymbol{m}}\right). \qquad (3.16)
$$

We define $\Sigma_{\boldsymbol{m}}$, $\boldsymbol{\mu}_{\boldsymbol{m}}$ and $\boldsymbol{\sigma}_{\boldsymbol{m}}$ as

$$\Sigma_{\boldsymbol{m}} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \ldots & \sigma_n^2 \end{pmatrix}, \quad \boldsymbol{\mu}_{\boldsymbol{m}} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \boldsymbol{\sigma}_{\boldsymbol{m}} = \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix},$$

then

$$\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} = \begin{pmatrix} \sigma_1^{-1} & 0 & \ldots & 0 \\ 0 & \sigma_2^{-1} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_n^{-1} \end{pmatrix} \left(\boldsymbol{m}_t^{(s,l)} - \boldsymbol{\mu}_{\boldsymbol{m}}\right) \sim \mathrm{MN}\left(\mathbb{0}_n, \mathbb{R}_{\boldsymbol{m}}\right), \tag{3.17}$$

where $\mathbb{0}_n^T = (0, \ldots, 0)$ and $\mathbb{R}_{\boldsymbol{m}}$ is the correlation matrix for the vector $\boldsymbol{m}_t^{(s,l)}$, i.e.,

$$\mathbb{R}_{\boldsymbol{m}} = \begin{pmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} & \ldots & \frac{\sigma_{1n}}{\sigma_1\sigma_n} \\ \frac{\sigma_{12}}{\sigma_1\sigma_2} & 1 & \ldots & \frac{\sigma_{2n}}{\sigma_2\sigma_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1n}}{\sigma_1\sigma_n} & \frac{\sigma_{2n}}{\sigma_2\sigma_n} & \ldots & 1 \end{pmatrix}.$$

Suppose that $\phi_{\mathbb{R}_{\boldsymbol{m}}}(m_1, \ldots, m_n)$ is a probability density function for $\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} \sim \mathrm{MN}\left(\mathbb{0}_n, \mathbb{R}_{\boldsymbol{m}}\right)$. For any hyper-rectangle $H = [a_1, b_1] \times \ldots \times [a_n, b_n]$ (it doesn't matter whether the boundary is open or close for one-dimension of the hyper-rectangle), we define

$$\Phi_{\mathbb{R}_{\boldsymbol{m}}}(H) = Pr\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} \in H\right) = \int_{a_1}^{b_1} dm_1 \int_{a_2}^{b_2} dm_2 \ldots \int_{a_n}^{b_n} \phi_{\mathbb{R}_{\boldsymbol{m}}}(m_1, \ldots, m_n)\, dm_n, \tag{3.18}$$

and

$$\boldsymbol{\Psi}_{\mathbb{R}_{\boldsymbol{m}}}(H) = \int_{a_1}^{b_1} dm_1 \int_{a_2}^{b_2} dm_2 \ldots \int_{a_n}^{b_n} \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \phi_{\mathbb{R}_{\boldsymbol{m}}}(m_1, \ldots, m_n)\, dm_n. \tag{3.19}$$

Based on Equation (3.7), we define $\Omega_i(\mu, \sigma)$ as

$$\Omega_i(\mu, \sigma) = \frac{\Omega_i - \mu}{\sigma} = \begin{cases} \left(-\infty, -\frac{\epsilon+\mu}{\sigma}\right), & \text{if } i = 1, \\ \left[-\frac{\epsilon+\mu}{\sigma}, -\frac{\mu}{\sigma}\right), & \text{if } i = 2, \\ \left[-\frac{\mu}{\sigma}, \frac{\epsilon-\mu}{\sigma}\right], & \text{if } i = 3, \\ \left(\frac{\epsilon-\mu}{\sigma}, \infty\right), & \text{if } i = 4. \end{cases} \tag{3.20}$$

Let $d = \{1, 2, 3, 4\}$ and we define

$$\Omega_{(i_1,\ldots,i_n)}(\boldsymbol{\mu_m}, \boldsymbol{\sigma_m}) = \Omega_{i_1}(\mu_1, \sigma_1) \times \ldots \times \Omega_{i_n}(\mu_n, \sigma_n), \tag{3.21}$$

where $i_k \in d$, $k = 1, \ldots, n$.

Given an initial wealth w_0, a constant rate of interest r and an investment horizon T. Let $\epsilon > 0$ be the investor specified risk tolerance. Let $\boldsymbol{\delta}_{(i_1,\ldots,i_n)}$ be the vector of n asset allocation parameters. Let $\boldsymbol{\eta}_t$ be the vector based multi-asset generalized moving average crossover (MGMA) strategy. We state the following Propositions for the MGMA strategy.

Proposition 3.1 *The vector based expected values of* $\boldsymbol{\eta}_t^T$ *are independent of time* t, *i.e.,*

$$E\left(\boldsymbol{\eta}_t^T\right) = \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right).$$

Proof.

Base on Equations (3.9), (3.17), (3.20) and (3.21),

$$\begin{aligned} E\left(\boldsymbol{\eta}_t^T\right) &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T P\left(\boldsymbol{m}_t^{(s,l)} \in \Omega_{(i_1,\ldots,i_n)}\right) \\ &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T P\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} \in \Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right) \\ &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right). \end{aligned}$$

The Proposition is proved.

Proposition 3.2 *The expected value of* $\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t$ *is independent of time* t, *i.e.,*

$$E\left(\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right) = \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\delta}_{(i_1,\ldots,i_n)} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right).$$

Proof:

Base on Equations (3.9), (3.17), (3.20) and (3.21),

$$\begin{aligned} E\left(\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right) &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\delta}_{(i_1,\ldots,i_n)} P\left(\boldsymbol{m}_t^{(s,l)} \in \Omega_{(i_1,\ldots,i_n)}\right) \\ &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\delta}_{(i_1,\ldots,i_n)} P\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} \in \Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right) \\ &= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\delta}_{(i_1,\ldots,i_n)} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right). \end{aligned}$$

The Proposition is proved.

Proposition 3.3 *The expected value of $\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t$ is independent of time t, i.e.,*

$$E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t\right) = \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U} \boldsymbol{\mu}_{\boldsymbol{x}} \Phi_{\mathbb{R}_{\boldsymbol{m}}} \left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}}, \boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right)$$

$$+ \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U} \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_n \end{pmatrix} \boldsymbol{\Psi}_{\mathbb{R}_{\boldsymbol{m}}} \left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}}, \boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right).$$

Proof:

Based on Equation (3.16) and Proposition 3.1,

$$\begin{aligned} E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t\right) &= E\left\{E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t \mid \boldsymbol{m}_t^{(s,l)}\right)\right\} = E\left\{\boldsymbol{\eta}_t^T \mathbb{U} E\left(\boldsymbol{x}_t \mid \boldsymbol{m}_t^{(s,l)}\right)\right\} \\ &= E\left\{\boldsymbol{\eta}_t^T \mathbb{U}\left[\boldsymbol{\mu}_{\boldsymbol{x}} + \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1}\left(\boldsymbol{m}_t^{(s,l)} - \boldsymbol{\mu}_{\boldsymbol{m}}\right)\right]\right\} \\ &= E\left(\boldsymbol{\eta}_t^T\right) \mathbb{U}\left(\boldsymbol{\mu}_{\boldsymbol{x}} - \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{\mu}_{\boldsymbol{m}}\right) + E\left(\boldsymbol{\eta}_t^T \mathbb{U} \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{m}_t^{(s,l)}\right), \end{aligned}$$

where

$$\begin{aligned} & E\left(\boldsymbol{\eta}_t^T\right) \mathbb{U}\left(\boldsymbol{\mu}_{\boldsymbol{x}} - \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{\mu}_{\boldsymbol{m}}\right) \\ &= \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U}\left(\boldsymbol{\mu}_{\boldsymbol{x}} - \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{\mu}_{\boldsymbol{m}}\right) \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}}, \boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right) \\ &= \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U} \boldsymbol{\mu}_{\boldsymbol{x}} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}}, \boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right) \\ &- \sum_{i_1 \in d, \ldots, i_n \in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U} \Delta_{\boldsymbol{x}\boldsymbol{m}} \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{\mu}_{\boldsymbol{m}} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}}, \boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right), \end{aligned}$$

and based on Equations (3.9) and (3.17),

$$E\left(\boldsymbol{\eta}_t^T \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1}\boldsymbol{m}_t^{(s,l)}\right) = \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1} E\left(\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}}\left(\boldsymbol{m}_t^{(s,l)}\right)\boldsymbol{m}_t^{(s,l)}\right),$$

where

$$\begin{aligned}
&E\left(\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}}\left(\boldsymbol{m}_t^{(s,l)}\right)\boldsymbol{m}_t^{(s,l)}\right)\\
&= E\left(\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}(\boldsymbol{\mu_m},\boldsymbol{\sigma_m})}\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}}\right)\left(\begin{pmatrix}\sigma_1 & 0 & \ldots & 0\\ 0 & \sigma_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \sigma_n\end{pmatrix}\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}} + \boldsymbol{\mu_m}\right)\right)\\
&= \begin{pmatrix}\sigma_1 & 0 & \ldots & 0\\ 0 & \sigma_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \sigma_n\end{pmatrix} E\left(\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}(\boldsymbol{\mu_m},\boldsymbol{\sigma_m})}\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}}\right)\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}}\right)\\
&+ E\left(\mathbf{1}_{\Omega_{(i_1,\ldots,i_n)}(\boldsymbol{\mu_m},\boldsymbol{\sigma_m})}\left(\boldsymbol{Z}_{\mathbb{R}_{\boldsymbol{m}}}\right)\boldsymbol{\mu_m}\right)\\
&= \begin{pmatrix}\sigma_1 & 0 & \ldots & 0\\ 0 & \sigma_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \sigma_n\end{pmatrix} \boldsymbol{\Psi}_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right) + \boldsymbol{\mu_m}\Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right),
\end{aligned}$$

which implies

$$E\left(\boldsymbol{\eta}_t^T \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1}\boldsymbol{m}_t^{(s,l)}\right)$$

$$= \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1} \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_n \end{pmatrix} \boldsymbol{\Psi}_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}},\boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right)$$

$$+ \sum_{i_1\in d,\ldots,i_n\in d} \boldsymbol{\delta}_{(i_1,\ldots,i_n)}^T \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma_{\boldsymbol{m}}^{-1}\boldsymbol{\mu}_{\boldsymbol{m}}\Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu}_{\boldsymbol{m}},\boldsymbol{\sigma}_{\boldsymbol{m}}\right)\right).$$

The Proposition is proved.

Proposition 3.4 *Let $\boldsymbol{\lambda}$ be a constant vector for MGMA strategy $\boldsymbol{\eta}_t$ when $t < l$, i.e., $\boldsymbol{\lambda}^T = (\lambda_1,\ldots,\lambda_n)$ where λ_k is a constant for $k = 1,\ldots,n$ and $\sum_{k=1}^n \lambda_k \leq 1$. Let $\epsilon > 0$ be the investor specified risk tolerance, then the investor's expected log-utility of wealth at the end of investment period T is*

$$E(\log w_T) = a_6 + (T-l)\left[E\left(\boldsymbol{\eta}_t^T\right)(\boldsymbol{\alpha} - r\mathbb{1}_n) - \frac{1}{2}E\left(\boldsymbol{\eta}_t^T\mathbb{V}_p\mathbb{V}_p^T\boldsymbol{\eta}_t\right) + E\left(\boldsymbol{\eta}_t^T\mathbb{U}\boldsymbol{x}_t\right)\right], \tag{3.22}$$

where $\mathbb{1}_n^T = (1,\ldots,1)$ and a_6 is a constant depending on l, i.e.,

$$a_6 = \log w_0 + rT + l\left[\boldsymbol{\lambda}^T(\boldsymbol{\alpha} - r\mathbb{1}_n) - \frac{1}{2}\left(\boldsymbol{\lambda}^T\mathbb{V}_p\mathbb{V}_p^T\boldsymbol{\lambda}\right) - \boldsymbol{\lambda}^T\mathbb{U}\Theta^{-1}\boldsymbol{\beta}\right].$$

By Propositions 3.1, 3.2 *and* 3.3, *Equation* (3.22) *can be rewritten as*

$$
\begin{aligned}
&E\left(\log w_T\right) \\
&= a_6 + \sum_{i_1\in d,\ldots,i_n\in d} (T-l)\,\boldsymbol{\delta}^T_{(i_1,\ldots,i_n)} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right)\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) \\
&- \sum_{i_1\in d,\ldots,i_n\in d} \frac{1}{2}(T-l)\,\boldsymbol{\delta}^T_{(i_1,\ldots,i_n)} \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\delta}_{(i_1,\ldots,i_n)} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right) \\
&+ \sum_{i_1\in d,\ldots,i_n\in d} (T-l)\,\boldsymbol{\delta}^T_{(i_1,\ldots,i_n)} \mathbb{U}\boldsymbol{\mu_x} \Phi_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right) \\
&+ \sum_{i_1\in d,\ldots,i_n\in d} (T-l)\,\boldsymbol{\delta}^T_{(i_1,\ldots,i_n)} \mathbb{U}\Delta_{\boldsymbol{xm}}\Sigma^{-1}_{\boldsymbol{m}}
\begin{pmatrix}
\sigma_1 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma_n
\end{pmatrix}
\boldsymbol{\Psi}_{\mathbb{R}_{\boldsymbol{m}}}\left(\Omega_{(i_1,\ldots,i_n)}\left(\boldsymbol{\mu_m},\boldsymbol{\sigma_m}\right)\right).
\end{aligned}
\tag{3.23}
$$

Proof:

Based on Equations (3.1) and (3.2), the budget constraint for the multi-asset portfolio follows

$$
\begin{aligned}
\frac{dw_t}{w_t} &= \boldsymbol{\eta}_t^T \left(\mathrm{diag}\left(\boldsymbol{p}_t\right)\right)^{-1} d\boldsymbol{p}_t + \left(1 - \boldsymbol{\eta}_t^T \mathbb{1}_n\right) r dt \\
&= \boldsymbol{\eta}_t^T \left(\mathrm{diag}\left(\boldsymbol{p}_t\right)\right)^{-1} \mathrm{diag}\left(\boldsymbol{p}_t\right)\left\{\left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t\right)dt + \mathbb{V}_p d\boldsymbol{b}_t\right\} + \left(1 - \boldsymbol{\eta}_t^T \mathbb{1}_n\right) r dt \\
&= r dt + \boldsymbol{\eta}_t^T \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - r\mathbb{1}_n\right)dt + \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t,
\end{aligned}
$$

then Equation (3.12) is proved. Since $(dt)^2 = o(dt), dt d\boldsymbol{b}_t = o(dt)$ and $d\boldsymbol{b}_t d\boldsymbol{b}_t^T = dt\mathbb{I}_n$, where $\mathbb{0}_n^T = (0, \ldots, 0)$ and $\mathbb{I}_n$ is the identity matrix,

$$\left(\frac{dw_t}{w_t}\right)^2 = \left(\boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t\right)^2 = \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t d\boldsymbol{b}_t^T \mathbb{V}_p^T \boldsymbol{\eta}_t = \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{I}_n dt \mathbb{V}_p^T \boldsymbol{\eta}_t = \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt,$$

which implies

$$\begin{aligned}
& d\left(\log w_t\right) \\
& = \left(\log w_t\right)' dw_t + \frac{1}{2}\left(\log w_t\right)'' \left(dw_t\right)^2 = \frac{dw_t}{w_t} + \frac{1}{2}\left(-\frac{1}{w_t^2}\right)\left(dw_t\right)^2 = \frac{dw_t}{w_t} - \frac{1}{2}\left(\frac{dw_t}{w_t}\right)^2 \\
& = rdt + \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - r\mathbb{1}_n\right) dt + \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t - \frac{1}{2}\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt \\
& = \left(r + \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - r\mathbb{1}_n\right) - \frac{1}{2}\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right) dt + \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t,
\end{aligned}$$

since $\log w_t \mid_0^T = \log w_T - \log w_0 = \int_0^T d\left(\log w_t\right)$ and by Equation (3.9) with $T \geq l$,

$$\begin{aligned}
& \log w_T \\
& = \log w_0 + \int_0^T d\left(\log w_t\right) \\
& = \log w_0 + \int_0^T rdt + \int_0^T \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) dt + \int_0^T \boldsymbol{\eta}_t^T \mathbb{U}\boldsymbol{x}_t dt - \frac{1}{2}\int_0^T \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt \\
& + \int_0^T \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t \\
& = \log w_0 + rT + \int_0^l \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) dt + \int_l^T \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) dt + \int_0^l \boldsymbol{\eta}_t^T \mathbb{U}\boldsymbol{x}_t dt \\
& + \int_l^T \boldsymbol{\eta}_t^T \mathbb{U}\boldsymbol{x}_t dt - \frac{1}{2}\int_0^l \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt - \frac{1}{2}\int_l^T \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt + \int_0^T \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t \\
& = \log w_0 + rT + \boldsymbol{\lambda}^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) l + \int_l^T \boldsymbol{\eta}_t^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) dt + \int_0^l \boldsymbol{\lambda}^T \mathbb{U}\boldsymbol{x}_t dt \\
& + \int_l^T \boldsymbol{\eta}_t^T \mathbb{U}\boldsymbol{x}_t dt - \frac{1}{2}\left(\boldsymbol{\lambda}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\lambda}\right) l - \frac{1}{2}\int_l^T \boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t dt + \int_0^T \boldsymbol{\eta}_t^T \mathbb{V}_p d\boldsymbol{b}_t,
\end{aligned}$$

which implies

$$
\begin{aligned}
&E\left(\log w_T\right) = \log w_0 + rT \\
&+\boldsymbol{\lambda}^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) l - \frac{1}{2}\left(\boldsymbol{\lambda}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\lambda}\right) l + \int_0^l \boldsymbol{\lambda}^T \mathbb{U} E\left(\boldsymbol{x}_t\right) dt + \int_l^T E\left(\boldsymbol{\eta}_t^T\right) dt \left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) \\
&-\frac{1}{2}\int_l^T E\left(\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right) dt + \int_l^T E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t\right) dt + \int_0^T E\left(\boldsymbol{\eta}_t^T\right) \mathbb{V}_p E\left(d\boldsymbol{b}_t\right).
\end{aligned}
$$

By Propositions 3.1, 3.2 and 3.3, we note that $E\left(\boldsymbol{\eta}_t^T\right)$, $E\left(\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right)$ and $E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t\right)$ are all independent of time t. Since $E\left(d\boldsymbol{b}_t\right) = \mathbb{0}_n$ and $E\left(\boldsymbol{x}_t\right) = -\Theta^{-1}\boldsymbol{\beta}$ by Lemma 3.2, we derive

$$
E\left(\log w_T\right) = a_6 + (T-l)\left[E\left(\boldsymbol{\eta}_t^T\right)\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) - \frac{1}{2}E\left(\boldsymbol{\eta}_t^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\eta}_t\right) + E\left(\boldsymbol{\eta}_t^T \mathbb{U} \boldsymbol{x}_t\right)\right],
$$

where $\mathbb{1}_n^T = (1, \ldots, 1)$ and a_6 is a constant depending on l, i.e.,

$$
a_6 = \log w_0 + rT + l\left[\boldsymbol{\lambda}^T\left(\boldsymbol{\alpha} - r\mathbb{1}_n\right) - \frac{1}{2}\left(\boldsymbol{\lambda}^T \mathbb{V}_p \mathbb{V}_p^T \boldsymbol{\lambda}\right) - \boldsymbol{\lambda}^T \mathbb{U}\Theta^{-1}\boldsymbol{\beta}\right],
$$

then Equation (3.22) is proved. If one put Propositions 3.1, 3.2 and 3.3 into Equation (3.22), then Equation (3.23) is proved.

Now that we have derived the equation for the expected log-utility of wealth $E\left(\log w_T\right)$, we can calculate optimal estimates of the asset allocation parameters for the MGMA strategy. In order to achieve this goal, we need to maximize $E\left(\log w_T\right)$ with respect to asset allocation parameters $\boldsymbol{\delta}_{(i_1,\ldots,i_n)}$. Suppose that the investor specific risk tolerance $\epsilon = \epsilon_0$, then for kth stock, we solve following equation for optimal estimates of $\delta^*_{k,(i_1,\ldots,i_n)}$, i.e.,

$$\left.\frac{\partial E\left(\log w_T\right)}{\partial \delta_{k,(i_1,\ldots,i_n)}}\right|_{\epsilon=\epsilon_0,\delta_{k,(i_1,\ldots,i_n)}=\delta^*_{k,(i_1,\ldots,i_n)}} = 0. \tag{3.24}$$

We also restrict $\delta_{k,(i_1,\ldots,i_n)} \in [0,1]$, which means there are no-borrowing and no-short-sale constrains, then the optimal estimates of $\boldsymbol{\delta}_{(i_1,\ldots,i_n)}$ are

$$\boldsymbol{\delta}^*_{(i_1,\ldots,i_n)} = \begin{pmatrix} \delta^*_{1,(i_1,\ldots,i_n)} \\ \vdots \\ \delta^*_{n,(i_1,\ldots,i_n)} \end{pmatrix}. \tag{3.25}$$

Notice that the optimal estimates of $\boldsymbol{\delta}^*_{(i_1,\ldots,i_n)}$ are functions of the investor specified risk tolerance ϵ. The results illustrate that the MGMA is a better investment strategy compared with the MA strategy for multi-asset portfolio because it has higher expected utility of wealth for the investor.

3.4 An Investment Algorithm for Multi-Asset Portfolio

We propose an investment algorithm for multi-asset portfolio. The algorithm will be tested on simulation data and real data in sections 3.5 and 3.6 to evaluate performance of the MGMA strategy. The algorithm contains following steps:

Step 1. Set investment parameters w_0, r and T, ϵ, $\boldsymbol{\lambda}$, s and l.

Step 2. Compute model parameters $\boldsymbol{\mu_x}$, $\boldsymbol{\mu_m}$, $\Sigma_{\boldsymbol{x}}$, $\Delta_{\boldsymbol{xm}}$, $\Sigma_{\boldsymbol{m}}$, $\boldsymbol{\sigma_m}$ and $\mathbb{R}_{\boldsymbol{m}}$.

Step 3. Compute $\boldsymbol{\delta}^*_{(i_1,\ldots,i_n)}$ and $E\left(\log w_T\right)$.

Step 4. Calculate $\boldsymbol{y}_t$, $\boldsymbol{m}_t^{(s)}$, $\boldsymbol{m}_t^{(l)}$ and $\boldsymbol{m}_t^{(s,l)}$.

Step 5. Allocate the wealth among n risky assets and one risk-free asset according to $\boldsymbol{\delta}^*_{(i_1,\ldots,i_n)}$.

Step 6. The holding risky assets are sold at the end of the investment horizon T.

3.5 Simulation Studies

We present several numerical examples based on simulated two-asset portfolio and simulated three-asset portfolio. The investment algorithm is tested and compared with the MA strategy as benchmark. For two-asset portfolio, the simplified MGMA strategy for a two-asset portfolio in table 3.2 will be used. The simulation results for three-asset portfolio are also presented.

3.5.1 Data Generating Process

We propose a data generating process for $\boldsymbol{p}_t$ and $\boldsymbol{x}_t$ time series. In order to ensure $\boldsymbol{p}_t$ be always positive, we first generate log transformed price $\boldsymbol{y}_t$ time series from Equation (3.13). We then take exponential transformation of $\boldsymbol{y}_t$ to calculate the price $\boldsymbol{p}_t$ time series. In order to ensure correlation matrix between $\boldsymbol{b}_t$ and $\boldsymbol{z}_t$ be $\mathbb{V}_{\boldsymbol{bz}}$, we use Lemma 3.13. We first generate two independent standard Brownian motions. We then transform them with a symmetric matrix Γ, which is derived from $\mathbb{V}_{\boldsymbol{bz}}$. This ensures the transformed time series be multi-dimensional standard Brownian motion with correlation matrix $\mathbb{V}_{\boldsymbol{bz}}$. Recall Equation (3.3),

$$d\boldsymbol{x}_t = (\boldsymbol{\beta} + \Theta \boldsymbol{x}_t)\, dt + \mathbb{V}_x d\boldsymbol{z}_t,$$

then

$$\boldsymbol{x}_t - \boldsymbol{x}_{t-dt} = (\boldsymbol{\beta} + \Theta \boldsymbol{x}_t)\, dt + \mathbb{V}_x d\boldsymbol{z}_t,$$

which implies

$$(\mathbb{I}_q - dt\Theta)\, \boldsymbol{x}_t = \boldsymbol{x}_{t-dt} + dt\boldsymbol{\beta} + \mathbb{V}_x d\boldsymbol{z}_t, \tag{3.26}$$

we solve above equation for $\boldsymbol{x}_t$ given initial $\boldsymbol{x}_0$, where $\mathbb{I}_q$ is identity matrix. Recall Equation (3.13),

$$d(\boldsymbol{y}_t) = \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - \frac{1}{2}\boldsymbol{\gamma}\right) dt + \mathbb{V}_p d\boldsymbol{b}_t,$$

where

$$\boldsymbol{\gamma} = \begin{pmatrix} \sum_{i=1}^{n} \left(v_{1i}^{p}\right)^2 \\ \vdots \\ \sum_{i=1}^{n} \left(v_{ni}^{p}\right)^2 \end{pmatrix},$$

then

$$\boldsymbol{y}_t - \boldsymbol{y}_{t-dt} = \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - \frac{1}{2}\gamma\right) dt + \mathbb{V}_p d\boldsymbol{b}_t,$$

which implies

$$\boldsymbol{y}_t = \boldsymbol{y}_{t-dt} + \left(\boldsymbol{\alpha} + \mathbb{U}\boldsymbol{x}_t - \frac{1}{2}\gamma\right) dt + \mathbb{V}_p d\boldsymbol{b}_t, \tag{3.27}$$

we solve above equation for $\boldsymbol{y}_t$ given initial $\boldsymbol{y}_0$ with calculated $\boldsymbol{x}_t$. The following pseudo codes describe data generating process.

Step 1. Set parameters n, q, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\mathbb{U}$, Θ, $\mathbb{V}_p$, $\mathbb{V}_x$, $\mathbb{V}_{\boldsymbol{bz}}$, $\boldsymbol{x}_0$, $\boldsymbol{y}_0$, T and dt.

Step 2. Calculate number of points, i.e., $N = T/dt$.

Step 3. Solve symmetric matrix Γ from $\mathbb{V}_{\boldsymbol{bz}}$. Solve γ from $\mathbb{V}_p$.

Step 4. Simulate independent standard Brownian motions $\boldsymbol{b}_t^{(0)}$ and $\boldsymbol{z}_t^{(0)}$. Each standard Brownian motion time series contains $N+1$ points.

Step 5. Compute standard Brownian motions $\boldsymbol{b}_t$ and $\boldsymbol{z}_t$ from Γ, $\boldsymbol{b}_t^{(0)}$ and $\boldsymbol{z}_t^{(0)}$.

Step 6. Calculate $d\boldsymbol{z}_t$ and $d\boldsymbol{b}_t$.

Step 7. For each time t (from 1 to N), recursively simulate the predictive variables $\boldsymbol{x}_t$ and n log prices $\boldsymbol{y}_t$.

Step 8. Take exponential transformation of $\boldsymbol{y}_t$ to get the price $\boldsymbol{p}_t$ time series.

3.5.2 Simulation Results for Two-Asset Portfolio

The simulated two-asset portfolio data is generated using parameters below.

$$\boldsymbol{\beta} = \begin{pmatrix} 0.0100 \\ 0.6542 \end{pmatrix}, \quad \Theta = \begin{pmatrix} -0.253 & 0 \\ 0 & 0.1438 \end{pmatrix}, \quad \mathbb{V}_x = \begin{pmatrix} 0.012 & 0 \\ 0 & 0.3356 \end{pmatrix},$$

and

$$\boldsymbol{\alpha} = \begin{pmatrix} 0.0310 \\ -0.0742 \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbb{U} = \begin{pmatrix} 2.0720 & 0.0150 \\ 0.0235 & 0.0181 \end{pmatrix},$$

and

$$\mathbb{V}_p = \begin{pmatrix} 0.195 & 0.100 \\ 0.100 & 0.495 \end{pmatrix}, \quad \mathbb{V}_{\boldsymbol{bz}} = \begin{pmatrix} -0.073 & 0.0050 \\ 0.001 & -0.9083 \end{pmatrix},$$

The simulation runs 1,000 times. Each time series contains 97,500 observed points.

The simulation studies are performed under two scenarios ($s = 5$ & $l = 30$ vs. $s = 5$ & $l = 10$). We set initial wealth $w_0 = 1,000,000$ and interest rate $r = 0$. Under each scenario, we test the MGMA strategy based on $\epsilon = 0.005, 0.01$ and 0.05 and compare with the MA strategy. The MGMA strategy performance results are provided in tables 3.4 and 3.5. We first report the theoretical expected log-utility of wealth $E(\log W_T)^*$ based on Equation (3.23) with percentage increase of the expected log-utility of wealth compared with the MA strategy. We then report numerical summaries calculate form the simulation results, including the expected log-utility of wealth $E(\log W_T)$, the expected of wealth $E(W_T)$ and the expected return on asset ratio $E(ROA\ \%)$ etc.

Table 3.4: MGMA strategy performance summary for scenario 1 on simulated two-asset portfolio (1000 run; s = 5; l = 30)

	MA	$MGMA(\epsilon = 0.005)$	$MGMA(\epsilon = 0.01)$	$MGMA(\epsilon = 0.05)$
a_1^*	na	0.03322929	0.03320624	0.033097262
a_2^*	na	0.03281636	0.03236742	0.028855955
a_3^*	na	1	1	1
a_4^*	na	1	1	1
a_5^*	na	1	1	1
$E(\log W_T)^*$	13.847667	13.896920	13.905998	13.945960
$\Delta\%\ E(\log W_T)^*$	na	0.36%	0.42%	0.71%
$E(\log W_T)$	13.794745	13.837687	13.845006	13.866623
$\log E(W_T)$	13.829619	13.861375	13.866273	13.886932
$E(W_T)$	1,014,208	1,046,932	1,052,073	1,074,033
$E(ROA\ \%)$	**1.42%**	**4.69%**	**5.21%**	**7.40%**
$SD(W_T)$	283,524	233,340	223,184	221,949
$MAX(W_T)$	3,058,106	2,273,877	2,186,479	2,176,345
$MIN(W_T)$	541,872	510,691	539,512	629,311
$MEDIAN(W_T)$	959,840	1,010,483	1,017,564	1,050,674
$E(TRANS\ \#)$	25	68	67	52

Table 3.5: MGMA strategy performance summary for scenario 2 on simulated two-asset portfolio (1000 run; s = 5; l = 10)

	MA	$MGMA(\epsilon = 0.005)$	$MGMA(\epsilon = 0.01)$	$MGMA(\epsilon = 0.05)$
a_1^*	na	0.034401876	0.034369298	0.072961411
a_2^*	na	0.033698847	0.032947417	0.113863840
a_3^*	na	1	1	1
a_4^*	na	1	1	1
a_5^*	na	1	1	1
$E(\log W_T)^*$	13.843826	13.916324	13.938047	13.966346
$\Delta\%\ E(\log W_T)^*$	na	0.52%	0.68%	0.89%
$E(\log W_T)$	13.786813	13.847683	13.863594	13.876244
$\log E(W_T)$	13.825506	13.870658	13.885702	13.901359
$E(W_T)$	1,010,045	1,056,696	1,072,714	1,089,641
$E(ROA\ \%)$	**1.00%**	**5.67%**	**7.27%**	**8.96%**
$SD(W_T)$	298,598	231,436	231,640	250,256
$MAX(W_T)$	3,633,894	2,354,587	2,202,291	2,467,362
$MIN(W_T)$	387,856	535,477	543,935	513,012
$MEDIAN(W_T)$	965,330	1,029,843	1,043,520	1,058,534
$E(TRANS\ \#)$	56	157	144	68

Notice that the MGMA strategy for two-asset portfolio not only can increase the investor's expected log-utility of wealth, but also increase the investor's expected wealth and the expected return on asset ratio from the simulation results. Under scenario 1, the expected log-utility of wealth increase in range 0.36% to 0.71%. The expected return ratio increases from benchmark return 1.42% to 4.69%, 5.21% and 7.40% respectively. Under scenario 2, the expected log-utility of wealth increase in range 0.52% to 0.89%. The expected return ratio increases from benchmark return 1.00% to 5.67%, 7.27% and 8.96% respectively.

3.5.3 Simulation Results for Three-Asset Portfolio

The simulated three-asset portfolio time series data is generated using parameters below.

$$\boldsymbol{\beta} = \begin{pmatrix} 0.010 \\ 0.065 \\ 0.185 \end{pmatrix}, \quad \Theta = \begin{pmatrix} -0.253 & 0 & 0 \\ 0 & -1.1438 & 0 \\ 0 & 0 & -1.89 \end{pmatrix}, \quad \mathbb{V}_x = \begin{pmatrix} 0.012 & 0 & 0 \\ 0 & 0.3356 & 0 \\ 0 & 0 & 0.134 \end{pmatrix},$$

and

$$\boldsymbol{\alpha} = \begin{pmatrix} 0.0310 \\ -0.0742 \\ -0.0945 \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbb{U} = \begin{pmatrix} 1.2720 & 0.0150 & 1.500 \\ 1.0235 & 1.0181 & 0.512 \\ 0.5000 & 0.0200 & 0.145 \end{pmatrix},$$

and

$$\mathbb{V}_p = \begin{pmatrix} 0.195 & 0.100 & 0.200 \\ 0.100 & 0.495 & 0.345 \\ 0.200 & 0.345 & 0.271 \end{pmatrix}, \qquad \mathbb{V}_{\boldsymbol{bz}} = \begin{pmatrix} -0.073 & 0.001 & -0.10 \\ 0.001 & -0.108 & 0.09 \\ -0.050 & 0.040 & 0.10 \end{pmatrix},$$

The simulation runs 1,000 times. Each time series contains 97,500 observed points by our settings.

The simulation studies are performed under two scenarios ($s = 5$ & $l = 30$ vs. $s = 5$ & $l = 10$). We set initial wealth $w_0 = 1,000,000$ and interest rate $r = 0$. Under each scenario, we test the MGMA strategy based on $\epsilon = 0.001$ and 0.0005 and compare with the MA strategy. The MGMA strategy performance results are provided in tables 3.6 and 3.7. We first report the theoretical expected log-utility of wealth $E(\log W_T)^*$ based on Equation (3.23) and percentage increase of the expected log-utility of wealth compared with the MA strategy. We then report numerical summaries calculate from the simulation results, including the expected log-utility of wealth $E(\log W_T)$, the expected of wealth $E(W_T)$ and the expected return on asset ratio $E(ROA\ \%)$ etc. By using no-borrowing and no-short-sale constrains, we can reduce the parameters of MGMA strategy for three-asset portfolio from 192 to 37 for implementation. For easy illustration, we do not report the optimal asset allocation parameters in simulation summary tables.

Notice that the MGMA strategy for three-asset portfolio not only can increase

Table 3.6: MGMA strategy performance summary for scenario 1 on simulated three-asset portfolio (1000 run; s = 5; l = 30)

	MA	$MGMA(\epsilon = 0.001)$	$MGMA(\epsilon = 0.005)$
$E(\log W_T)^*$	13.832331	13.911899	13.918328
$\Delta\%\ E(\log W_T)^*$	na	0.58%	0.62%
$E(\log W_T)$	13.785474	13.863884	13.868901
$\log E(W_T)$	13.834971	13.900779	13.904022
$E(W_T)$	1,019,651	1,089,009	1,092,547
$E(ROA\ \%)$	**1.97%**	**8.90%**	**9.25%**
$SD(W_T)$	339,954	308,718	303,463
$MAX(W_T)$	2,926,612	2,764,480	2,696,956
$MIN(W_T)$	366,703	421,502	458,013
$MEDIAN(W_T)$	946,339	1,038,858	1,043,233
$E(TRANS\ \#)$	36	101	101

the investor's expected log-utility of wealth, but also increase the investor's expected wealth and the expected return on asset ratio from the simulation results. Under scenario 1, the expected log-utility of wealth increase in range 0.58% to 0.62%. The expected return ratio increases from benchmark return 1.97% to 8.90% and 9.25%

Table 3.7: MGMA strategy performance summary for scenario 2 on simulated three-asset portfolio (1000 run; s = 5; l = 10)

	MA	$MGMA(\epsilon = 0.001)$	$MGMA(\epsilon = 0.005)$
$E(\log W_T)^*$	13.823737	13.914620	13.934576
$\Delta\%\ E(\log W_T)^*$	na	0.66%	0.80%
$E(\log W_T)$	13.771318	13.857019	13.875335
$\log E(W_T)$	13.827102	13.894495	13.909474
$E(W_T)$	1,011,659	1,082,187	1,098,520
$E(ROA\ \%)$	**1.17%**	**8.22%**	**9.85%**
$SD(W_T)$	356,262	307,575	297,192
$MAX(W_T)$	3,360,435	2,643,002	2,794,930
$MIN(W_T)$	419,986	479,525	504,198
$MEDIAN(W_T)$	937,137	1,025,448	1,053,388
$E(TRANS\ \#)$	82	240	237

respectively. Under scenario 2, the expected log-utility of wealth increase in range 0.66% to 0.80%. The expected return ratio increases from benchmark return 1.17% to 8.22% and 9.85% respectively.

3.6 Real Data Applications

We present several real data analysis based on high-frequency exchange traded fund (ETF) data. The investment algorithm is tested and compared with the benchmark. The simplified MGMA strategy for a two-asset portfolio in table 3.2 will be used. The MA strategy for a two-asset portfolio in table 3.3 as benchmark strategy.

We use PowerShares QQQ Trust Series 1 (QQQ) and SPDR S&P 500 ETF Trust (SPY). These are exchange-traded funds incorporated in the USA. QQQ ETF tracks performance of the Nasdaq 100 Index. It holds large cap U.S. stocks and tends to focus on technology and consumer sector. The holdings are weighted by market capitalization. As of October 6, 2017, there are 107 holding companies. The top 3 holding companies are Apple Inc (AAPL, 11.57%), Microsoft Corp (MSFT, 8.44%) and Amazon.com Inc (AMZN, 6.86%). SPY ETF tracks the S&P 500 Index. The Trust consists of a portfolio representing all 500 stocks in the S&P 500 Index. It holds predominantly large-cap U.S. stocks. It is structured as a Unit Investment Trust and pays dividends on a quarterly basis. The holdings are weighted by market capitalization. As of October 6, 2017, the top 3 holding companies are Apple Inc (AAPL, 3.67%), Microsoft Corp (MSFT, 2.68%) and Facebook Inc. Class A (FB, 1.87%).

3.6.1 An Algorithm to Estimate Model Parameters

In order to use the investment algorithm for multi-asset portfolio on real data, we need to estimate model parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\mathbb{U}$, Θ, $\mathbb{V}_p$, $\mathbb{V}_x$ and $\mathbb{V}_{\boldsymbol{bz}}$. There is no such algorithm in literature due to complex model settings. We propose an algorithm to fill the gap. Without loss generality, we describe the algorithm by using general model for two-asset portfolio. Based on Equation (3.2),

$$\begin{aligned}\frac{dp_{1t}}{p_{1t}} &= (\alpha_1 + u_{11}x_{1t} + u_{12}x_{2t})\,dt + e^p_{1t},\\ \frac{dp_{2t}}{p_{2t}} &= (\alpha_2 + u_{21}x_{1t} + u_{22}x_{2t})\,dt + e^p_{2t},\end{aligned} \tag{3.28}$$

and Equation (3.3),

$$\begin{aligned}dx_{1t} &= (\beta_1 + \theta_{11}x_{1t} + \theta_{12}x_{2t})\,dt + e^x_{1t},\\ dx_{2t} &= (\beta_2 + \theta_{21}x_{1t} + \theta_{22}x_{2t})\,dt + e^x_{2t},\end{aligned} \tag{3.29}$$

where

$$\boldsymbol{e}^p_t = \begin{pmatrix} e^p_{1t} \\ e^p_{2t} \end{pmatrix} = \begin{pmatrix} v^p_{11}db_{1t} + v^p_{12}db_{2t} \\ v^p_{21}db_{1t} + v^p_{22}db_{2t} \end{pmatrix} = \mathbb{V}_p d\boldsymbol{b}_t \sim \mathrm{MN}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix},\quad dt\mathbb{V}_p\mathbb{V}_p^T\right], \tag{3.30}$$

and

$$\boldsymbol{e}^x_t = \begin{pmatrix} e^x_{1t} \\ e^x_{2t} \end{pmatrix} = \begin{pmatrix} v^x_{11}dz_{1t} + v^x_{12}dz_{2t} \\ v^x_{21}dz_{1t} + v^x_{22}dz_{2t} \end{pmatrix} = \mathbb{V}_x d\boldsymbol{z}_t \sim \mathrm{MN}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix},\quad dt\mathbb{V}_x\mathbb{V}_x^T\right]. \tag{3.31}$$

Let $dt\mathbb{V}_p\mathbb{V}_p^T \triangleq \Sigma_{e^p}$, then it is easy to check the log-likelihood function for $\boldsymbol{e}_t^p$ is

$$l\left(\Sigma_{e^p} \mid \boldsymbol{e}_t^p\right) = -\frac{T}{2}\log\left|\Sigma_{e^p}\right| - \frac{1}{2}\sum_{t=1}^{T}\left\{\left(\boldsymbol{e}_t^p\right)^T \Sigma_{e^p}^{-1}\boldsymbol{e}_t^p\right\} - T\log(2\pi), \tag{3.32}$$

and let $dt\mathbb{V}_x\mathbb{V}_x^T \triangleq \Sigma_{e^x} = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix}$, then the log-likelihood function for $\boldsymbol{e}_t^x$ is

$$\begin{aligned} l\left(\Sigma_{e^x} \mid \boldsymbol{e}_t^x\right) &= -\frac{T}{2}\log\left|\Sigma_{e^x}\right| - \frac{1}{2}\sum_{t=1}^{T}\left\{\left(\boldsymbol{e}_t^x\right)^T \Sigma_{e^x}^{-1}\boldsymbol{e}_t^x\right\} - T\log(2\pi) \\ &= -\frac{T}{2}\log\left(v_1v_2\right) - \frac{1}{2}\sum_{t=1}^{T}\left\{\frac{\left(e_{1t}^x\right)^2}{v_1} + \frac{\left(e_{2t}^x\right)^2}{v_2}\right\} - T\log(2\pi). \end{aligned} \tag{3.33}$$

Let $Cov(d\boldsymbol{b}_t, d\boldsymbol{z}_t) \triangleq \Sigma_{\boldsymbol{bz}}$, it is also easy to verify that

$$\Sigma_{\boldsymbol{bz}} = dt\mathbb{V}_{\boldsymbol{bz}}. \tag{3.34}$$

Then, the algorithm contains following steps:

Step 1. Given a dt, calculate dp_{1t}, dp_{2t}, dx_{1t} and dx_{2t} (for $t > 1$) based on the historical time series.

Step 2. Use least square estimation method to estimate parameters $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{u}_{11}$, $\hat{u}_{12}$, $\hat{u}_{21}$ and $\hat{u}_{22}$ by minimizing

$$\sum_{i=1}^{2}\sum_{t=2}^{T}\left[\frac{dp_{it}}{p_{it}} - \left(\alpha_i + u_{i1}x_{1t} + u_{i2}x_{2t}\right)dt\right]^2.$$

Step 3. Let $\Theta = \text{diag}\,(\theta_{11}, \theta_{22})$, use least square estimation method to estimate parameters $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\theta}_{11}$ and $\hat{\theta}_{22}$ by minimizing

$$\sum_{i=1}^{2} \sum_{t=2}^{T} \left[dx_{it} - (\beta_i + \theta_{i1} x_{1t} + \theta_{i2} x_{2t})\, dt\right]^2 .$$

Step 4. Calculate $\hat{\boldsymbol{e}}_t^p$ and $\hat{\boldsymbol{e}}_t^x$ from p_{1t}, p_{2t}, x_{1t} and x_{2t}, $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{u}_{11}$, $\hat{u}_{12}$, $\hat{u}_{21}$, $\hat{u}_{22}$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\theta}_{11}$, $\hat{\theta}_{12}$, $\hat{\theta}_{21}$ and $\hat{\theta}_{22}$.

Step 5. Use maximize likelihood estimation method and set

$$\frac{\partial l\left(\Sigma_{e^p} \mid \hat{\boldsymbol{e}}_t^p\right)}{\partial \Sigma_{e^p}} = 0,$$

to estimate $\widehat{\Sigma}_{e^p}$ from $\hat{\boldsymbol{e}}_t^p$. Since $\widehat{\mathbb{V}}_p = \left(\frac{1}{dt}\widehat{\Sigma}_{e^p}\right)^{\frac{1}{2}}$, we can estimate parameters $\hat{v}_{11}^p$, $\hat{v}_{12}^p$, $\hat{v}_{21}^p$ and $\hat{v}_{22}^p$.

Step 6. Use maximize likelihood estimation method and set

$$\frac{\partial l\left(\Sigma_{e^x} \mid \hat{\boldsymbol{e}}_t^x\right)}{\partial \Sigma_{e^x}} = 0,$$

to estimate $\widehat{\Sigma}_{e^x}$, $\hat{v}_{11}^x$, $\hat{v}_{12}^x$, $\hat{v}_{21}^x$ and $\hat{v}_{22}^x$ from $\hat{\boldsymbol{e}}_t^x$.

Step 7. Calculate $d\hat{\boldsymbol{b}}_t$ and $d\hat{\boldsymbol{z}}_t$ from $\widehat{\mathbb{V}}_p$, $\widehat{\mathbb{V}}_x$, $\hat{\boldsymbol{e}}_t^p$ and $\hat{\boldsymbol{e}}_t^x$.

Step 8. Calculate $\widehat{\Sigma}_{\boldsymbol{bz}}$ from $d\hat{\boldsymbol{b}}_t$ and $d\hat{\boldsymbol{z}}_t$. Then the estimated parameter is $\widehat{\mathbb{V}}_{\boldsymbol{bz}} = \frac{1}{dt}\widehat{\Sigma}_{\boldsymbol{bz}}$.

3.6.2 Case 1: MGMA Strategy on High-Frequency Exchange Traded Fund with Observed Predictive Variables

We collect daily second-level QQQ ETF, SPY ETF, MSFT and AAPL price time series for this study. The QQQ ETF price time series and SPY ETF price time series are used as the vector based ETF price $\boldsymbol{p}_t$. The MSFT and AAPL stock price time series are used as the vector predictive variable $\boldsymbol{x}_t$. The collection period is daily trading time from 9:30 am to 4:00 pm (Eastern Time) to ensure high liquid market. We divide QQQ ETF and SPY ETF time series into two data: vector based ETF price $\boldsymbol{p}_t$ training data (9:30 am to 3:00 pm, which contains 19,800 seconds) and vector based ETF price $\boldsymbol{p}_t$ test data (3:00 pm to 4:00 pm, which contains 3,601 seconds). We use the MSFT and AAPL price time series as vector based predictive variable $\boldsymbol{x}_t$ training data (9:30 am to 3:00 pm, which contains 19,800 seconds). We set initial wealth $w_0 = 10,000$ and interest rate $r = 0$. Suppose that the investor's risk tolerance is 0.000001. We restrict a_1, a_2, a_3, a_4 and a_5 in $[0, 1]$, s in $5, 10$ and l in $30, 60, 90, 120, 180, 240$. We use training data to choose model parameters with the highest return. We first report the MGMA strategy performance summary for QQQ ETF and SPY ETF on training data, then we report the MGMA strategy evaluation summary for QQQ ETF and SPY ETF on test data. Our study spans five days from 10/2/2017 to 10/6/2017.

Let us use 10/2/2017 as an example first, then we will report all results for 5 days. Second-level QQQ ETF and SPY ETF price time series on day 1 (10/02/2017) are provided in figure 3.1. The MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) training data is provided in table 3.8. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) test data is provided in table 3.9.

Notice that (1) The MGMA strategy can increase daily return ratio from 0.09498% to 0.24542% on training data, which equals to increase annual return ratio by 46.1%; (2) The MGMA strategy can increase daily return ratio from 0.06668% to 0.08534% on test data, which equals to increase annual return ratio by 4.8%.

We repeat this study for four more days (10/03/2017 to 10/06/2017). The MGMA strategy performance summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data is provided in table 3.10. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data is provided in table 3.11.

The MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data is provided in figure 3.2. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data is provided in figure 3.3.

Figure 3.1: Case 1: Second-level QQQ ETF and SPY ETF prices time series on day 1 (10/02/2017)

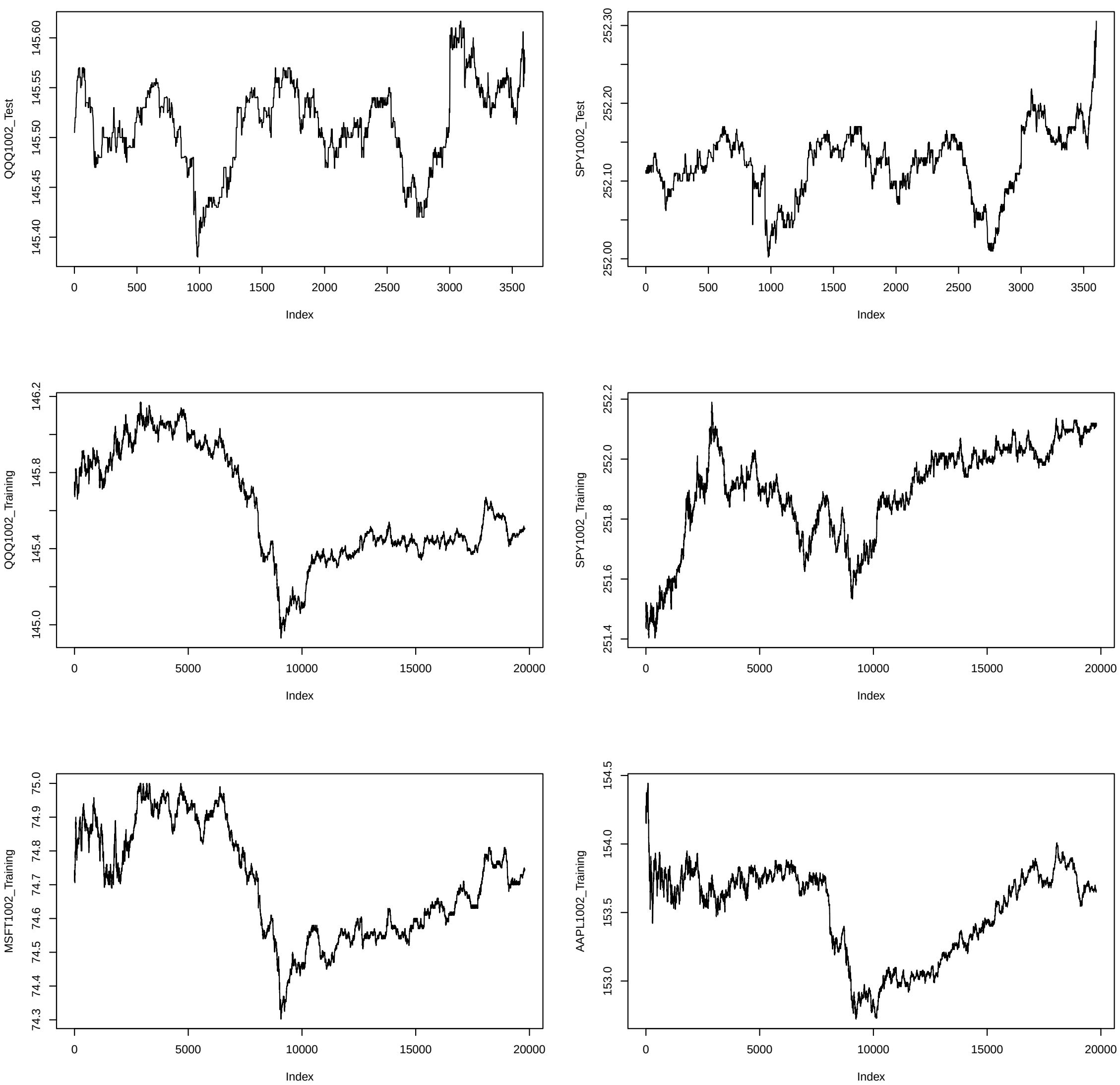


Table 3.8: Case 1: MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) training data

training data	*day*	10/2/2017
	time	$9:30\ am - 3:00\ pm$
	T	$19,800\ seconds$
	dt	$1\ second$
tuned parameters	s	10
	l	180
	a_1^*	0.3287
	a_2^*	0.9611
	a_3^*	0.0352
	a_4^*	0.7160
	a_5^*	1
backward MA	$E(W_T)$	10,009.49768
	return ratio (%)	**0.09498%**
	trans num	504
backward MGMA	$E(W_T)$	10,024.54175
	return ratio (%)	**0.24542%**
	trans num	655

Table 3.9: Case 1: MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) test data

test data	*day*	10/2/2017
	time	$3:00\ pm - 4:00\ pm$
	T	$3,601\ seconds$
	dt	$1\ second$
tuned parameters	s	10
	l	180
	a_1^*	0.3287
	a_2^*	0.9611
	a_3^*	0.0352
	a_4^*	0.7160
	a_5^*	1
$forward\ MA$	$E(W_T)$	10,006.66819
	return ratio (%)	**0.06668%**
	trans num	67
$forward\ MGMA$	$E(W_T)$	10,008.53428
	return ratio (%)	**0.08534%**
	trans num	82

Table 3.10: Case 1: MGMA strategy performance summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data

training data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$	$9:30\ am - 3:00\ pm$
	T	$19,800\ seconds$	$19,800\ seconds$	$19,800\ seconds$	$19,800\ seconds$
	dt	$1\ second$	$1\ second$	$1\ second$	$1\ second$
tuned parameters	s	10	10	10	10
	l	60	240	180	120
	a_1^*	0.1250	0.0742	0.0737	0.8537
	a_2^*	0.2346	0.7111	0	0.2780
	a_3^*	0	0	1	0
	a_4^*	0	0	1	0
	a_5^*	1	1	1	1
backward MA	$E(W_T)$	9,955.07379	10,001.35849	10,002.15772	9,997.54490
	return ratio (%)	**-0.44926%**	**0.01358%**	**0.02158%**	**-0.02455%**
	trans num	1,117	421	535	589
backward MGMA	$E(W_T)$	9,974.19921	10,010.97889	10,015.65524	10,016.17125
	return ratio (%)	**-0.25801%**	**0.10979%**	**0.15655%**	**0.16171%**
	trans num	1,557	534	708	794

Notice that the MGMA strategy in general can outperform the MA strategy for both backward investment on training data and forward investment on test data.

Table 3.11: Case 1: MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data

test data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$
	T	$3,601\ seconds$	$3,601\ seconds$	$3,601\ seconds$	$3,601\ seconds$
	dt	$1\ second$	$1\ second$	$1\ second$	$1\ second$
tuned parameters	s	10	10	10	10
	l	60	240	180	120
	a_1^*	0.1250	0.0742	0.0737	0.8537
	a_2^*	0.2346	0.7111	0	0.2780
	a_3^*	0	0	1	0
	a_4^*	0	0	1	0
	a_5^*	1	1	1	1
$forward\ MA$	$E(W_T)$	9,990.21402	9,998.84086	10,003.58666	10,001.47801
	return ratio (%)	**-0.09786%**	**-0.01159%**	**0.03587%**	**0.01478%**
	trans num	210	85	92	111
$forward\ MGMA$	$E(W_T)$	9,991.06108	10,000.43623	10,008.31762	9,999.77159
	return ratio (%)	**-0.08939%**	**0.00436%**	**0.08318%**	**-0.00228%**
	trans num	318	127	120	141

3.6.3 Case 2: MGMA Strategy on High-Frequency Exchange Traded Fund without Observed Predictive Variables

We collect daily second-level QQQ ETF and SPY ETF price time series for this study. The collection period is daily trading time from 9:30 am to 4:00 pm (Eastern Time) to ensure high liquid market. We divide QQQ ETF and SPY ETF time

Figure 3.2: Case 1: MGMA strategy performance summary plot for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data

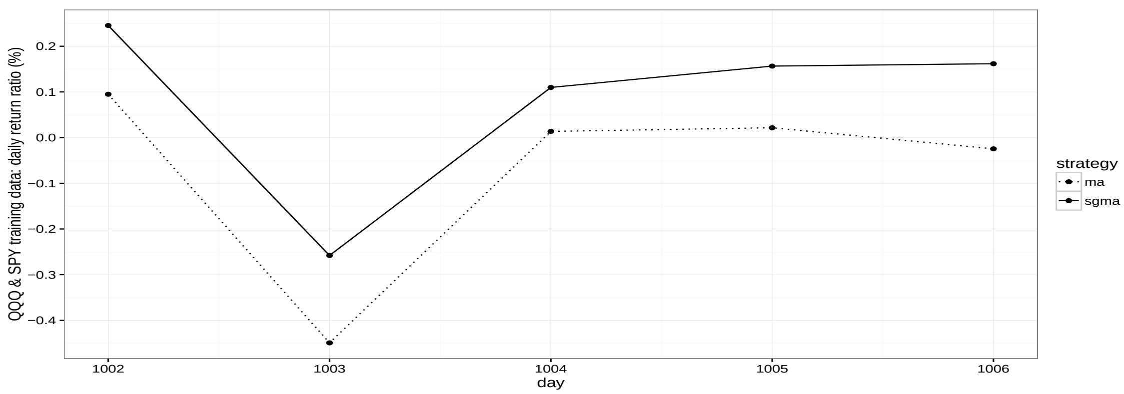
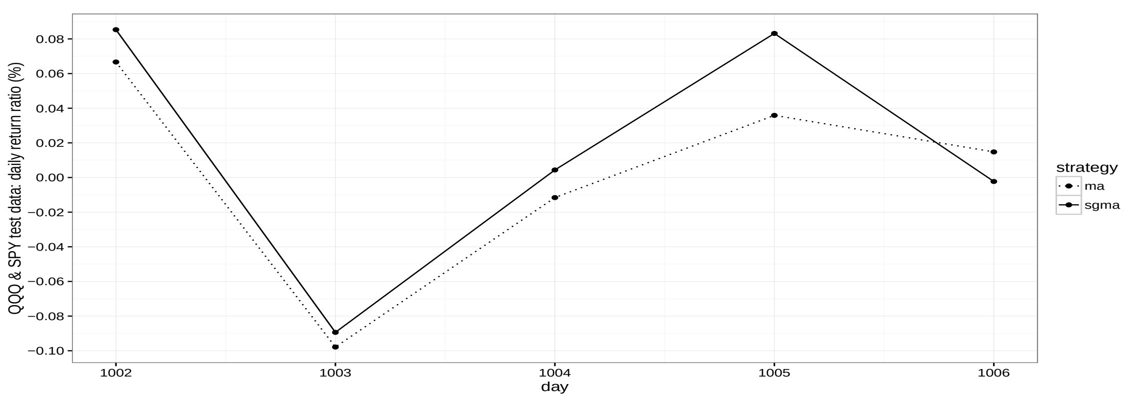


Figure 3.3: Case 1: MGMA strategy evaluation summary plot for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data



series into three data: vector based predictive variable $\boldsymbol{x}_t$ training data (9:30 am to 12:15 pm, which contains 9,900 seconds), vector based ETF price $\boldsymbol{p}_t$ training data (12:15 pm to 3:00 pm, which contains 9,900 seconds) and vector based ETF price $\boldsymbol{p}_t$ test data (3:00 pm to 4:00 pm, which contains 3,601 seconds). We set initial wealth $w_0 = 10,000$ and interest rate $r = 0$. Suppose that the investor's risk tolerance is 0.000001. We restrict a_1, a_2, a_3, a_4 and a_5 in $[0, 1]$, s in $5, 10$ and l in $30, 60, 90, 120, 180, 240$. We use training data to choose model parameters with the highest return. We first report the MGMA strategy performance summary for QQQ ETF and SPY ETF on training data, then we report the MGMA strategy evaluation summary for QQQ ETF and SPY ETF on test data. Our study spans five days from 10/2/2017 to 10/6/2017.

Let us use 10/2/2017 as an example first, then we will report all results for 5 days. Second-level QQQ ETF and SPY ETF price time series on day 1 (10/02/2017) is provided in figure 3.4. The MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) training data is provided in table 3.12. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) test data is provided in table 3.13.

Notice that (1) The MGMA strategy can increase daily return ratio from 0.12898% to 0.34159% on training data, which equals to increase annual return ratio by 70.8%;

Figure 3.4: Case 2: Second-level QQQ ETF and SPY ETF prices time series on day 1 (10/02/2017)

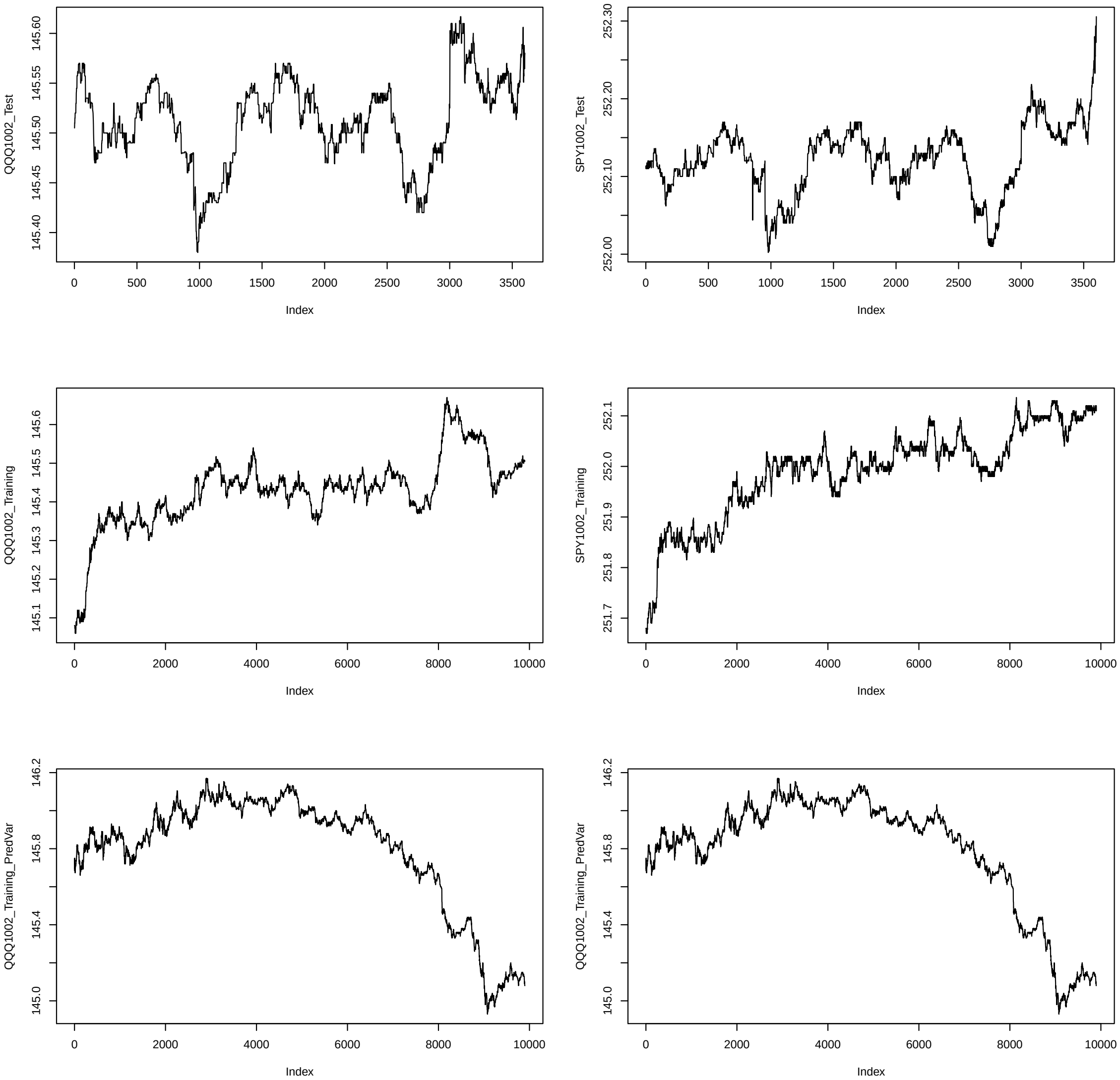


Table 3.12: Case 2: MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) training data

training data	*day*	10/2/2017
	time	$12:15\ pm - 3:00\ pm$
	T	$9,900\ seconds$
	dt	$1\ second$
tuned parameters	s	10
	l	180
	a_1^*	0.9550
	a_2^*	0.1178
	a_3^*	0.9603
	a_4^*	0.6566
	a_5^*	1
backward MA	$E(W_T)$	10,012.89790
	return ratio (%)	**0.12898%**
	trans num	253
backward MGMA	$E(W_T)$	10,034.15884
	return ratio (%)	**0.34159%**
	trans num	354

(2) The MGMA strategy can increase daily return ratio from 0.06668% to 0.07846% on test data, which equals to increase annual return ratio by 3.0%.

We repeat this study for four more days (10/03/2017 to 10/06/2017). The

Table 3.13: Case 2: MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) test data

test data	*day*	10/2/2017
	time	$3:00\ pm - 4:00\ pm$
	T	$3,601\ seconds$
	dt	$1\ second$
tuned parameters	s	10
	l	180
	a_1^*	0.9550
	a_2^*	0.1178
	a_3^*	0.9603
	a_4^*	0.6566
	a_5^*	1
forward MA	$E(W_T)$	10,006.66819
	return ratio (%)	**0.06668%**
	trans num	67
forward MGMA	$E(W_T)$	10,007.84608
	return ratio (%)	**0.07846%**
	trans num	82

MGMA strategy performance summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data is provided in table 3.14. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data is provided in table 3.15.

Table 3.14: Case 2: MGMA strategy performance summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) training data

training data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$	$12:15\ pm - 3:00\ pm$
	T	$9,900\ seconds$	$9,900\ seconds$	$9,900\ seconds$	$9,900\ seconds$
	dt	$1\ second$	$1\ second$	$1\ second$	$1\ second$
tuned parameters	s	10	10	10	10
	l	240	180	30	240
	a_1^*	0.8019	0.4776	0.4748	0.8657
	a_2^*	0.0606	0.2472	0.2444	0.1369
	a_3^*	0.1330	0.0768	0.3815	0
	a_4^*	0.2808	0.6948	1	0
	a_5^*	0	1	1	1
backward MA	$E(W_T)$	9,973.93692	9,993.24879	9,967.95729	9,994.78793
	return ratio (%)	**-0.26063%**	**-0.06751%**	**-0.32043%**	**-0.05212%**
	trans num	291	224	842	225
backward MGMA	$E(W_T)$	9,980.20191	9,998.79192	9,987.20274	10,000.99791
	return ratio (%)	**-0.19798%**	**-0.01208%**	**-0.12797%**	**0.00998%**
	trans num	406	290	1,227	280

The MGMA strategy performance summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data is provided in figure 3.5. The MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data is provided in figure 3.6. Notice that the MGMA strategy in general can outperform the MA strategy for both backward

Table 3.15: Case 2: MGMA strategy evaluation summary for QQQ ETF and SPY ETF on day 2 (10/03/2017) to day 5 (10/06/2017) test data

test data	*day*	10/03/2017	10/04/2017	10/05/2017	10/06/2017
	time	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$	$3:00\ pm - 4:00\ pm$
	T	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*	$3,601$ *seconds*
	dt	1 *second*	1 *second*	1 *second*	1 *second*
tuned parameters	s	10	10	10	10
	l	240	180	30	240
	a_1^*	0.8019	0.4776	0.4748	0.8657
	a_2^*	0.0606	0.2472	0.2444	0.1369
	a_3^*	0.1330	0.0768	0.3815	0
	a_4^*	0.2808	0.6948	1	0
	a_5^*	0	1	1	1
forward MA	$E(W_T)$	10,004.41162	9,996.02456	9,998.02785	10,003.14985
	return ratio (%)	**0.04412%**	**-0.03975%**	**-0.01972%**	**0.03150%**
	trans num	62	94	308	86
forward MGMA	$E(W_T)$	10,004.72732	9,998.96704	10,000.42527	10,002.81663
	return ratio (%)	**0.04727%**	**-0.01033%**	**0.00425%**	**0.02817%**
	trans num	76	139	472	110

investment on training data and forward investment on test data. It is expected that case 2 study shows under-performance compared with case 1 study because case 1 study involves additional information from Microsoft and Apple time series.

Figure 3.5: Case 2: MGMA strategy performance summary plot for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) training data

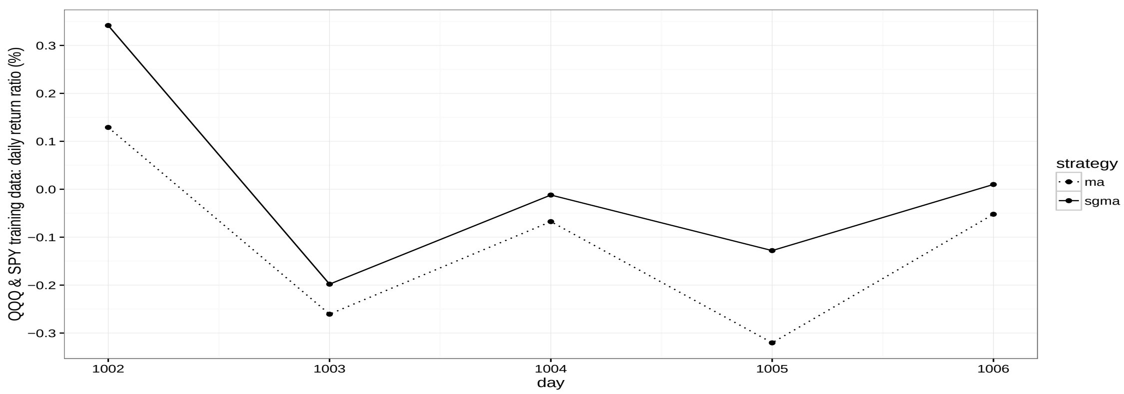
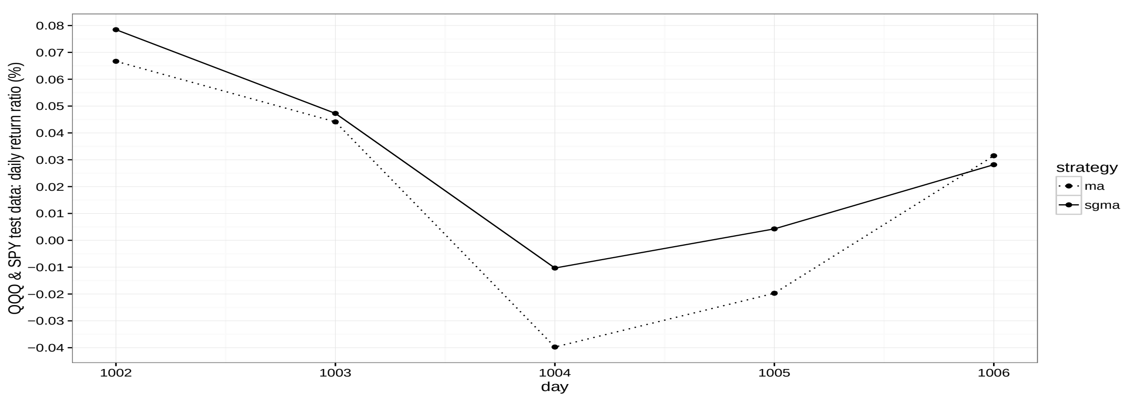


Figure 3.6: Case 2: MGMA strategy evaluation summary plot for QQQ ETF and SPY ETF on day 1 (10/02/2017) to day 5 (10/06/2017) test data



3.7 Conclusion

Base on the simulation studies and real data analysis, we observe followings: (1) The MGMA strategy can provide more investment options with the investor's risk tolerance, which can overcome the well-known drawback from the MA strategy; (2) The MGMA strategy can increase the investor's expected log-utility of wealth compared with the MA strategy; (3) The MGMA strategy is also able to increase the investor's expected wealth compared with the MA strategy.

4 Generalized Regime-Switching Model with its Application

In this chapter, we study statistical modeling with its application to capture the financial market behavior. The SGMA and MGMA strategies involve the risk tolerance specified by investor. Intuitively, an investor should take more risk in the upward market than in the downward market. This motivates us to study the movements of stocks or equity indices. The Markov regime-switching model is widely used for this type analysis. Notice that model parameters estimated under normality assumption might not be stable and the corresponding change-point detection algorithm introduced by Guo *et al.* (2011) might face some challenges when either the empirical distribution is heavy-tailed or observed data contain outliers. We relax the normality assumption and propose a generalized regime-switching (generalized RS) model. We assume error terms follow a general class of density functions. The generalized RS model can improve the stability of model parameter estimation and the corresponding change-point detection algorithm for financial time series. We adopt four assumptions for the algorithm: (1) The stock market always moves from a period of rising stock prices (bull phase) to a period of declining stock prices (bear phase) or vise versa. (2) No matter what phase the market is in, stock or index price always fluctuates and can be described by two regimes. For example, a lower-return regime and higher-return regime. (3) The time series may be non-stationary but it is locally stationary. (4) There is a relatively long grace period after the market moves into a new phase.

Chapter 4 is organized as follows. Section 4.1 introduces the generalized RS model and its likelihood function. Simulation studies in section 4.1.2 demonstrate that the generalized RS model can improve the stability of model parameter estimation. In section 4.2, a change-point detection algorithm using the generalized RS model is provided. Simulation studies are conducted in section 4.3 to evaluate the performance of the algorithm. In section 4.4, we apply the change-point detection algorithm to Hang Seng monthly index data and test the performance. The conclusion is presented in section 4.5.

4.1 The generalized RS Model with its Likelihood Function

4.1.1 The Methodology

Let P_t be the price of a stock or an equity index at a particular month t. The log-return of the $(t+1)$th month is denoted by $Y_t = \omega \log(P_{t+1}/P_t)$, where ω is a suitable constant. For example, $\omega = 100$. Under a generalized Markov regime-switching (generalized RS) model, the monthly log-return Y_t is assumed to be in one of K different regimes or states $\{r_t\}$ at any given time t, where $r_t = 1, 2, \ldots, K$. Assume that $\{Y_t\}$ satisfies the following Markov state process, i.e.,

$$Y_t = \mu_{r_t} + \sigma_{r_t} e_t, \tag{4.1}$$

where μ_r and σ_r^2 are mean and variance of Y_t when it stays in regime r_t, and e_t are independent and identically distributed (I.I.D.) error terms with density g. We will discuss a general class of density functions for g later. The model in Equation (4.1) shows that $\{Y_t\}$ are serially uncorrelated series given the regime at time t. For demonstration purposes, we consider a two-regime generalized RS model, i.e., $r_t = 1$ or 2 as $K = 2$. Let $\{\mu_1, \sigma_1\}$ and $\{\mu_2, \sigma_2\}$ be expected log-returns and corresponding volatilities of two different regimes, respectively. Let $\mathbb{P}$ be a transition matrix for the two-regime generalized RS model as it follows a stationary Markov process. The element p_{ij} of $\mathbb{P}$ is a transition probability from regime i to regime j, i.e.,

$$p_{ij} = P\left[r_t = j \mid r_{t-1} = i\right], \tag{4.2}$$

where $i = 1, 2$ and $j = 1, 2$. Define regime 1 as the low-return state and regime 2 as the high-return state, then $\mu_1 < \mu_2$. Thus, there are 6 parameters to determine a two-regime generalized RS model, i.e., $\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{11}, p_{22}\}$. In addition, a stationary distribution $\boldsymbol{\pi} = \left(\pi_1, \pi_2\right)$ of the two regimes can be defined as

$$\boldsymbol{\pi} \cdot \mathbb{P} = \boldsymbol{\pi}, \tag{4.3}$$

which implies that

$$\pi_1 = p_{21}/(p_{21} + p_{12}) \;\;\&\;\; \pi_2 = p_{12}/(p_{21} + p_{12}). \tag{4.4}$$

It means that the process lies in regime i with probability π_i, $i = 1, 2$, with $\pi_1 + \pi_2 = 1$ at any time when no historical information is available.

We use maximum likelihood estimation (MLE) method to estimate model parameters. For a two-regime generalized RS model in Equation (4.1), we have the set of parameters $\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{11}, p_{22}\}$. Let $I_t = \{y_1, \ldots, y_t\}$ represent the information available through time t and $I_0 = \varnothing$. The likelihood function based on $\boldsymbol{y} = (y_1, \ldots, y_T)$ is then given as follows:

$$L\left(\Theta \mid y_1, \ldots, y_T\right) = \prod_{t=1}^{T} f\left(y_t \mid I_{t-1}, \Theta\right), \tag{4.5}$$

then the maximum likelihood estimate of Θ is the value $\widehat{\Theta}$ that maximizes $L(\Theta \mid y_1, \ldots, y_T)$, i.e.,

$$\widehat{\Theta} = \arg\max_{\Theta} L(\Theta \mid y_1, \ldots, y_T) = \arg\max_{\Theta} \prod_{t=1}^{T} f(y_t \mid I_{t-1}, \Theta). \tag{4.6}$$

For a given I_{t-1}, $f(y_t \mid I_{t-1}, \Theta)$ is the sum of $f(r_t, y_t \mid I_{t-1}, \Theta)$ over all possible values of r_t given by:

$$f(y_t \mid I_{t-1}, \Theta) = \sum_{r_t=1}^{2} f(r_t, y_t \mid I_{t-1}, \Theta). \tag{4.7}$$

We apply the Hamilton filter (Hamilton, 1989) to estimate model parameters and compute filtered probabilities at each regime. To be more specific, for $t > 1$, the filter can be recursively calculated as below:

$$f(r_t, y_t \mid I_{t-1}, \Theta) = f(y_t \mid r_t, \Theta) \sum_{i=1}^{2} P(r_t \mid r_{t-1} = i, \Theta) P(r_{t-1} = i \mid I_{t-1}, \Theta), \tag{4.8}$$

where $P(r_{t-1} \mid I_{t-1}, \Theta)$ is the filtered probability recursively updated as below:

$$P(r_{t-1} \mid I_{t-1}, \Theta) = \frac{f(r_{t-1}, y_{t-1} \mid I_{t-2}, \Theta)}{f(y_{t-1} \mid I_{t-2}, \Theta)}, \tag{4.9}$$

and for $t \leq 1$,

$$P(r_t = i \mid \Theta) = \pi_i, \quad i = 1, 2. \tag{4.10}$$

We now discuss the functional form of $f(y_t \mid r_t, \Theta)$ in Equation (4.8). Guo *et al.* (2011) assume that $f(y_t \mid r_t, \Theta)$ is the density function for a Normal distribution because e_t for any time t is assumed to follow a Normal distribution $N(0, 1)$. However, parameters estimated under normality assumption might not be stable when either the empirical distribution is heavy-tailed or observed data contain outliers. To rectify this problem and improve the stability of model parameter estimation, we assume error terms e_t follow a general class of density functions g. In sprit of M-estimation, suppose

$$g(x) = \frac{1}{c}\ e^{-\rho(x)}, \tag{4.11}$$

where $\rho(x)$ is a M-estimation ρ-function and c is a normalizing constant satisfying

$$c = \int e^{-\rho(x)}\ dx. \tag{4.12}$$

Then $f\left(y_t \mid r_t, \Theta\right)$ is the density function given by

$$f\left(y_t \mid r_t, \Theta\right) = \frac{1}{\sigma_{r_t}} g\left(\frac{y_t - \mu_{r_t}}{\sigma_{r_t}}\right) = \frac{1}{c\ \sigma_{r_t}} e^{-\rho\left(\frac{y_t - \mu_{r_t}}{\sigma_{r_t}}\right)}. \tag{4.13}$$

If one put Equations (4.7)-(4.13) into Equation (4.6), the maximum likelihood estimator $\widehat{\Theta}$ of the two-regime generalized RS model can be re-written as

$$\widehat{\Theta} = \arg\max_{\Theta} \prod_{t=1}^{T} \left\{\sum_{j=1}^{2} \frac{1}{c\ \sigma_j} e^{-\rho\left(\frac{y_t-\mu_j}{\sigma_j}\right)} \sum_{i=1}^{2} P_{ij} P\left(r_{t-1} = i \mid I_{t-1}, \Theta\right)\right\}. \tag{4.14}$$

Given that logarithm is an increasing function, we can take log transformation and Equation (4.14) can be re-written as

$$\widehat{\Theta} = \arg\min_{\Theta} \sum_{t=1}^{T} -\log\left\{\sum_{j=1}^{2} \frac{1}{c\ \sigma_j} e^{-\rho\left(\frac{y_t-\mu_j}{\sigma_j}\right)} \sum_{i=1}^{2} P_{ij} P\left(r_{t-1} = i \mid I_{t-1}, \Theta\right)\right\}. \tag{4.15}$$

We then use a numeric optimization method to compute the maximum likelihood estimator $\widehat{\Theta}$ of the two-regime generalized RS model by recursively computing the filtered probability $P\left(r_{t-1} \mid I_{t-1}, \Theta\right)$. However, this is computational intensive. Instead, we derive an explicit form without using filtered probability. The justification of the explicit form is presented as below:

$$\widehat{\Theta} = \arg\max_{\Theta} \left\{ \boldsymbol{\pi} \left(\prod_{t=1}^{T} \mathbb{P}\mathbb{G}_t\left(y_t\right) \right) \mathbb{1} \right\}, \tag{4.16}$$

where

$$\boldsymbol{\pi} = \begin{pmatrix} \pi_1, \pi_2 \end{pmatrix}, \qquad \mathbb{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \qquad \mathbb{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

and

$$\mathbb{G}_t\left(y_t\right) = \begin{pmatrix} f\left(y_t \mid r_t = 1, \Theta\right) & 0 \\ 0 & f\left(y_t \mid r_t = 2, \Theta\right) \end{pmatrix} = \begin{pmatrix} \frac{1}{c\ \sigma_1} e^{-\rho\left(\frac{y_t-\mu_1}{\sigma_1}\right)} & 0 \\ 0 & \frac{1}{c\ \sigma_2} e^{-\rho\left(\frac{y_t-\mu_2}{\sigma_2}\right)} \end{pmatrix}.$$

Let $F_{t,j} = f\left(r_t = j, y_t \mid I_{t-1}, \Theta\right)$ where $j = 1, 2$. Then

$$\begin{aligned}
F_{t,j} = f\left(r_t = j, y_t \mid I_{t-1}, \Theta\right) &= f\left(y_t \mid r_t = j, \Theta\right) P\left(r_t = j \mid I_{t-1}, \Theta\right) \\
&= f\left(y_t \mid r_t = j, \Theta\right) \sum_{i=1}^{2} P\left(r_t = j \mid r_{t-1} = i, \Theta\right) P\left(r_{t-1} = i \mid I_{t-1}, \Theta\right) \\
&= f\left(y_t \mid r_t = j, \Theta\right) \sum_{i=1}^{2} p_{ij} P\left(r_{t-1} = i \mid I_{t-1}, \Theta\right) \\
&= f\left(y_t \mid r_t = j, \Theta\right) \sum_{i=1}^{2} p_{ij} \frac{f\left(r_{t-1} = i, y_{t-1} \mid I_{t-2}, \Theta\right)}{f\left(y_{t-1} \mid I_{t-2}, \Theta\right)} \\
&= f\left(y_t \mid r_t = j, \Theta\right) \frac{\sum_{i=1}^{2} p_{ij} f\left(r_{t-1} = i, y_{t-1} \mid I_{t-2}, \Theta\right)}{\sum_{k=1}^{2} f\left(r_{t-1} = k, y_{t-1} \mid I_{t-2}, \Theta\right)} \\
&= f\left(y_t \mid r_t = j, \Theta\right) \frac{\sum_{i=1}^{2} p_{ij} F_{t-1,i}}{\sum_{k=1}^{2} F_{t-1,k}},
\end{aligned}$$

where

$$F_{1,j} = f\left(r_1 = j, y_1 \mid \Theta\right) = f\left(y_1 \mid r_1 = j, \Theta\right) P\left(r_1 = j \mid \Theta\right) = f\left(y_1 \mid r_i = j\right) \pi_j.$$

Let

$$\mathbb{F}_t = \begin{pmatrix} F_{t,1}, & F_{t,2} \end{pmatrix}, \qquad \mathbb{1} = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix},$$

which implies

$$f\left(y_{t-1} \mid I_{t-2}, \Theta\right) = \sum_{k=1}^{2} F_{t-1,k} = \mathbb{F}_{t-1}\mathbb{1}.$$

We obtain

$$F_{t,j} = f\left(y_t \mid r_t = j, \Theta\right) \frac{\sum_{i=1}^{2} p_{ij} F_{t-1,i}}{\mathbb{F}_{t-1}\mathbb{1}},$$

then

$$
\begin{aligned}
\mathbb{F}_t &= \begin{pmatrix} F_{t,1}, & F_{t,2} \end{pmatrix} \\
&= \begin{pmatrix} f\left(y_t \mid r_t = 1, \Theta\right) \frac{\sum_{i=1}^{2} p_{i1} F_{t-1,i}}{\mathbb{F}_{t-1}\mathbb{1}}, & f\left(y_t \mid r_t = 2, \Theta\right) \frac{\sum_{i=1}^{2} p_{i2} F_{t-1,i}}{\mathbb{F}_{t-1}\mathbb{1}} \end{pmatrix} \\
&= \frac{1}{\mathbb{F}_{t-1}\mathbb{1}} \left(\sum_{i=1}^{2} p_{i1} F_{t-1,i} f\left(y_t \mid r_t = 1, \Theta\right), \ \sum_{i=1}^{2} p_{i2} F_{t-1,i} f\left(y_t \mid r_t = 2, \Theta\right) \right) \\
&= \frac{1}{\mathbb{F}_{t-1}\mathbb{1}} \left(\sum_{i=1}^{2} p_{i1} F_{t-1,i}, \ \sum_{i=1}^{2} p_{i2} F_{t-1,i} \right) \begin{pmatrix} f\left(y_t \mid r_t = 1, \Theta\right) & 0 \\ 0 & f\left(y_t \mid r_t = 2, \Theta\right) \end{pmatrix} \\
&= \frac{1}{\mathbb{F}_{t-1}\mathbb{1}} \begin{pmatrix} F_{t-1,1}, & F_{t-1,2} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} f\left(y_t \mid r_t = 1, \Theta\right) & 0 \\ 0 & f\left(y_t \mid r_t = 2, \Theta\right) \end{pmatrix} \\
&= \frac{\mathbb{F}_{t-1}\mathbb{P}\mathbb{G}_t\left(y_t\right)}{\mathbb{F}_{t-1}\mathbb{1}},
\end{aligned}
$$

where

$$
\mathbb{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \quad \mathbb{G}_t\left(y_t\right) = \begin{pmatrix} f\left(y_t \mid r_t = 1, \Theta\right) & 0 \\ 0 & f\left(y_t \mid r_t = 2, \Theta\right) \end{pmatrix},
$$

and

$$
\mathbb{F}_1 = \begin{pmatrix} F_{1,1}, & F_{1,2} \end{pmatrix} = \begin{pmatrix} f\left(y_1 \mid r_i = 1\right) \pi_1 & f\left(y_1 \mid r_i = 2\right) \pi_2 \end{pmatrix} = \boldsymbol{\pi} \mathbb{P} \mathbb{G}_1\left(y_1\right),
$$

with

$$
\boldsymbol{\pi} = \begin{pmatrix} \pi_1, & \pi_2 \end{pmatrix},
$$

which implies

$$
\mathbb{F}_t = \frac{\boldsymbol{\pi} \left(\prod_{t=1}^{T} \mathbb{P}\mathbb{G}_t\left(y_t\right) \right)}{\prod_{t=1}^{T-1} \left(\mathbb{F}_t \mathbb{1}\right)}.
$$

Therefore

$$\widehat{\Theta} = \arg\max_{\Theta} \prod_{t=1}^{T} f\left(y_t \mid I_{t-1}, \Theta\right) = \arg\max_{\Theta} \prod_{t=1}^{T} \left(\mathbb{F}_t \mathbb{1}\right) = \arg\max_{\Theta} \left\{ \boldsymbol{\pi} \left(\prod_{t=1}^{T} \mathbb{P}\mathbb{G}_t\left(y_t\right) \right) \mathbb{1} \right\}.$$

Recall that Guo *et al.* (2011) use maximum likelihood estimation (MLE) with Normal distribution, it can be seen that the maximum likelihood estimator $\widehat{\Theta}$ in Guo *et al.* (2011) is a special case of Equation (4.14). In rest of this chapter, we use Huber's ρ-function (ρ_h), which is defined by

$$\rho_h(x) = \begin{cases} x^2, & \text{if } |x| \leq k, \\ 2k|x| - k^2, & \text{if } |x| > k. \end{cases} \tag{4.17}$$

where $k = 1.345$. This is equivalent to use Huber's least favorable distribution for $f\left(y_t \mid r_t, \Theta\right)$. In next subsection, we conduct simulation studies to investigate the performance of maximum likelihood estimation (MLE) using Huber's least favorable distribution.

4.1.2 Simulation Study

We perform three different simulation studies. Each simulation runs 1,000 times. The mean of estimated parameters is summarized. The simulated time series data are generated using parameters $\mu_1 = -1, \mu_2 = 0.5, \sigma_1 = 3, \sigma_2 = 1, p_{11} = 0.95$ and $p_{22} = 0.95$ for 20-year period. Data generating process (DGP) is described in section (4.3.1). The details are given below:

- Study 1: Maximum likelihood estimator $\widehat{\Theta}$ for a regime-switching normal time series without outliers. Table 4.1 displays that the maximum likelihood estimator $\widehat{\Theta}$ in Guo *et al.* (2011) by using Normal distribution and maximum likelihood estimator $\widehat{\Theta}$ in our method by using Huber's least favorable distribution are similar and closed to true values of the parameters.

Table 4.1: Regime-switching normal time series without outliers

Θ	$\widehat{\Theta}$	MLE with Normal distribution	MLE with Huber's least favorable distribution
-1	$\hat{\mu}_1$	-1.0131	-1.0538
0.5	$\hat{\mu}_2$	0.5020	0.5098
3	$\hat{\sigma}_1$	2.9787	2.5838
1	$\hat{\sigma}_2$	0.9920	0.8920
0.95	$\hat{p}_{11}$	0.9269	0.9316
0.95	$\hat{p}_{22}$	0.9548	0.9581
0.50	$\hat{\pi}_1$	0.3863	0.3836
0.50	$\hat{\pi}_2$	0.6137	0.6164

- Study 2: Maximum likelihood estimators $\widehat{\Theta}$ for a regime-switching normal time series with one random outlier. Table 4.2 displays that the maximum likelihood estimator $\widehat{\Theta}$ in Guo *et al.* (2011) by using Normal distribution is impacted by outliers, while the maximum likelihood estimator $\widehat{\Theta}$ in our method by using

Huber's least favorable distribution is stable and closed to true values of the parameters.

Table 4.2: Regime-switching normal time series with one random outlier

Θ	$\widehat{\Theta}$	MLE with Normal distribution	MLE with Huber's least favorable distribution
-1	$\hat{\mu}_1$	-2.0452	-1.2233
0.5	$\hat{\mu}_2$	0.8623	0.5002
3	$\hat{\sigma}_1$	9.8801	3.4536
1	$\hat{\sigma}_2$	4.4043	1.1053
0.95	$\hat{p}_{11}$	0.7331	0.9159
0.95	$\hat{p}_{22}$	0.9525	0.9529
0.50	$\hat{\pi}_1$	0.3049	0.3852
0.50	$\hat{\pi}_2$	0.6951	0.6148

- Study 3: Maximum likelihood estimator $\widehat{\Theta}$ for a regime-switching heavy-tailed (T-3) time series. Table 4.3 displays that the maximum likelihood estimator $\widehat{\Theta}$ in our method by using Huber's least favorable distribution is still stable and closed to true values of the parameters.

The simulation studies demonstrate that maximum likelihood estimator $\widehat{\Theta}$ in our method by using Huber's least favorable distribution is stable when either the empirical distribution is heavy-tailed or observed data contain outliers.

Table 4.3: Regime-switching heavy-tailed (T-3) time series

Θ	$\widehat{\Theta}$	MLE with Normal distribution	MLE with Huber's least favorable distribution
-1	$\hat{\mu}_1$	-1.2807	-1.0595
0.5	$\hat{\mu}_2$	0.4469	0.4789
3	$\hat{\sigma}_1$	5.6937	3.8169
1	$\hat{\sigma}_2$	2.1418	1.2104
0.95	$\hat{p}_{11}$	0.7352	0.8825
0.95	$\hat{p}_{22}$	0.8953	0.9326
0.50	$\hat{\pi}_1$	0.3415	0.3830
0.50	$\hat{\pi}_2$	0.6585	0.6170

4.2 Change-Point Detection Algorithm

Guo *et al.* (2011) define change-points as the time points that segment a time series if data in two neighboring segments are modeled by the same model with different parameters or different models. They also propose a change-point detection algorithm to identify the change-points. Let $\widehat{\Theta}_T$ be the maximum likelihood estimate of Θ based on a sample indexed from 1 to T. $\widehat{\Theta}_{T+1}$ is expected to be different with $\widehat{\Theta}_T$ if y_{T+1} is from a different market phase. Therefore, a significant change in state-dependent parameters would indicate a change-point in the time series. Guo *et al.*

(2011) use a two-regime Markov regime-switching log-normal (RSLN) model to fit state-dependent parameters of the time series. However, their approach faces some challenges when either the empirical distribution is heavy-trailed or observed data contain outliers. To rectify this problem and improve the change-point detection algorithm, we use a two-regime generalized RS model to fit state-dependent parameters of the time series. The main difference is that state-dependent parameters of the index data are estimated by maximum likelihood estimation (MLE) with Huber's least favorable distribution not with Normal distribution. The parameter of regime 1, i.e., $\mu_1\pi_1$, is still used to identify possible change-points as model parameter estimation for regime 1 changes more dramatically than those for regime 2. We claim there is a change-point if the changes in parameter estimates are significant.

We provide detailed description of the change-point detection algorithm using the generalized RS model. Suppose we have a time series $\{y_1, \ldots, y_n\}$, where n is the length of the time series. Assume k-th $(1 < k < n)$ change occurs at time t_k, which means y_{t_k} is the k-th change point of the entire series. We use MLE with Huber's least favorable distribution to estimate state-dependent parameters based on $\{y_{t_k}, \ldots, y_{\min\{t_k+c-1,n\}}\}$, where c is a pre-defined constant according to the nature of data. By using the numerical values of estimated state-dependent parameters, we simulate a regime-switching process $\{\tilde{y}_{k_1}, \ldots, \tilde{y}_{k_c}\}$. We combine $\{\tilde{y}_{k_1}, \ldots, \tilde{y}_{k_c}\}$

with $\{y_{t_k}, \ldots, y_n\}$ to generate a new series $\{\tilde{y}_{k_1}, \ldots, \tilde{y}_{k_c}, y_{t_k}, \ldots, y_n\}$, denoted as $\{y^*_{t_k-c}, \ldots, y^*_n\}$. For each time $j \geq 1$, we can calculate estimated $\hat{\mu}_1^{(j)}\hat{\pi}_1^{(j)}$, denoted as $\hat{\tau}_j$, respectively, using MLE with Huber's least favorable distribution on series $\{y^*_{t_k-c}, \ldots, y^*_{t_k+j-1}\}$. If $\hat{\tau}_j$ is the first estimation with significant change compared to its previous ones $\hat{\tau}_1, \ldots, \hat{\tau}_{j-1}$, then t_{k+1} is equal to $t_k + j - 1$, and $y_{t_{k+1}}$ is the $(k+1)$-th change-point. We implement following criteria to detect the significant change of $\hat{\tau}_j$. First, based on $\{\hat{\tau}_{(\max\{1,(j-m)\})}, \ldots, \hat{\tau}_{(j-1)}\}$ to calculate average $\bar{\tau}$ and standard deviation σ_τ, where m is a memory duration determined from the feature of data. If $\hat{\tau}_j$ falls outside a standard deviations of $\bar{\tau}$ and passes a pre-chosen constant grace period ν, i.e.,

$$\hat{\tau}_j \notin [\bar{\tau} - a\sigma_\tau, \bar{\tau} + a\sigma_\tau] \ \& \ j > \nu \tag{4.18}$$

then $t_{k+1} = t_k + j - 1$ is a change-point. The change-point detection algorithm is given as follows:

Step 1. Set $j = 1, k = 0, t_0 = 0$.

Step 2. Set $i = t_k + 1$.

Step 3. Use MLE with Huber's least favorable distribution to estimate state-dependent parameters from $\{y_i, \ldots, y_{\min\{i+c-1,n\}}\}$.

Step 4. Simulate a regime-switching process $\{\tilde{y}_{k_1}, \ldots, \tilde{y}_{k_c}\}$ using the estimated state-dependent parameters in step 3.

Step 5. Use MLE with Huber's least favorable distribution to estimate parameters from $\{\tilde{y}_{k_1}, \ldots, \tilde{y}_{k_c}, y_i, \ldots, y_{i+j-1}\}$, and calculate $\hat{\tau}_j$, $\bar{\tau}$ and σ_τ.

Step 6. If $\hat{\tau}_j \notin [\bar{\tau} - a\sigma_\tau, \bar{\tau} + a\sigma_\tau]$ & $j > \nu$, then $k = k+1$, $t_k = i + j - 1$, go back to step 2; otherwise, set $j = j + 1$, go back to step 5.

Step 7. If $j > n - i + 1$, stop.

4.3 Simulation Studies

We conduct simulation studies to investigate the performance of the change-point detection algorithm using the generalized RS model. We first introduce a data generating process, and then we present simulation study results.

4.3.1 Data Generating Process

The following pseudo codes describe a two-state regime-switching normal time series $\{\tilde{y}_1, \ldots, \tilde{y}_T\}$ data generating process (DGP). Similar procedure can be implemented for a two-state regime-switching heavy-tailed (T-3) time series. Let $\{u_1, \ldots, u_T\}$ represent a random uniform series, and use series $\{r_1, \ldots, r_T\}$ to record the regime status at each time t. Then *(i).* At time $t = 1$,

If $u_1 \leq \pi_1$, then $r_1 = 1$ and generate a random normal $\tilde{y}_1 \sim N(\mu_1, \sigma_1)$;

Otherwise, then $r_1 = 2$ and generate a random normal $\tilde{y}_1 \sim N(\mu_2, \sigma_2)$.

(ii). At time $t = 2, \ldots, T$,

If $r_{t-1} = 1$, then

If $u_t \leq p_{11}$, then $r_t = 1$ and generate a random normal $\tilde{y}_t \sim N(\mu_1, \sigma_1)$,

Otherwise, then $r_t = 2$ and generate a random normal $\tilde{y}_t \sim N(\mu_2, \sigma_2)$;

Otherwise, then

If $u_t \leq p_{21}$, then $r_t = 1$ and generate a random normal $\tilde{y}_t \sim N(\mu_1, \sigma_1)$,

Otherwise, then $r_t = 2$ and generate a random normal $\tilde{y}_t \sim N(\mu_2, \sigma_2)$.

4.3.2 Simulation Study Results

We perform two simulation studies. Each simulation runs 1,000 times. In simulation study 1, we generate two regime-switching normal time series $\{\tilde{y}_1, \ldots, \tilde{y}_{60}\}$ using parameters $\mu_1 = -2, \mu_2 = 4, \sigma_1 = 6, \sigma_2 = 5, p_{11} = 0.75$ and $p_{22} = 0.95$ and $\{\tilde{y}_{61}, \ldots, \tilde{y}_{120}\}$ using parameters $\mu_1 = -8, \mu_2 = 4, \sigma_1 = 6, \sigma_2 = 5, p_{11} = 0.85$ and $p_{22} = 0.85$. Then we combine the two time series to obtain $Y_1 = \{\tilde{y}_1, \ldots, \tilde{y}_{120}\}$. In this setting, no outlier in data and data contain a change-point at 61th. We apply the algorithm to time series Y_1. In simulation study 2, we randomly enlarge a value $\tilde{y}_i$ as an outlier in time series Y_1, so that we obtain a new time series Y_2. There is a random outlier in data and data contain a change-point at 61th. We also apply the algorithm to time series Y_2. We define following three cases of performance evaluation by a given detective tolerant constant α:

- Accurate-Detection (AD): Detect only one change-point, and the detected change-point is in range of $[61 - \alpha, 61 + \alpha]$. This means that the algorithm can accurately detect change-point.

- Over-Detection (OD): Detect more than one change-point, but at least one change-point is in range of $[61 - \alpha, 61 + \alpha]$. This means that the algorithm can accurately detect change-point but also with some false detections.

- No-Detection (ND): Otherwise. This means that the algorithm can not accurately detect change-point.

The performance summary of change-point detection algorithm of simulation studies (1,000 runs with $\alpha = 5$) is provided in table 4.4 below.

From the simulation results, we observe: (1) When time series does not contain outliers, both algorithms have reasonable detective ability because accurate-detection results are over 70% and no-detection results are less than 5%. The algorithm using

Table 4.4: Performance summary of change-point detection algorithm of simulation studies (1,000 runs with $\alpha = 5$)

Methods	MLE with Normal distribution			MLE with Huber's least favorable distribution		
Detection %	Accurate-Detection	Over-Detection	No-Detection	Accurate-Detection	Over-Detection	No-Detection
Simulation 1	72%	24%	4%	79%	19%	2%
Simulation 2	44%	34%	22%	69%	23%	8%

MLE with Huber's least favorable distribution performs slightly better (+7% for accurate-detection); (2) When time series contains outliers, the algorithm using MLE with Huber's least favorable distribution has better performance. The accurate-detection results are closed to 70% and no-detection results are still less than 10%. The algorithm using MLE with Normal distribution is largely impacted by outliers. Notice that the performance slightly decrease compared with simulation 1 results. A more robust ρ-function might solve this observation. This topic might be considered as future research.

4.4 Real Market Time Series Application

Let $\{P_t\}$ be the monthly price time series and let $Y_t = 100 \times \log(P_t/P_{t-1})$ be the monthly log-return time series. We apply the change-point detection algorithm to Hang Seng monthly index data from January 1988 to March 2015.

The time series sample must contain a sufficient number of observations in order to ensure a good estimation of model parameters. Therefore, we set $c = 120$ and $m = 12$. $m = 12$ indicates the memory period is one year and $c = 120$ indicates the sample for parameter estimation is 10-year period. We set a grace period $\nu = 12$ to reflect that the market needs long time to move between different phase, i.e., bear market and bull market. We also set $a = 3$ for Equation (4.18) due to control chart method.

The change-points identified by the algorithm using MLE with Huber's least favorable distribution compared with the change-points identified by the algorithm using MLE with Normal distribution are shown in table 4.5.

Table 4.5: Summary of detected change-points on Hang Seng monthly index data

Estimation Method	Detected change-points
MLE with Normal distribution	18, 33, 85, 116, 238, 285
MLE with Huber's least favorable distribution	31, 60, 118, 188, 238, 285

We visualize the detected change-points on Hang Seng monthly index price and log-return time series in figure 4.1. We also visualize the estimated expected returns of state 1, i.e., $\hat{\tau}_j$, with boundary in figure 4.2.

We note that for Hang Seng monthly index price time series, both algorithms

Figure 4.1: Detected change-points on Hang Seng price and log-return time series

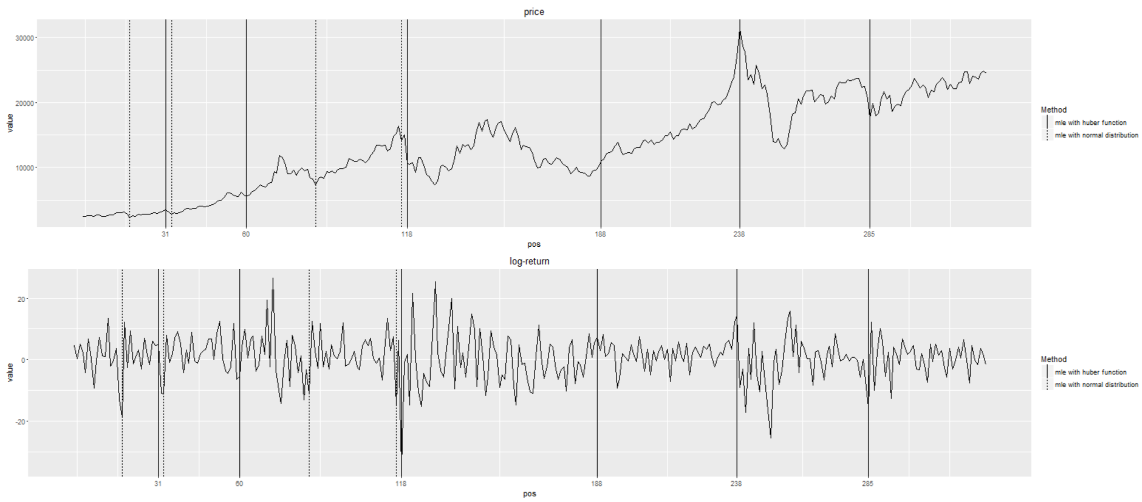
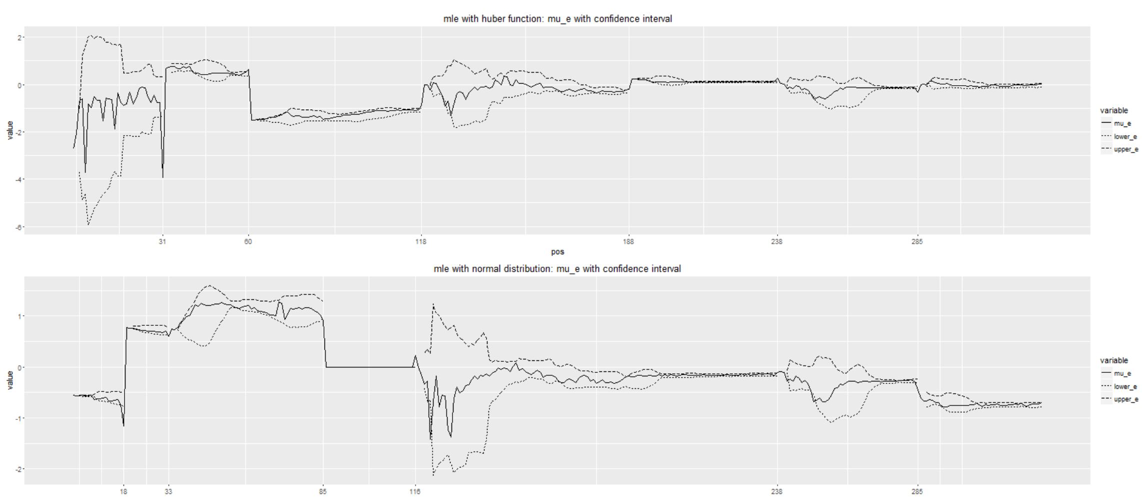


Figure 4.2: Estimated expected returns of state 1 for detected change-points



can identify four change-points at $31(33), 116(118), 238$ and 285. The change-point detection algorithm using MLE with Huber's least favorable distribution detects two change-points at 60 and 188 while the algorithm using MLE with Normal distribution detects two change-points at 18 and 85. Visually, there is small trend change around time 18 but relatively large trend changes around 188. The change-points detected by the algorithm using MLE with Huber's least favorable distribution are described as follows:

- $\hat{\tau}_1 = 31$, which corresponds to June 1990. After "Black Monday" on October 19, 1987, many countries including Hong Kong experienced economic growth. The Hang Seng index entered a slow bull market. This phase lasted 6 years, and in total increased 565% with average annual increased 37%.

- $\hat{\tau}_2 = 60$, which corresponds to December 1992. The Hang Seng index started to rapidly increase as economic continued growth. Citic Pacific became the first red chip to join the Hang Seng Index constituent stocks on August 1992. The Hang Seng index reached a new historical highest $(12, 599)$ after 12 months, which was the end of this bull market.

- $\hat{\tau}_3 = 118$, which corresponds to October 1997. The largest one-day historical increase (in percentage terms) was on October 29, 1997. This mainly attributable to hedge fund activities and the rally after the plunge on the previous day. The day before yesterday, the Thai baht was attacked and the hedge funds speculated Hong Kong stocks. The Hang Seng index consecutively reached historical highest on May, June and July and reached a new historical highest with $16,820$ on August, then rapidly decreased to less than $10,000$, which ended this bull market. In the same time, Asian Financial Crisis began and the Asian economic "miracle" crashed.

- $\hat{\tau}_4 = 188$, which corresponds to August 2003. Three month ago, the Hang Seng index decreased to the lowest with $8,331$, indicating the end of last bear market. The market entered a new bull market. In general, it is believed that before 2006 Hong Kong Hang Seng has a very strong synergy effect with the US stocks except in the crisis of Southeast Asia in 1997. The reason is that the United States is the largest export market in Hong Kong. The period between 2000 and 2002 was considered a major downturn period for US stock market and also for Hong Kong market. The Federal Reserve lowered interest rates six times in a row from January to June 2001 and five times again in the next three years, which stimulate economic recovery from 2003.

- $\hat{\tau}_5 = 238$, which corresponds to October 2007. The Hang Seng index climbed to the most historical highest with $31,352$, indicating an ending bull market. January 2008 started a new worldwide financial crisis caused mainly by the US subprime mortgage crisis.

- $\hat{\tau}_6 = 285$, which corresponds to September 2011. After a relative long up and down period, the market entered a stable increase period for next bull market.

4.5 Conclusion

In this chapter, we propose a generalized Markov regime-switching (generalized RS) model. Our study demonstrates that the generalized RS model can improve the stability of model parameter estimation when either the empirical distribution is heavy-tailed or observed data contain outliers. Financial market behavior changes at some points in time. It is not suitable to use a single regime-switching model to model entire time series. Guo *et al.* (2011) propose a change-point detection algorithm to identify change-points in the time series. We use the generalized RS model to improve the change-point detection algorithm. Simulation studies demonstrate that the change-point detection algorithm using the generalized RS model can improve accuracy of identifying the change-points when either the empirical distribution is heavy-tailed or observed data contain outliers. The real data analysis identified six change-points on Hang Seng index. Corresponding segments match the bear-bull phases observed in market well.

5 Conclusion and Future Work

5.1 Conclusion

In this dissertation, we study statistical models and their applications on financial time series data.

In chapter 2, we study statistical modeling and moving average strategy for single-asset portfolio. We propose a single-asset generalized moving average crossover (S-GMA) strategy. Our study demonstrates that the SGMA strategy can provide more investment options which can solve the well-known problem from the MA strategy. Simulation studies demonstrate that the SGMA strategy can increase both the investor's expected log-utility of wealth and the investor's expected wealth compared with the MA strategy. Two high-frequency ETF real data analysis demonstrate that the SGMA strategy can outperform the MA strategy for both backward investment on training data and forward investment on test data.

In chapter 3, we extend our research to multi-asset portfolio. We propose a multi-asset generalized moving average crossover (MGMA) strategy. Our study demonstrates that the MGMA strategy can provide more investment options with the investor's risk tolerance. Simulation studies demonstrate that the MGMA strategy can increase both the investor's expected log-utility of wealth and the investor's expected wealth. Two high-frequency ETF real time examples demonstrate that the MGMA strategy can outperform the MA strategy for both backward investment on training data and forward investment on test data.

In chapter 4, we study the regime-switching model with its application to capture the financial market behaviors. We propose a generalized Markov regime-switching (generalized RS) model. Our study demonstrates that the generalized RS model can improve the stability of the model parameter estimations when either the empirical distribution is heavy-tailed or observed data contain outliers. We then use the generalized RS model to improve the corresponding change-point detection algorithm. Simulation studies demonstrate that the change-point detection algorithm using the generalized RS model can improve accuracy of identifying the change-points when either the empirical distribution is heavy-tailed or observed data contain outliers. The real data analysis identified six change-points on Hang Seng index. Corresponding segments match the bear-bull phases observed in market well.

5.2 Future Work

In the area of moving average trading strategy, there are three possible future research directions. First one is to improve the stability and accuracy of model parameter estimation, especially for multi-asset portfolio. Second one is to develop a more robust criteria to estimate model parameters from training data. Lastly, different utility function can be considered and approximately solutions might be adopted.

In the area of regime-switching model with its application to detect change-points, one possible future research direction is to explore different choice of the ρ functions other than Huber ρ_h function. This might further improve the performance of the corresponding algorithm.

Bibliography

Ang A. and Bekaert G. (2006). Stock Return Predictability: Is it There?. *The Review of Financial Studies*, **20(3)**, 651–707.

Bauwens L., Dufays A. and Rombouts J.V.K (2014). Marginal likelihood for markov-switching and change-point GARCH models. *Journal of Econometrics*, **178**, 508–522.

Buraschi A., Porchia P. and Trojani F. (2010). Correlation Risk and Optimal Portfolio Choice. *The Journal of Finance*, **LXV1**, 393–420.

Brock W., Lakonishok J. and LeBaron B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, **47(5)**, 1731–1764.

Campbell J.Y. and Thompson S.B. (2008). Predicting the equity premium out of sample: Can anything beat the historic alaverage?. *The Review of Financial Studies*, **21(4)**, 1509–1531.

Cespa G. and Vives X. (2012). Dynamic trading and asset prices: keynes vs. hayek. *Review of Economic Studies*, **79**, 539–580.

Chang E.J., Lima E.J.A. and Tabak B.M. (2004). Testing for predictability in emerging equity markets. *Emerging Markets Review*, **5(3)**, 295–316.

Chen, J.H., Huang Y. and Wang P.M. (2014) Composite likelihood under hidden markov model. *Statistica Sinica*.

Chiu, M.C. and Wong H.Y. (2014) Mean-variance portfolio selection with correlation risk. *Journal of Computational and Applied Mathematics*, **263**, 432–444.

Garlappi L. and Skoulakis G. (2009). Numerical solutions to dynamic portfolio problems: the case for value function iteration using taylor approximation. *Computational Economics*, **33(2)**, 193–207.

Guerin P. and Marcellino M. (2013). Markov-switching MIDAS models. *Journal of Business and Economic Statistics*, **31(1)**, 45–56.

Gunasekarage A. and Power D.M (2001). The profitability of moving average trading rules in south asian stock markets. *Emerging Markets Review*, **2(1)**, 17–33.

Guo B.B., Wu Y.H., Xie H. and Miao B.Q. (2011). A segmented regime-switching model with its application to stock market indices. *Journal of Applied Statistics*, **38(10)**, 2241–2252.

Hamilton J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57(2)**, 357–384.

Hamilton J.D and Susmel R. (1994). Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics*, **64(1-2)**, 307–333.

Hardy M.R (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, **5(2)**, 41–53.

Huang L. and Liu H. (2007). Rational in attention and portfolio selection. *Journal of Finance*, **62**, 1999–2040.

Ito A. (1999). Profits on technical trading rules and time-varying expected returns: evidence from pacific-basin equity markets. *Pacific-Basin Finance Journal*, **7(3-4)**, 282–330.

Keim D.B. and Stambaugh R.F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, **17(2)**, 357–390.

Legendre F. and Togola D. (2016). Explicit solutions to dynamic portfolio choice problems: A continuous-time detour. *Economic Modelling*, **58**, 627–641.

Lioui A. and Poncet P. (2016). Understanding dynamic mean variance asset allocaiton. *European Journal of Operational Research*, **254**, 320–337.

Lo A. and MacKinlay C. (1999). A Non-Random Walk Down Wall Street. *Princeton University Press, Princeton, NJ.*

Lo A., Mamaysky H. and Wang J.(2000). Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, **55(4)**, 1705–1765.

Mohr E., Schmidt G. and Jansen S. (2013). A comparative study of heuristic conversion algorithms, genetic programming and return predictability on the german market. *EVOLVE - A Bridge between Probability*, **447**, 393–414.

Neely A., Gregory M. and Platts K. (2014). Performance measurement system design. *International Journal of Operations and Production Management*, **25(12)**, 1228–1263.

Stambaugh,R.(1999). Predictive regressions. *Journal of Financial Economics*, **54**, 375–421.

Tabak B.M. and Lima E.J.A (2009). Market efficiency of brazilian exchange rate: evidence from variance ratio statistics and technical trading rules. *European Journal of Operational Research*, **194(3)**, 814–820.

Wu H., Ma C. and Yue S. (2017). Momentum in strategic asset allocation. *International Review of Economics and Finance*, **47**, 115–127.

Zhu Y.Z. and Zhou G.F. (2009). Technical analysis: an asset allocation perspective on the use of moving averages. *Journal of Financial Economics*, **92**, 519–544.

Zou W. and Chen J.H. (2013). Markov regime-switching model in crude oil market - comparison of composite likelihood and full likelihood. *Canadian Journal of Statistics*, **41(2)**, 353–367.

A Appendix

A.1 Matrix Exponential

We present the matrix exponential definition and properties in this section.

Let $\mathbb{A}$ is an $n \times n$ real or complex matrix. The exponential of matrix $\mathbb{A}$ denoted by $e^{\mathbb{A}}$ is also the $n \times n$ matrix given by the power series

$$e^{\mathbb{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbb{A}^k = \mathbb{I}_n + \mathbb{A} + \frac{1}{2!}\mathbb{A}^2 + \dots,$$

where $\mathbb{A}^0$ is defined to be the identity matrix $\mathbb{I}_n$ with the same dimension as matrix $\mathbb{A}$. The above series always converges, so the exponential of matrix $\mathbb{A}$ is well-defined. If $\mathbb{A}$ is a 1×1 matrix, the matrix exponential of $\mathbb{A}$ is a 1×1 matrix whose single element is the ordinary exponential of the single element of $\mathbb{A}$.

The following properties are used in the derivation of the Lemmas and Propositions:

- $e^0 = \mathbb{I}$

- $e^{a\mathbb{A}}e^{b\mathbb{A}} = e^{(a+b)\mathbb{A}}$
- $e^{\mathbb{A}}e^{-\mathbb{A}} = \mathbb{I}$
- If $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$, then $e^{\mathbb{A}}e^{\mathbb{B}} = e^{\mathbb{B}}e^{\mathbb{A}} = e^{(\mathbb{A}+\mathbb{B})}$
- If $\mathbb{B}$ is invertible, then $e^{\mathbb{B}\mathbb{A}\mathbb{B}^{-1}} = \mathbb{B}e^{\mathbb{A}}\mathbb{B}^{-1}$
- $e^{\left(\mathbb{A}^T\right)} = \left(e^{\mathbb{A}}\right)^T$, where $\mathbb{A}^T$ denotes the transpose of $\mathbb{A}$. It follows that if $\mathbb{A}$ is symmetric, then $e^{\mathbb{A}}$ is also symmetric
- $\mathrm{d}\left(e^{t\mathbb{A}}\right) = \mathbb{A}e^{t\mathbb{A}}\mathrm{d}t$ where $\mathbb{A}$ is a constant matrix.