Equilibrium strategies for Mean-variance problem

Zeyu He

School of Mathematics

University of Leeds

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

February 2018

Acknowledgements

I owe special thanks to my supervisors: Dr. Jan Palczewski, Dr. Georgios Aivaliotis and Prof. Alexander Yu. Veretennikov, for their intellectual support and for giving me the motivation I needed to complete this thesis. In fact, without their encouragement, I would not have contemplated embarking upon a PhD. To them, I am eternally grateful.

I would like to express my thanks to my colleagues and staffs at the School of Mathematics, with special thanks to Jeanne Shuttleworth. Of course I would not have been able to undertake this PhD if it were not for the financial support I received from the University of Leeds. Thus, to them, I am also extremely grateful.

I would also like to thank my parents, Wei Wang and Anthony George Raymond Green, who have supported me throughout my time in University. A big thank you goes to my fiancée, Handan Lu, for putting up with me over the last few months and for making me all those dinners so that I could keep on working.

Last but not least, I would like to thank Dr James Fung who, throughout this process, have provided me with much needed *fun time*.

Abstract

This research is devoted to study equilibrium strategies in a game theoretical framework for the mean-variance problem. The thesis explores the investment behaviour and interlinks between different types of equilibrium strategies.

In order to find the open-loop strategy in discrete time, we incorporate the idea based on Hu *et al.* (2012) and the concept of open-loop strategies in engineering study. In engineering study, there are two types of strategies: open-loop and closed-loop control strategies. We find the interpretations for both strategies in a Nash equilibrium context from a financial perspective. This thesis extends the literature by providing the existence and uniqueness of the solution of open-loop equilibrium strategy in discrete time. Our findings point to the causes of different equilibrium strategies in the existing literature.

We show the common issue of equilibrium strategies, i.e. that the amount of money invested in risky assets decays to 0 as time moves away from the maturity. Furthermore, the closed-loop strategy tends to a negative limit depending on the assets' Sharpe ratio. We call this phenomenon as *Mean-variance puzzle*. The reason is that the variance term penalises the wealth changes quadratically as well as the expectation only increases linearly. By drawing in the concepts from behavioural economics, we solve this puzzle by using the present-biased preference. The advantage of the present-biased preference is that equilibrium investors have the flexibility to adjust their risk attitude based on their anticipated future.

We simulate three types of control strategies existing in the literature and compare the investment performance. Furthermore, we evaluate the performance with respect to different rebalancing periods.

Abbreviations

M^T :	transpose of a matrix M.
$\mathbf{b}_t$:	a column vector of return factor for n risky assets $[b_t^1, \ldots, b_t^n]^T$ where $b_t^i = \frac{S_t^i}{S_{t-1}^i}$.
$\mathbf{B}_t$:	a column vector of excess return factor for n risky assets $[(b_{t+1}^1 - e^{r_t^0}), (b_{t+1}^2 - e^{r_t^0}), \ldots, (b_{t+1}^n - e^{r_t^0})]^T$ at time t.
$\mathcal{A}$:	the set of admissible strategies.
$o(\varepsilon^k)$:	a family of ξ^ε such that $\lim_{\varepsilon\to 0} \operatorname{ess\,sup} \left\lvert\frac{\xi^\varepsilon}{\varepsilon^k}\right\rvert = 0$.
$O(\varepsilon^k)$:	a family of ξ^ε such that $\lim_{\varepsilon\to 0} \operatorname{ess\,sup} \left\lvert\frac{\xi^\varepsilon}{\varepsilon^k}\right\rvert < c$ for some finite constant c.
$\frac{\mathcal{F}_s \otimes \mathbb{B}(\mathbb{R})}{\mathbb{B}(\mathbb{R}^n)}$:	a set of mearuable functions $f : (X, Y) \to Z$, where $(X, \mathcal{F}_s)$, $(Y, \mathbb{B}(\mathbb{R}))$ and $(Z, \mathbb{B}(\mathbb{R}^n))$ are measurable spaces.
$\mathcal{L}^2(\mathcal{F}_t)$:	the set of $\mathcal{F}_t$-measurable function with $\mathbb{E}[\lvert f\rvert^2] < \infty$.
$\mathcal{L}_p$:	the Banach space of Borel measurable functions summable with order p.
$\mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^n)$:	the set of $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes $f = \{f_s : t < s < T\}$ with $\mathbb{E}[\int_t^T \lvert f_s\rvert^2\, ds] < \infty$.

$\mathcal{L}^{\infty}_{\mathcal{F}}(t,T;\mathbb{R}^n)$:	the set of essentially bounded $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes.
$\mathcal{L}^{2}_{\mathcal{F}}(\Omega, C(t,T;\mathbb{R}^n))$:	the set of continuous $\{\mathcal{F}_t\}_{s\in[t,T]}$-adapted processes $f = \{f_s : t \leq s \leq T\}$ with $\mathbb{E}[\sup_{s\in[t,T]} \lvert f_s\rvert^2] \leq \infty$.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The modern portfolio selection theory dates back to Markowitz's Nobel prize-winning work (Markowitz, 1952). Before the 1950s, there was little systematic theory for the financial market. Markowitz (1952) worked in single-period mean-variance portfolio selection problem, which became the foundation for modern financial portfolio theory. The first pioneering contribution made by Markowitz is to discuss the trade-off between return and risk of the portfolio, where the risk is measured by the variance of the portfolio return. A second significant contribution is that Markowitz's model provides the foundation for developing the Capital Asset Pricing Model, which is introduced by Sharpe (1964) during the 1960s, where the Beta risk in CAPM is now used widely in all the stock exchange-listed companies.

In the single period of Markowitz's framework, he identifies the return with the expectation of the portfolio and the risk with the variance of the portfolio. As the trade-off between return and risk, the problem focuses on how to minimise portfolio's variance subject to a given level of return. Then, the strategy is said to be *optimal* if the portfolio achieves minimum variance subject to a prescribed expected level of return X. The set of all optimal pairs $(\mathbb{E}[X], \mathrm{Var}(X))$ is called *Mean-Variance frontier*. Following the Markovitz's work, Merton (1971) extends his result to the case in the single-period when short selling is allowed.

Ever since the appearance of the Markovitz's framework, this subject has attracted a huge amount of researches on itself. However, the investor faces several practical difficulties in multi-period time as the tradeoff information between return and risk becomes less intuitive. The multi-period mean-variance problem has not been further investigated until the results from the paper Li & Ng (2000). The strategy developed by Li & Ng (2000) is referred to pre-commitment strategy because the investor pre-commits to a strategy which is only optimal at the time of evaluation.

However, the pre-commitment strategy has its own disadvantages. Under the pre-commitment strategy, the investor will find a different optimal strategy if he evaluates his plan at a different time. In other words, if the investor is keen to follow the best strategy at all times, he has to re-evaluate his plan continuously to ensure that he is making the best decision. In modern financial market, there are thousands of different stocks available in the market. Furthermore, if an investor decides to follow one strategy and wants to be satisfied at any time, then pre-commitment strategy does not meet his requirement. Therefore, the pre-commitment strategy can be very computationally expensive.

In game theory, the Nash equilibrium is a concept related to a non-cooperative game among two and more players. The players in the game are assumed to know strategies of the other players. Every player in the game makes the best decision for himself, based on what he expects the others will do. For example, suppose two cars are driving towards a junction from perpendicular directions. The traffic is red for car A and green for car B. The question is whether they want to break the law if there is no police officer. As car B knows that car A will stop in front of traffic light, then car B will make the decision that he will drive through this junction. If not, both cars would not move. On the other side, as car A knows that car B will drive through the junction, then the best decision for A is to wait before the junction. Otherwise, they will crash into each other. Therefore, the traffic light can be regarded as a law of Nash equilibrium in that both players are willing to follow it.

In game theory context, the time-consistent mean-variance investor is playing the optimisation game with future-self. The investor is finding the best responding strategy based on the strategies which are expected to be chosen by his future-selves. Strotz (1955) suggested a time-consistent strategy should take into account investor's own belief or insight. Such *own belief* coincides the idea of anticipating future strategy in the game theoretic framework. Therefore, the Nash subgame equilibrium approach has been proposed by Bjork & Murgoci (2010) and Basak & Chabakauri (2010) in order to obtain a time-consistent strategy.

Suppose the equilibrium investor at time t is looking for the best strategy sequence $(u_t, u_{t+1}, \ldots, u_{T-1})$. He has to pre-determine his future strategies $(u_{t+1}, \ldots, u_{T-1})$. However, the key difference between the open-loop and closed-loop equilibrium strategies is: the open-loop investors consider their future strategies at time $(u_{t+1}, \ldots, u_{T-1})$ as a fix strategy (i.e. a fixed amount of money invested in the stock) whereas the open-loop investors modify the future strategies according to the different choice of strategy u_t at time t. These modified future strategy, denote as $(u^m_{t+1}, \ldots, u^m_{T-1})$, are only used to evaluate the strategy u_t at time t. Once the strategy u_t is chosen, then the investor combine u_t with the pre-determined future strategies $(u_{t+1}, \ldots, u_{T-1})$ to give a closed-loop equilibrium strategy. Therefore, the open-loop equilibrium investor can be viewed as a *fixed recipe* investor whereas the closed-loop equilibrium investor is a *flexible recipe* investor.

Stochastic control is a sub-field of control theory, which plays an important role in economic problems. Stochastic control was developed in the late 1950s and the early 1960s. During that period, stochastic control problems were formulated to solve engineering problems, typically, which involve a dynamical system in the presence of randomness. For instance, the dynamics of the system can be described as a stochastic differential equation and the randomness described as Brownian motion. The engineers can control the direction of dynamics and the size of the randomness in the system. In addition, the performance of control system is assessed by a function of future value of a stochastic process. Therefore, the stochastic control problem is to find a control law which optimises the random quantity (see Fleming & Rishel (2012)).

In engineering study, a manager can adjust different elements which are connected to the system to produce the desired output. This system is referred to control system. The control systems are classified as open loop or closed loop. The closed loop system is also known as feedback control system. A system is called an open loop control system if the control action is independent of the output of the system. Alternatively, if a system is called closed loop system if a manager determines the input of system based on the responses (or output) generating from the previous input quantity.

In both one period and multi-period models, the strategy is said to be *optimal* if this strategy gives the highest rewards at the end. The main approach in solving such optimisation problems is to derive a set of necessary conditions that must be satisfied by any optimal control. Suppose that the investor needs to maximise his objectives, if the investor decides to deviate from the optimal strategy by a quantity ε, then the functional rewards will be worse off than if he had not changed. The deviation quantity ε is called *perturbation*. In the classical calculus of variations, the optimisation problem can be solved by considering the first-order condition with respect to this perturbation. However, in the equilibrium context, the investor is not only to find a strategy that maximises the rewards, but also he has to be satisfied with this strategy at any time.

This thesis has three purposes: firstly, to get insights for the equilibrium control strategies by looking at discrete time problems in detail. In order to show optimality in the equilibrium setting, we need to use a perturbation and investigate the effect of different types of perturbations to equilibrium strategies. Generally speaking, if we obtain the equilibrium strategies that turn out to be different, we can confirm the fact that the perturbation plays an important role of determining the equilibrium strategy. Later, we extend our study into continuous setting by looking at a random perturbation process. Secondly, this thesis aims to look at how open-loop strategy compares to the closed-loop strategy and also the pre-commitment strategy. Following the result from previous motivation, we can find a family of different equilibrium controls by setting the perturbation differently. In practice, it is important to see if there exists any equilibrium control that will out-perform the rest. Third, this thesis examines whether equilibrium strategy can be a better solution than the pre-commitment as the pre-commitment strategy is widely used in financial investment.

1.2 Research questions

Research question 1

Providing the open-loop equilibrium strategy in continuous time setting, this study explores a discrete version of open-loop equilibrium, and obtain the necessary and sufficient condition for discrete open-loop strategies in the form of a uniqueness and existence theorem.

Research question 2

By using a similar construction technique for the closed-loop equilibrium strategies, this study examines the consistency of closed-loop equilibrium with the definition introduced by Bjork & Murgoci (2010). This is to understand and provide an interpretation of equilibrium strategies and different types of perturbations from a financial prospective.

Research question 3

The simulated equilibrium strategies in the existing literature show a down-trend in the investment proportion as time moves away from the maturity. Do equilibrium strategies converge to any limit as time moves backwards? The decay in equilibrium strategies makes no practical sense for the independent and identical investment environment. Therefore, is it possible to solve this issue?

Research question 4

Compared to pre-commitment strategies, we would like to find whether the outcomes of equilibrium strategies for the long-term investments have been improved or not. The comparison will be tested on the basis of the situation whether the re-evaluation is allowed. Furthermore, we would like to implement the approach developed from *Research question 3* and assess the improvement of the investment performance.

1.3 Overview of the thesis

This thesis is organised as follows: In Chapter 2, we review the previous work for multi-period mean-variance problem which covers the pre-commitment strategy by Li & Ng (2000), open-loop equilibrium strategy by Hu *et al.* (2012) and closed-loop equilibrium strategy by Bjork & Murgoci (2010). Chapter 3 describes our approach to construct open-loop equilibrium strategies in discrete time. We derive a uniqueness and existence of open-loop strategy. Finally we present an example to show the discrete open-loop equilibrium strategies converge to the continuous strategies. Chapter 4 applies same technique used in Chapter 3 to derive the closed-loop in discrete time. Further, we show the closed-loop equilibrium strategies decay to a negative limit as time goes to infinity. In, Chapter 5, we relax the assumption on the type of perturbation for open-loop equilibrium strategy in continuous time. Chapter 6.2 addresses the issue of decays in equilibrium strategy and uses the present-biased concept from economics to fix the issue. Chapter 7 uses numerical tools to assess the performance of different types of equilibrium strategies. In Chapter 8, we provide a summary of our approach on studying the equilibrium strategies and give a financial interpretation of the rationale of the equilibrium investors.

Chapter 2

Literature review

2.1 Mean-varaince analysis in single period time setting

The central issue in the mean-variance formulation is the optimisation of quadratic criteria. In the literature, there are various formulations of mean-variance problem. The classical mean-variance portfolio selection problem is formulated as:

$$\min_{u} \quad \mathrm{Var}(X_T^u) \tag{P1}$$

$$\text{s.t.} \quad \mathbb{E}(X_T^u) = c, \tag{$*$}$$

$$\sum_i u_i = 1. \tag{$\star$}$$

where $u \in \mathbb{R}^n$ denotes the proportion of investment allocated in ith asset and X denotes the investor's wealth level. Since different assets yield different returns at terminal time, the increment of investor's wealth depends on the way of allocating u for each assets. It can be noticed that we use the original formulation in Markowitz (1952) instead of maximising the expected return subject to the risk.

In this formulation, the investor splits the objective (minimise the portfolio variance) and reward condition (∗). The Lagrangian of such optimisation problem is

$$L(u, c, \lambda_1, \rho_1) = \mathrm{Var}(X_T^u) - \lambda_1(\mathbb{E}[X_T^u] - c) - \rho_1(\sum_i u_i - 1).$$

We refer to the optimal multipliers λ_1 and ρ_1 as reward multiplier and budget multipler, respectively. From the Lagrangian, the first-order necessary condition yields the optimal triple of (u^*, λ^*, ρ^*) subject to a desired level c. On the other hand, the mean-variance problem can also be read as a concave quadratic optimisation problem, which can be formulated as:

$$\begin{aligned} \max \quad & \lambda_2\,\mathbb{E}[X_T^u] - \frac{1}{2}\mathrm{Var}(X_T^u) \\ \text{s.t.} \quad & \sum_i u_i = 1, \end{aligned} \tag{P2}$$

for some constant λ_2. The classic formulation is equivalent to the quadratic optimisation problem if and only if the chosen constant λ_2 is equal to the optimal reward multiplier λ_1. The Markowitz framework provides an important theoretical justification of the parameters λ_1 and λ_2. Markowitz (1968) describes this parameter as risk aversion which represents the rational behaviour or investor's preference under uncertainty. Several measures of risk aversion parameters have been justified and explored by Mossin (1968); Rubinstein (1973); Tobin (1958) for the relation between the utility functions and the corresponding risk attitude. It is generally assumed that the risk aversion parameter is a constant. However, Arrow (1971) argues the constant risk aversion leads the invariance of demand for risky assets as the present wealth changes (i.e. the strategy does not depend on wealth level). Arrow (1971) believes that the risk aversion should decrease as the present wealth increases. This idea will be discussed later in the equilibrium approach.

2.2 Mean-variance analysis in multi-period time setting

Following the formulation from $(P2)$, the extension of objective functional J in the multi-period framework becomes

$$J(t, x_t; u) = \mathbb{E}_{t,x}[X_T^u] - \gamma \mathrm{Var}_{t,x}(X_T^u),$$

where we assume the investor is sitting at time t with wealth level x and λ is postive constant that penalises the risk. The conditional expectation and variance are defined by $\mathbb{E}[X_T^u|X_t = x]$ and $\mathrm{Var}(X_T^u|X_t = x)$, respectively. The main issue in such framework is that the Bellman's principle does not hold and dynamic programming cannot be applied. When the objective functional is linear in terms of conditional expectation, the Markov property can be used to decomposit the original problem into several sub-problems with respect to different time points. Thus, the investor can first optimise the sub-problems and then solve the original problem based on the previously computed solution. The obtained sequential control strategies will form an optimal control law for the optimisation problem. However, due to the term $\big(\mathbb{E}[X]\big)^2$ arising from variance term, there is no global optimal strategy in multi-period that optimises the functional at every point in time.

There are three types of time-inconsistent problem in the literature. First, the mean-variance optimisation problem is time-inconsistent. We recall the variance formula:

$$\mathrm{Var}(x) = \mathbb{E}[X^2] - E^2[X].$$

We have the term $E^2[X]$, which is not an expected value of a non-linear function of wealth, but instead a non-linear function of expected wealth. Second, the objective functional depends on the the wealth of evaluation point. For example, the endogenous utility function has the following form

$$E_{t,x}\big[\ln(X_T^u - x + c)\big], \quad c > 0.$$

The different evaluation time leads to different wealth level x. Since utility function changes, the resulting strategy u will not necessarily be the same. The final time-inconsistent situation is when the utility function depends on the evaluation time. For example,

$$\mathbb{E}_{t,x}\Big[\int_t^t \rho(s-t)f(X_s)ds\Big].$$

Thus, if the optimality of obtained strategy depends on the time of the evaluating the problem, then such problem is called time inconsistent problem. Previous research suggests there are three different methods dealing with the time inconsistent problem.

2.2.1 Pre-commitment strategy

The first type of the strategy is called pre-commitment which is developed by Li & Ng (2000). The investor fixes an initial point and tries to find the control strategy which optimises the objective with respect to this fixed time point. Li & Ng (2000)'s work is a breakthrough as it is the first paper that provides the analytical result in multi-period discrete time setting. They use the embedding technique to embed the problem ($P2$) into an auxiliary problem, which is time-consistent, and obtain the optimal strategies for the auxiliary problem. Furthermore, they show the solution set for the original problem is a subset of the solution for auxiliary problem, and provide a necessary condition for the solution of auxiliary problem attaining the optimum. Based on Li & Ng (2000)'s study, Zhou & Li (2000) introduce the continuous stochastic linear-quadratic framework which is a generalisation of the mean-variance problem. The authors combine the portfolio selection problem and stochastic control theory to expand this topic into a more complicated situation. For example, random parameters (Lim & Zhou, 2002), regime switching market (Zhou & Yin, 2003), incomplete market (Lim, 2004), short-selling prohibition (Li *et al.*, 2002), and Skorokhod embedding problem (Ankirchner *et al.*, 2015). Besides the embedding technique developed by Li & Ng (2000), there is an alternative approach called martingale approach. This approach can be divided into two sub-problems. Firstly, the investor tries to find the targeting terminal wealth level which optimises the objective function. Secondly, the investor searches for a strategy which replicates this terminal threshold. Pliska (1982) first applies martingale approach to solve stochastic control problem. Afterwards, according to Li & Ng (2000)'s work, Bielecki *et al.* (2005) use martingale approach to solve the mean-variance problem with the short-selling constraint. Compared with the embedding approach, martingale approach appears to be more straightforward to solve the problem with constraints.

2.2.2 Equilibrium strategy

The second type of the strategy is called equilibrium strategy, which is first introduced by Basak & Chabakauri (2010). Basak & Chabakauri (2010) study the time consistent strategy by following the suggestion from Strotz (1955): an investor chooses "the best plan among those that he will actually follow." Therefore, Basak & Chabakauri (2010) investigate local behaviour of the mean-variance strategy. The total law of the variance enables them to derive recursive Hamiltonian-Jacob-Bellman equation to characterise the strategy. The main observation in their study is that they characterise a relationship between the HJB equation and the expected total gains or losses from the stock investment. This observation allows them to obtain the optimal time consistent strategy in two parts: myopic and intertemporal investment terms. The myopic term is in the form of discounted optimal mean-variance strategy for single period model. The intertemporal investment term is expressed by the sensitivity of the expected total gains or losses from the stock investment with respect to the stock price and wealth, respectively. Basak & Chabakauri (2010) examine this sensitivity of anticipated gain under the risk-neutral measures and provide a characterisation of this intertemporal investment demand under the risk-neutral measures. Basak & Chabakauri (2010) provide the insights and benefits to the mean-variance time-consistent strategy. One disadvantage could be the optimality of time-consistent strategy is not defined mathematically throughout their work.

Bjork & Murgoci (2010) are the first to formulate the time consistent strategy in game theoretic manner. It is motivated by game theory. In discrete time, the investor at different times can be viewed as different players. Given the initial value (t, x), the player t plays a non-cooperative game with the player $t+1, ..., T$. The reward to player t is given by the performance functional, and this reward only depends on the action chosen by himself and the actions chosen by players $t+1, ..., T$. The player t will make the decision based on the strategies chosen by the players $t+1, ..., T$.

In view of financial perspective, the objective functional changes at different time t which can be understood as different investment preference, the problem can be viewed as a game between the investor and his future self. Due to the fact that the objective rewards for the investor depend on the future strategies of his incarnations, every incarnation of the investor will determine the best strategy for his own objective given his best conjecture about the others. This concept coincides with the idea that "choose the best plan among those that he will actually follow (Strotz, 1955)." In discrete time, the results in forming a time-consistent policy by backward recursion from the terminal. For example, the non-exponential discounting problem (Ekeland & Lazrak, 2006) and the mean-variance problem (Björk *et al.*, 2014). Wang & Forsyth (2011) study the numerical method to obtain the equilibrium strategy. The time-consistent policy in continuous formulation can be approached via two different directions: discrete time approximation and continuous time formulation. In continuous formulation, Bjork & Murgoci (2010) give a mathematical definition of equilibrium strategy. The Markovian structure enables the authors to analyse the local behaviour of a general class of time-inconsistent functionals and derive a local strategy. The local optimality yields a system of partial differential equations (PDE), which is referred to extended Hamilton Jacobi Bellman (HJB) system. They prove a verification theorem for the extended HJB system to ensure that the solutions satisfying the extended HJB are the desired value function. On the other hand, Czichowsky (2013) studies the same problem by focusing on the mean-variance functional via discrete approximation. He proves that the discrete time formulation solution converges to the continuous case. Björk *et al.* (2017) attempt to understand the convergence theory for the general time-inconsistent problem. Yong (2012) studies the existence and the convergence of equilibrium discrete formulation solutions for hyperbolic discounting problems. Unfortunately, to the best of our knowledge, the existence and uniqueness of the solutions for the extended HJB system of general functional is still an open question.

Within previous research, by applying this approach to mean-variance framework with constant risk aversion, the obtained equilibrium control does not depend on the wealth state. This phenomenon agrees with the result from single-period model by Arrow (1971) where the constant risk aversion leads to the invariance of demand for risky assets as present wealth changes. However, in practice, the investor might change his risk attitude if any circumstances change. e.g. wealth level, investment environment or the time remaining to retire. Henderson & Hobson (2013) point out that the risk aversion can also depend on the timing. Having realised this, Björk *et al.* (2014) suggest a state-dependent risk aversion parameter for mean-variance problem and focus the risk aversion parameter λ in the form of $\frac{c}{x_t}$ for a positive constant c and wealth level x_t. This is because, as Arrow (1971) pointed out, risk aversion decreases as present wealth increases.

Björk *et al.* (2014) derive extended HJB system for a general state-dependent risk aversion. Having focused on risk aversion in the form of $\frac{c}{x_t}$, they deduce further that the equilibrium control can be written in the feedback form and linear in term of wealth state (e.g. $u(t,x) = a(t)x$). They provide a sufficient condition (an integral equation) that $a(t)$ must satisfy and prove the uniqueness of the solution for the integral equation. They propose a numerical scheme to compute the time-coefficient $a(t)$ from the integral equation and show the convergence theory for such scheme. As an extension of the model, Bensoussan *et al.* (2014) explore the same problem with no short-selling. Besides the state-dependent risk aversion, another way to have a wealth dependent equilibrium strategy is when there are only risky assets available in the market, which is studied by Lam *et al.* (2016). The reason to consider the risky asset only is due to the stochastic nature of real interest rates and the inflation risk. A risk-free asset hardly exists in a long investment horizon. Aside from portfolio selection problem, there are several applications of equilibrium approach developed in Bjork & Murgoci (2010). For example, a number of research (Li *et al.*, 2012; Zeng & Li, 2011; Zeng *et al.*, 2013) has studied the reinsurance problem for mean-variance objective reinsurer. Yao *et al.* (2013) and Wei *et al.* (2013) look at asset–liability management problem with endogenous liabilities and regime switching, respectively.

Hu *et al.* (2012) design the game theoretic formulation to the general dynamic setting. Due to the nature of general dynamic system, the impact of deviating from a strategy at a specific time cannot be fully characterised until all the information is observed. Hu *et al.* (2012) use a random variable as the perturbation to the wealth process. They study the differences between perturbed functional and unperturbed functional. They later introduce two families of backward stochastic differential equations (BSDEs), which characterises this difference between the functionals. This characterisation allows them to deduce a necessary condition for the equilibrium strategy for general dynamic. However, regarding the mean-variance problem with the same set-up, the equilibrium strategy developed in Hu *et al.* (2012) is not the same as the closed-loop equilibrium strategy in Bjork & Murgoci (2010). Due to the feature of general dynamic system, the strategy in Hu *et al.* (2012) is referred to an open-loop equilibrium strategy. Hu *et al.* (2015) take a step further and investigate connections between a family of BSDEs and the equilibrium condition. They reduce the family of BSDEs into a single BSDE, which enables them to have a deeper understanding of a specific structure of the necessary condition. They obtain a sufficient condition and provide the uniqueness of open-loop equilibrium strategy for linear-quadratic problem. Alia *et al.* (2016) attempt to solve the reinsurance problem by using the open-loop equilibrium approach.

2.2.3 Dynamic optimal strategy

Another approach to solve the time-inconsistent issue is called dynamic optimal strategy, which is first introduced in Pedersen & Peskir (2017). Pedersen & Peskir (2017) point out a key drawback of pre-commitment strategy that it uses the initial points (t, x_t) to determine all future optimal strategy u_s for all $s > t$. However, at least in Markovian framework, the investor should forget the past information and rely on the most updated information to evaluate the strategy. The notion of dynamic optimality yields a strategy in a time-consistent manner that it optimises the objective at each new point in time and wealth state. Intuitively, dynamic optimality requires the investor to solve infinitely many optimal control problems based on the different initial time points and wealth levels. In the game theoretic point of view, the dynamic optimality considers incarnations of the investors achieving their own objective rewards in an aggressive way. Pedersen & Peskir (2017) derive a closed-form solution for dynamic optimal strategy. The dynamic optimal strategy shares a lot of similarities with pre-commitment strategy. In particular, they both achieve the same expected terminal wealth or terminal value function. However, at the initial point, the value function of dynamic optimal strategy strictly dominates the value function of pre-commitment strategy. Vigna (2016) further investigates the difference in value function between dynamic optimal and pre-commitment strategy and shows that this difference is decreasing through time until it vanishes at the terminal. Vigna (2016) also studies the value functions between pre-commitment and equilibrium strategy and shows the value function of pre-commitment out-performs the equilibrium value function until terminal time. Pedersen & Peskir (2017) study a constrained case where the dynamic wealth is bounded from above. The authors identify the behaviour that the dynamic optimal wealth process remains below the upper bound until it hits the target at terminal. Pedersen & Peskir (2015) explore a case when the dynamic wealth is constrained from above and below. Pedersen & Peskir (2015) show that the terminal wealth under the dynamic optimal strategy can only take two values from the upper and lower bounds.

2.3 Problems

Although the above literature review on the time consistent approach for the mean-variance portfolio selection problem has broadened us with the understanding on the rationale of time-consistent problem, in particular, the equilibrium approach. However, this method has some limitations.

First of all, in the classic control theory with Markovian system, the open-loop strategy and the closed-loop strategy are identical. However, the current results lead the equilibrium investor into two different directions: open-loop equilibrium (Hu *et al.* (2012)) and closed-loop equilibrium (Bjork & Murgoci (2014)) with completely different trading strategies. The majority of previously paper focus on the application of a particular type of equilibrium including Liang & Song (2015), Zeng & Li (2011) and Alia *et al.* (2016), but they do not consider the differences between the two approaches. Vigna (2016) compares the value function between pre-commitment, closed-loop and dynamic optimal strategies. A number of paper use the closed-loop equilibrium to define the notion of equilibrium strategy. Hence, their analysis and numeric results may underestimate the performance of the equilibrium strategy. Basak & Chabakauri (2010) provide insights of the rationale of a closed-loop investor and the structure of a closed-loop equilibrium strategy. However, little research has investigated the open-loop strategy in discrete formulation. To complete the study about open-loop strategy and understand the rationale of open-loop investor, I will focus on the open-loop equilibrium in the discrete time formulation. Studying the discrete formulation allows us to investigate the structure of open-loop equilibrium strategy. By doing so, the effect of dynamic wealth can be identified and the differences between the rationale of the equilibrium can be studied.

Secondly, the existing results about equilibrium approach indicate that the proportion of wealth invested in stocks decreases as the investor moves away from maturity. The explanation of this phenomenon is complex. Considering the pre-commitment, the martingale approach prescribes an optimal terminal wealth and the investor increases the risky asset investment towards the end of the investment period in order to reach the benchmark. In mean-variance framework, the risk is measured by the variance which is a quadratic function of the wealth dynamic while the return is measured by the expectation which is a linear function of the wealth dynamic. This is consistent with the equilibrium approach. The variance term is not a coherent risk measure (Artzner *et al.*, 1999) in which the risk/cost arising from each time period is not additive. Therefore, the starting control strategy has more impact on the volatility of the wealth dynamic than the future strategies. Similarly, the equilibrium strategy prescribes the small stock investment when the investor is far away from the maturity as well as large stock investment when the investor is near the maturity. This finding is also supported by Aivaliotis & Palczewski (2014). They study the mean-variance objective functional depending on the whole trajectory of the control process instead of terminal wealth. However, when the investment period is sufficiently long, compared to the pre-commitment strategy, the equilibrium strategy behaves differently at the starting point. To our best knowledge, little study has looked at the behaviour of trajectory of the equilibrium strategy.

Finally, Markowitz (1968) discusses the advantages and limitations of using variance as the risk measure. The main issue is that the variance penalises the wealth process for being too high as well as being too low. This limits the investment strategy to seek for a situation that is profitable. These considerations drive the research into asymmetric risk measures involving lower partial moments and semi-variance. Research in this area includes Markowitz *et al.* (1993), Bawa & Lindenberg (1977), Harlow & Rao (1989), Rockafellar & Uryasev (2000), and Krokhmal *et al.* (2002). More recently the idea of cash withdrawal technique was introduced and further developed by Cui *et al.* (2012). Cui *et al.* (2012) study the situation where investor decides to withdraw the cash dynamically to avoid the wealth level to be too high. The cash withdrawal strategy achieves the same efficient mean-variance pair of the terminal wealth as well as offers free cash flow by the withdrawal. In the context of equilibrium approach in the multi-period framework, the portfolio optimisation problem using semi-variance has not received much attention. Rudloff *et al.* (2014) consider the time-consistent strategy for mean-cvar problem based on nested objective functional. However, in my thesis, I will concentrate on the empirical study of the equilibrium approach to evaluate the performance. The reason is that, when the notion of optimality changes to equilibrium, it is less intuitive to determine the real objective for the equilibrium investor. It is no longer to optimise the mean-variance objective at every point in time. Instead, the investor plays a game with the future incarnations of himself. Therefore, comparing the equilibrium value function with the rest becomes less effective.

Chapter 3

Open-loop equilibrium strategy in discrete time

3.1 Introduction

Hu *et al.* (2012) first study the equilibrium strategy for linear-quadratic optimization problem, which is an extended study for the previous paper by Zhou & Li (2000). The equilibrium control problem in continuous time framework has been developed by using the spike variation in which the investor deviates the strategy in a short time horizon and study the infinitesimal of the objective functions. This chapter aims to use a similar approach to study the equilibrium strategy in discrete time setting. In Hu *et al.* (2012), the authors define the perturbation as a random variable in continuous time. As described earlier, the two different perturbations used in the existing research lead to different strategies. However, compare to the continuous case, it is more straightforward to investigate the rationale of the equilibrium strategy in discrete time. Therefore, this chapter presents the perturbed method in discrete time for obtaining the open-loop equilibrium strategy.

3.2 Overview of the model and the equilibrium condition

We consider a discrete time, finite horizon, ($t \in \mathbb{T} := \{0, 1, \ldots, T\}$) a general market with $(n+1)$ available assets: a risk-free bond S_t^0 with its deterministic growth factor r_t at time t and n risky stock assets with prices $\mathbf{S}_t = [S_t^1, \ldots, S_t^n]^T$ with its random return factor $b_t^i := \frac{S_t^i}{S_{t-1}^i}$ at time t for $i = 1, \ldots, n$, where $(\cdot)^T$ denotes a transformed vector. It is assumed that the first and second moment of $\mathbf{b}_t$ are known and the $\mathrm{Cov}(\mathbf{b}_t)$ is strictly positive definite. The information set at time t is given by

$$\mathcal{F}_t = \sigma(\mathbf{S}_0, \mathbf{S}_1, \ldots, \mathbf{S}_t).$$

A self-financing investor joins the market at any time $t \in \{0, 1, \ldots, T-1\}$ with the wealth X_t. Therefore, his entire investment strategy is described by a sequence of $\mathbb{R}^n$-valued random variables $\bar{\mathbf{u}}_t := (\mathbf{u}_t, \ldots, \mathbf{u}_{T-1})$, where $\mathbf{u}_s = [u_s^1, \ldots, u_s^n]^T$ and u_s^i is the amount of wealth invested in the i^{th} risky asset in period s. Let $\mathcal{A}$ denotes the admissible set and a strategy is said to be *admissible* $u_s \in \mathcal{A}$ if $\mathbf{u}_s$ is $\mathcal{F}_s$-measurable and $\mathbf{b}_{s+1}\mathbf{u}_s$ is square integrable for all s. Since the amount of money invested in the risk-free asset at time t is $X_t^{\bar{\mathbf{u}}} - \sum_{i=1}^n u_t^i$, the dynamics of investor's wealth follow

$$X_{t+1}^{\bar{\mathbf{u}}} = \sum_{i=1}^{n} \frac{u_t^i}{S_t^i} S_{t+1}^i + (X_t^{\bar{\mathbf{u}}} - \sum_{i=1}^{n} u_t^i)\, e^{r_t^0} = A_t X_t^{\bar{\mathbf{u}}} + \mathbf{B}_{t+1}^T \mathbf{u}_t, \tag{3.1}$$

where $A_t = e^{r_t^0}$ and $\mathbf{B}_{t+1} := [B_{t+1}^1, B_{t+1}^2, \ldots, B_{t+1}^n]^T = [(b_{t+1}^1 - e^{r_t^0}), (b_{t+1}^2 - e^{r_t^0}), \ldots, (b_{t+1}^n - e^{r_t^0})]^T$. When it is clear from the context, we will drop $\bar{\mathbf{u}}$ from $X_t^{\bar{\mathbf{u}}}$.

Suppose the investor entering the market at time t with the wealth $X_t \in L^2(\mathcal{F}_t)$ would like to maximize the following mean-variance objective functional over admissible investment strategies:

$$\operatorname*{ess\,sup}_{\bar{\mathbf{u}} \in \mathcal{A}} \left\{ X_t\, \mathbb{E}_t[X_T^{\bar{\mathbf{u}}}] - \frac{\gamma}{2} \mathrm{Var}_t[X_T^{\bar{\mathbf{u}}}] \right\}, \tag{3.2}$$

where $\gamma > 0$ will be called the investor's *risk aversion*, and $\mathbb{E}_t[\cdot]$ and $\mathrm{Var}_t(\cdot)$ denote $\mathbb{E}[\cdot|\mathcal{F}_t]$ and $\mathrm{Var}(\cdot|\mathcal{F}_t)$, respectively. Notice that under the assumption that $X_t > 0$

a.s., this problem is equivalent to the classical Markowitz mean-variance criterion for the return of the portfolio over $[t, T]$

$$\operatorname*{ess\,sup}_{\bar{\mathbf{u}}} \left\{ \mathbb{E}_t[X_T^{\bar{\mathbf{u}}}/X_t] - \frac{\gamma}{2} \mathrm{Var}_t[X_T^{\bar{\mathbf{u}}}/X_t] \right\}.$$

Criterion (3.2) can also be interpreted, c.f. Björk *et al.* (2014), as imposing a state-dependent risk aversion $\frac{\gamma}{2X_t}$ whereby the investor with a higher wealth would be more risk-seeking than the investor with a lower wealth.

For an admissible strategy $\bar{\mathbf{u}}_t$ and an initial wealth $X_t \in L^2(\mathcal{F}_t)$ at time t, denote

$$J(t, X_t; \bar{\mathbf{u}}_t) = X_t\, \mathbb{E}_t[X_T^{\bar{\mathbf{u}}}] - \frac{\gamma}{2} \mathrm{Var}_t[X_T^{\bar{\mathbf{u}}}].$$

3.3 Open-loop strategy for mean-variance problem

Definition 3.3.1. An admissible strategy $\bar{\mathbf{u}}_0^* \equiv (\mathbf{u}_0^*, \mathbf{u}_1^*, \ldots, \mathbf{u}_{T-1}^*)$ is an *open-loop equilibrium control* if for any t and any bounded $\mathbb{R}^n$-valued $\mathcal{F}_t$-measurable random variable ε

$$J(t, X_t^{\bar{\mathbf{u}}^{*,\varepsilon}}; \bar{\mathbf{u}}_t^{*,\varepsilon}) - J(t, X_t^{\bar{\mathbf{u}}^*}; \bar{\mathbf{u}}_t^*) \le 0, \quad \text{a.s.},$$

where $\bar{\mathbf{u}}_t^* = (\mathbf{u}_t^*, \mathbf{u}_{t+1}^*, \ldots, \mathbf{u}_{T-1}^*)$, and $\bar{\mathbf{u}}_t^{*,\varepsilon} = (\mathbf{u}_t^* + \varepsilon, \mathbf{u}_{t+1}^*, \ldots, \mathbf{u}_{T-1}^*)$.

The idea of the equilibrium concept is that, when the investor decides to change the investment at any single point $t = 0, \ldots, T-1$, the rewards for investor would be worse than if he had not changed the strategy.

Dynamics (3.1) allow for explicit expression for the terminal wealth. Given $X_t \in$

$L^2(\mathcal{F}_t)$ at time t and an admissible control $\bar{\mathbf{u}}_t$ we have

$$\begin{aligned}
X_T^{\bar{\mathbf{u}}} &= A_{T-1}X_{T-1}^{\bar{\mathbf{u}}} + \mathbf{B}_T^T\mathbf{u}_{T-1} \\
&= A_{T-1}\big(A_{T-2}X_{T-2}^{\bar{\mathbf{u}}} + \mathbf{B}_{T-1}^T\mathbf{u}_{T-2}\big) + \mathbf{B}_T^T\mathbf{u}_{T-1} \\
&= \Big(\prod_{i=T-2}^{T-1} A_i\Big)X_{T-2}^{\bar{\mathbf{u}}} + A_{T-1}\mathbf{B}_{T-1}^T\mathbf{u}_{T-2} + \mathbf{B}_T^T\mathbf{u}_{T-1}
\end{aligned}$$

Define $D_T = 1$ and $D_t = \prod_{i=t}^{T-1} A_i$, then

$$X_T^{\bar{\mathbf{u}}} = D_{T-2}X_{T-2}^{\bar{\mathbf{u}}} + D_{T-1}\mathbf{B}_{T-1}^T\mathbf{u}_{T-2} + D_T\mathbf{B}_T^T\mathbf{u}_{T-1}$$

Repeating the above procedure until time t, we obtain

$$X_T^{\bar{\mathbf{u}}} = D_tX_t + \sum_{s=t}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T\mathbf{u}_s \tag{3.3}$$

For a perturbed control $\bar{\mathbf{u}}_t^{\varepsilon}$, as constructed in Definition 3.3.1, the terminal wealth equals

$$\begin{aligned}
X_T^{\bar{\mathbf{u}}^{\varepsilon}} &= D_tX_t^{\bar{\mathbf{u}}} + \sum_{s=t+1}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T\mathbf{u}_s + D_{t+1}\mathbf{B}_{t+1}^T(\mathbf{u}_t + \varepsilon) \\
&= D_tX_t^{\bar{\mathbf{u}}} + \sum_{s=t}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T\mathbf{u}_s + D_{t+1}\mathbf{B}_{t+1}^T\varepsilon \\
&= X_T^{\bar{\mathbf{u}}} + D_{t+1}\mathbf{B}_{t+1}^T\varepsilon
\end{aligned} \tag{3.4}$$

The expression (3.4) shows that, the perturbed wealth process is differed from the original process by a term $D_{t+1}\mathbf{B}_{t+1}^T\varepsilon$. The extra term can be view as following: investing ε amount of money into the stock market at time t. After one-period time, the amount becomes $\mathbf{B}_{t+1}^T\varepsilon$. The investor then withdraws it from the stock market and invests it into the risk-free bond until terminal time.

Define $V_T^{\bar{\mathbf{u}}} = 0$ and

$$V_t^{\bar{\mathbf{u}}} = \sum_{s=t}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T\mathbf{u}_s, \qquad t = 0, \dots, T-1, \tag{3.5}$$

and introduce the notation that will be used to characterise equilibrium strategy:

$$\Sigma_t = \mathrm{Cov}_t[\mathbf{B}_{t+1}], \qquad \mu_t = \mathbb{E}_t[\mathbf{B}_{t+1}], \qquad \beta_t^{\bar{\mathbf{u}}} = \mathrm{Cov}_t[\mathbf{B}_{t+1}, \mathbb{E}_{t+1} V_{t+1}^{\bar{\mathbf{u}}}].$$

Formulae (3.3)-(3.4) play an important role in deriving open-loop equilibrium controls. In the following theorem, we prove the uniqueness of open-loop equilibrium strategy for a general wealth process (3.1). As a by-product, we obtain a recursive formula for backward computation of a candidate strategy. If this strategy satisfies the required integrability conditions, i.e., $\mathbf{B}_{t+1}\mathbf{u}_t \in L^2(\mathcal{F}_t)$ for all t then it is an open-loop equilibrium control. Otherwise, there is no admissible equilibrium strategy.

Theorem 3.3.2. *A open-loop equilibrium control $\bar{\mathbf{u}}_t^*$ satisfies*

$$\mathbf{u}_t^* = \frac{1}{\gamma D_{t+1}}\Sigma_t^{-1}\big(X_t^{\bar{\mathbf{u}}^*}\mu_t - \gamma\beta_t^{\bar{\mathbf{u}}^*}\big). \tag{3.6}$$

Proof. Fix an admissible control strategy $\bar{\mathbf{u}}$ and assume perturbed version of this admissible strategy $\bar{\mathbf{u}}^\varepsilon$ with a deterministic perturbation $\varepsilon \in \mathbb{R}^n$ applied at time t. Later, we can relax the condition for the perturbation ε. Recalling expressions (3.3)-(3.4), we notice that

$$\mathbb{E}_t[X_T^{\bar{\mathbf{u}}^\varepsilon}] = \mathbb{E}_t[X_T^{\bar{\mathbf{u}}}] + D_{t+1}\mu_t^T\varepsilon.$$

Recall the law of total variance,

$$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}(\mathbb{E}[Y|X]),$$

we obtain

$$
\begin{aligned}
\mathrm{Var}_t[X_T^{\bar{\mathbf{u}}^\varepsilon}] &= \mathbb{E}_t\left[\mathrm{Var}_{t+1}[X_T^{\bar{\mathbf{u}},\varepsilon}]\right] + \mathrm{Var}_t\left[\,\mathbb{E}_{t+1}[X_T^{\bar{\mathbf{u}},\varepsilon}]\right] \\
&= \mathbb{E}_t\left[\mathrm{Var}_{t+1}[V_{t+1}]\right] + \mathrm{Var}_t\left[D_t X_t + D_{t+1}\mathbf{B}_{t+1}^T(\mathbf{u}_t+\varepsilon) + \mathbb{E}_{t+1}[V_{t+1}]\right] \\
&= \mathbb{E}_t\left[\mathrm{Var}_{t+1}[V_{t+1}]\right] + \mathrm{Var}_t\left[\,\mathbb{E}_{t+1}[V_{t+1}]\right] + D_{t+1}^2(\mathbf{u}_t+\varepsilon)^T\Sigma_t(\mathbf{u}_t+\varepsilon) \\
&\quad + 2D_{t+1}\beta_t^T(\mathbf{u}_t+\varepsilon) \\
&= \mathrm{Var}_t[V_{t+1}] + D_{t+1}^2(\mathbf{u}_t+\varepsilon)^T\Sigma_t(\mathbf{u}_t+\varepsilon) + 2D_{t+1}\beta_t^T(\mathbf{u}_t+\varepsilon),
\end{aligned}
$$

where in β and V_t we skipped the superscript $\bar{\mathbf{u}}$. Therefore,

$$
\begin{aligned}
&J(t, X_t^{\bar{\mathbf{u}}}; \bar{\mathbf{u}}_t^\varepsilon) - J(t, X_t^{\bar{\mathbf{u}}}; \bar{\mathbf{u}}_t) \\
&= X_t^{\bar{\mathbf{u}}} D_{t+1}\mu_t^T\varepsilon - \frac{\gamma}{2}\Big(D_{t+1}^2(\mathbf{u}_t+\varepsilon)^T\Sigma_t(\mathbf{u}_t+\varepsilon) + 2D_{t+1}\beta_t^T(\mathbf{u}_t+\varepsilon) \\
&\quad - D_{t+1}^2\mathbf{u}_t^T\Sigma_t\mathbf{u}_t - 2D_{t+1}\beta_t^T\mathbf{u}_t\Big) \\
&= X_t^{\bar{\mathbf{u}}} D_{t+1}\mu_t^T\varepsilon - \gamma D_{t+1}^2\mathbf{u}_t^T\Sigma\varepsilon - \frac{\gamma}{2}D_{t+1}^2\varepsilon^T\Sigma_t\varepsilon - \gamma D_{t+1}\beta_t^T\varepsilon. \qquad (3.7)
\end{aligned}
$$

For $\bar{\mathbf{u}}$ to be an open equilibrium strategy, the above expression has to be a.s. non-positive for any ε, in particular, for deterministic ε. As Σ_t is strictly positive definite, the equation (3.7) is a strictly concave function in ε. The condition of non-positivity is satisfied if and only if the maximum is at $\varepsilon = \mathbf{0}$. Due to concavity, the first order condition is necessary and sufficient. The derivative with respect to ε reads

$$
X_t^{\bar{\mathbf{u}}}\mu_t - \gamma D_{t+1}\Sigma_t\mathbf{u}_t - \gamma D_{t+1}\Sigma_t\varepsilon - \gamma\beta_t.
$$

The first order condition at $\varepsilon = \mathbf{0}$ takes the form:

$$
X_t^{\bar{\mathbf{u}}}\mu_t - \gamma D_{t+1}\Sigma_t\mathbf{u}_t - \gamma\beta_t = \mathbf{0}.
$$

Therefore, $\mathbf{u}_t$ has to satisfy

$$\mathbf{u}_t = \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \big(X_t^{\bar{\mathbf{u}}} \mu_t - \gamma \beta_t \big). \tag{3.8}$$

For the sufficient condition, inserting the above expression (3.8) into (3.7) gives

$$J(t, X_t^{\bar{\mathbf{u}}}; \bar{\mathbf{u}}_t^{\varepsilon}) - J(t, X_t^{\bar{\mathbf{u}}}; \bar{\mathbf{u}}_t) = -\frac{\gamma}{2} D_{t+1}^2 \varepsilon^T \Sigma_t \varepsilon,$$

which is non-positive for all ε. Now, we relax ε to be a bounded $\mathcal{F}_t$-measurable random variable. Due to positive definiteness of Σ_t, the above expression is non-positive for all ε.

□

Condition (3.6) provides a recipe for constructing an open-loop equilibrium. At $T-1$, $V_T = 0$, so $\beta_T = \mathbf{0}$ and $\mathbf{u}_{T-1}$ is uniquely determined by the conditional moments of $\mathbf{B}_T$ given $\mathcal{F}_{T-1}$. Moving to $T-2$, the random variable V_{T-1} is known and therefore β_{T-1} can be computed, provided that the covariance is well-defined (it will be discussed later on). Equation (3.6) gives a unique characterisation of $\mathbf{u}_{T-2}$, and we continue down to $\mathbf{u}_0$.

The above procedure depends on the ability to compute the covariance β_t. If at each step, we could obtain an admissible strategy (i.e. $\mathbf{b}_{s+1}\mathbf{u}_s$ is square integrable for all s) and $V_t = \sum_{s=t}^{T-1} D_{s+1} \mathbf{B}_{s+1}^T \mathbf{u}_s$ is square-integrable. Therefore, the covariance β_t is well-defined. The control strategy exists if and only if, in each step of the above procedure, the calculated risky asset position $\mathbf{u}_t \in \mathcal{A}$, i.e. $\mathbf{B}_{t+1}^T \mathbf{u}_t \in L^2(\mathcal{F}_{t+1})$. It is not clear that the condition (3.6) yields a unique construction control sequence as the random variable V_t does depend on all the control $\mathbf{u}_s$ for $s \leq t$. We summarise these findings below.

Theorem 3.3.3. *Define K_t by the following:*

$$\mathbf{K}_t = \mathbf{E}_{t+1}^{-1} \Big(\frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \mu_t - \frac{1}{D_{t+1}} \Sigma_t^{-1} A_t \mathrm{Cov}_t \big(\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \big) \Big),$$

where

$$\mathbf{E}_{t+1} = \mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \big(\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1}^T \big)$$

and M_{t+1} is constructed by induction:

$$M_{t+1} := \Big[\sum_{s=t+2}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T \mathbf{K}_s \big(\prod_{n=t+1}^{s-1} (A_n + \mathbf{B}_{n+1}^T \mathbf{K}_n) \big) + D_{t+2}\mathbf{B}_{t+2}^T \mathbf{K}_{t+1} \Big]$$

and $M_T = 0$. If the above inverse of the matrix $\mathbf{E}_{t+1}$ is well-defined for all t, the control

$$\mathbf{u}_t^* = \mathbf{K}_t X_t^*,$$

where $X_t^* = \prod_{i=0}^{t-1} \big(A_i + \mathbf{B}_{i+1}^T \mathbf{K}_i\big) X_0$, is the unique open-loop equilibrium control.

Proof. The theorem 3.3.2 characterises a necessary and sufficient condition for equilibrium control. It allows us to construct the control recursively. However, β_t is not necessarily unique since it depends on the control sequence $\{u_0, \ldots, u_{T-1}\}$. We need to show that, under such construction, the equilibrium control is uniquely defined.

At time T-1, we have $\beta_{T-1} = 0$, so

$$\begin{aligned}
\mathbf{u}_{T-1}^* &= \frac{1}{\gamma D_T} \Sigma_{T-1}^{-1} \big(X_{T-1}^{\bar{\mathbf{u}}^*} \mu_{T-1} \big) \\
&= \big(\frac{1}{\gamma D_T} \Sigma_{T-1}^{-1} \mu_{T-1} \big) X_{T-1}^{\bar{\mathbf{u}}^*} \\
&:= \mathbf{K}_{T-1} X_{T-1}^{\bar{\mathbf{u}}^*},
\end{aligned}$$

where K_{T-1} is $\mathcal{F}_{T-1}$-measurable that depends only on the filtration of the stock price $\sigma(S_0, \ldots, S_{T-1})$.

Assume that $\mathbf{u}_s^* = \mathbf{K}_s X_s^{\bar{\mathbf{u}}^*}$ for all $s \geq t+1$ and $\mathbf{K}_s$ depends only on the filtration of the stock price $\sigma(S_0, \ldots, S_s)$, i.e., it does not depend on any of the control sequence $\{\mathbf{u}_0^*, \ldots, \mathbf{u}_{s-1}^*\}$. Therefore, we can rewrite the wealth at time $t+2, \ldots, T-1$ in terms of the wealth at time $t+1$ in the following way:

$$X_{t+2}^{\bar{\mathbf{u}}^*} = \big(A_{t+1} + \mathbf{B}_{t+2}^T \mathbf{K}_{t+1}\big) X_{t+1}^{\bar{\mathbf{u}}^*},$$

$$X_{t+3}^{\bar{\mathbf{u}}^*} = \big(A_{t+2} + \mathbf{B}_{t+3}^T \mathbf{K}_{t+2}\big)\big(A_{t+1} + \mathbf{B}_{t+2}^T K_{t+1}\big) X_{t+1}^{\bar{\mathbf{u}}^*}.$$

Therefore, for any $s \in \{t+2, \ldots, T-1\}$,

$$X_s^{\bar{\mathbf{u}}^*} = \big(\prod_{n=t+1}^{s-1} (A_n + \mathbf{B}_{n+1}^T \mathbf{K}_n)\big) X_{t+1}^{\bar{\mathbf{u}}^*}.$$

Then we have

$$\begin{aligned}
V_{t+1}^{\bar{\mathbf{u}}^*} &= \sum_{s=t+1}^{T-1} D_{s+1} \mathbf{B}_{s+1}^T \mathbf{u}_s^* \\
&= \sum_{s=t+1}^{T-1} D_{s+1} \mathbf{B}_{s+1}^T \mathbf{K}_s X_s^{\bar{\mathbf{u}}^*} \\
&= \sum_{s=t+2}^{T-1} D_{s+1} \mathbf{B}_{s+1}^T \mathbf{K}_s \Big[\Big(\prod_{n=t+1}^{s-1} (A_n + \mathbf{B}_{n+1}^T \mathbf{K}_n)\Big) X_{t+1}^{\bar{\mathbf{u}}^*}\Big] + D_{t+2} \mathbf{B}_{t+2}^T \mathbf{K}_{t+1} X_{t+1}^{\bar{\mathbf{u}}^*} \\
&= \Big[\sum_{s=t+2}^{T-1} D_{s+1} \mathbf{B}_{s+1}^T \mathbf{K}_s \big(\prod_{n=t+1}^{s-1} (A_n + \mathbf{B}_{n+1}^T \mathbf{K}_n)\big) + D_{t+2} \mathbf{B}_{t+2}^T \mathbf{K}_{t+1}\Big] \cdot X_{t+1}^{\bar{\mathbf{u}}^*} \\
&:= M_{t+1} \cdot X_{t+1}^{\bar{\mathbf{u}}^*},
\end{aligned}$$

where M_{t+1} is a $\mathcal{F}_{T-1}$-measurable random variable which depends only on stock price filtration $\sigma(S_0, \ldots, S_{T-1})$. Therefore

$$\begin{aligned}
\beta_t &= \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[V_{t+1}^{\bar{\mathbf{u}}^*}]\Big] \\
&= \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] X_{t+1}^{\bar{\mathbf{u}}^*}\Big] \\
&= \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}](A_t X_t^{\bar{\mathbf{u}}^*} + \mathbf{B}_{t+1}^T \mathbf{u}_t^*)\Big] \\
&= A_t X_t^{\bar{\mathbf{u}}^*} \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big] + \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1}\Big] \mathbf{u}_t^*.
\end{aligned}$$

Substituting β_t into (3.8), we obtain

$$
\begin{aligned}
\mathbf{u}_t^* &= \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \big(X_t^{\bar{\mathbf{u}}^*} \mu_t - \gamma \beta_t \big) \\
&= \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} X_t^{\bar{\mathbf{u}}^*} \mu_t - \frac{1}{D_{t+1}} \Sigma_t^{-1} \beta_t \\
&= \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} X_t^{\bar{\mathbf{u}}^*} \mu_t - \frac{A_t X_t^{\bar{\mathbf{u}}^*}}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \Big] \\
&\quad - \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1} \Big] \mathbf{u}_t^*.
\end{aligned}
$$

Rearranging above expression, we get

$$
\begin{aligned}
&\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1} \Big] \right) \mathbf{u}_t^* \\
&= \left(\frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \mu_t - \frac{A_t}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \Big] \right) X_t^{\bar{\mathbf{u}}^*}.
\end{aligned}
$$

Under the assumption that the matrix

$$
\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1} \big] \right)
$$

is invertible, then we have

$$
\begin{aligned}
\mathbf{u}_t^* &= \left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \mathbf{B}_{t+1} \big] \right)^{-1} \Bigg(\frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \mu_t \\
&\quad - \frac{A_t}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t \Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}] \Big] \Bigg) X_t^{\bar{\mathbf{u}}^*} \\
&= \mathbf{K}_t X_t^{\bar{\mathbf{u}}^*}
\end{aligned}
$$

where $\mathbf{K}_t$ is $\mathcal{F}_t$-measurable depends only on the stock price filtration $\sigma(S_0, \ldots, S_t)$

By induction, we can construct an unique strategy sequence $\mathbf{u} = (u_0, u_1, \ldots, u_{T-1})$. Since such construction agrees the form in equation (3.8), the sufficiency is followed by Theorem 3.3.2. □

The above theorem shows that the open-loop equilibrium control is linear in term of wealth X with a random coeffient $\mathbf{K}$. For $t \in (0, \ldots, T-1)$, any random coefficient $\mathbf{K}_t$ is independent of choices of the strategies $(u_0, \ldots, u_{t-1})$. Although we construct the equilibrium strategy backward, we can compute each random coefficient $\mathbf{K}_t$ for all t and the value of $\mathbf{K}_t$ is uniquely defined with respect to the stock price filtration.

Corollary 3.3.4. *Assume the excess return $\mathbf{B}_t$ is essentially bounded for all t, i.e. there exists a constant c such that $\operatorname{ess\,sup}_{\omega\in\Omega} |B_t| = c_1 < \infty$ for all t, and $\operatorname{ess\,sup}_{\omega} \|\Sigma^{-1}\|_u \leq c_2 < \infty$. Also the matrix E_t defined in theorem 3.3.3 is invertible for all t. Then an open-loop equilibium strategy is unique. Moreover, it is determined by a backward procedure stemming from (3.6) provided that at each step the calculated $\mathbf{u}_t$ satisfies that $\mathbf{B}_{t+1}^T \mathbf{u}_t$ is a square integrable random variable. (i.e. admissible strategy)*

Proof. Define the upper and lower bounds on $\mathcal{L}_2$ norm space as:

$$\|A\|_u = \max_{\|\xi\|_2=1} \|A\xi\|_2$$

$$\|A\|_l = \min_{\|\xi\|_2=1} \|A\xi\|_2$$

For any $t \in \{0, 1, \ldots, T-1\}$, there exists some constants c_1 and c_2 such that

$$\operatorname*{ess\,sup}_{\omega} |B_t| = c_1 < \infty,$$

$$\operatorname*{ess\,sup}_{\omega} \left\|\Sigma^{-1}\right\|_u \leq c_2 < \infty.$$

Recall that $\mathbf{u}_t^* = \mathbf{K}_t X_t^{\bar{\mathbf{u}}^*}$ where

$$\mathbf{K}_t := \Big(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\big]\Big)^{-1}$$
$$\Big(\frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \mu_t - \frac{A_t}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big]\Big).$$

Then, for any $t \in \{0, 1, \ldots, T-1\}$,

$$\mathbb{E}\left[\left(\mathbf{B}_{t+1}^T \mathbf{u}_t^*\right)^2\right] \leq c_1^2\, \mathbb{E}\left[\left(\mathbf{u}_t^*\right)^2\right]$$
$$= c_1^2\, \mathbb{E}\left[\mathbf{K}_t^T \mathbf{K}_t \big(X_t^{\bar{\mathbf{u}}^*}\big)^2\right].$$

We will prove by induction that $\mathbf{K}_t$ is bounded for any $t \in \{0, 1, \ldots, T-1\}$. At time $T-1$, since $\beta_{T-1} = 0$, then

$$\begin{aligned}
\left\|\mathbf{K}_{T-1}^T \mathbf{K}_{T-1}\right\|_u &\leq \frac{1}{\gamma^2 D_T^2} \left\|\mu_{T-1}^T \Sigma_{T-1}^{-1} \Sigma_{T-1}^{-1} \mu_{T-1}\right\|_u \\
&\leq \frac{1}{\gamma^2 D_T^2} \left\|\mu_{T-1}^T\right\|_u \left(\left\|\Sigma_{T-1}^{-1}\right\|_u\right)^2 \left\|\mu_{T-1}\right\|_u \\
&= \frac{c_1^2 c_2^2}{\gamma^2 D_T^2},
\end{aligned}$$

which is bounded from above. We make the following induction hypothesis: if $\{\mathbf{K}_{t+1}, \ldots, \mathbf{K}_{T-1}\}$ are bounded, then $\mathbf{K}_t$ is bounded. Denote $d = \max\left\{\left\|\mathbf{K}_{t+1}\right\|_u, \ldots, \left\|\mathbf{K}_{T-1}\right\|_u\right\}$. Since $\mathbf{K}_t$ is a product of two matrices, we can prove that the operator norm is bounded for each individual matrix. Since $\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\left[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\right]\right)^{-1}$ is a $n \times n$ square-matrix, then

$$\begin{aligned}
&\left\|\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\left[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\right]\right)^{-1}\right\|_u \\
&= \left\|\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\left[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\right]\right)\right\|_l^{-1}
\end{aligned}$$

Since the matrix $\left(\mathbf{I} + \frac{1}{D_{t+1}} \Sigma_t^{-1} \mathrm{Cov}_t\left[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\right]\right)$ is an $n \times n$ real symmetric matrix. Therefore, the matrix has n number of non-zero eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$.

Theorem. *Given $A \in \mathbb{C}^{n \times n}$, suppose $rank(A) = n$. Then*

$$\min_{\|x\|=1} \|Ax\|_2 = \sigma_{min}(A)$$

Proof. For any $\|x\| = 1$,

$$\begin{aligned}\|Ax\|_2 &= \|U\Sigma V'x\|_2 \\ &= \|\Sigma V'x\|_2 \\ &= \|\Sigma y\|_2 \text{ for } y = V'x.\end{aligned}$$

Then,

$$\|\Sigma y\|_2 = \left(\Sigma_{i=1}^n |\sigma_i y_i|^2\right)^2 \geq \sigma_{min}$$

□

Denote $\lambda_{min} = \min\{|\lambda_1|, \dots, |\lambda_n|\}$, then

$$\left\|\left(\mathbf{I} + \frac{1}{D_{t+1}}\Sigma_t^{-1}\mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\Big]\right)^{-1}\right\|_u \leq \frac{1}{\lambda_{min}}.$$

To prove the boundedness for

$$\left(\frac{1}{\gamma D_{t+1}}\Sigma_t^{-1}\mu_t - \frac{A_t}{D_{t+1}}\Sigma_t^{-1}\mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big]\right),$$

we will need the following estimates:

$$\begin{aligned}\|M_{t+1}\|_u &= \left\|\sum_{s=t+2}^{T-1} D_{s+1}\mathbf{B}_{s+1}^T\mathbf{K}_s\big(\prod_{n=t+1}^{s-1}(A_n + \mathbf{B}_{n+1}^T\mathbf{K}_n)\big) + D_{t+2}\mathbf{B}_{t+2}^T\mathbf{K}_{t+1}\right\|_u \\ &\leq \sum_{s=t+2}^{T-1}\left\|D_{s+1}\mathbf{B}_{s+1}^T\mathbf{K}_s\big(\prod_{n=t+1}^{s-1}(A_n + \mathbf{B}_{n+1}^T\mathbf{K}_n)\big)\right\|_u + \left\|D_{t+2}\mathbf{B}_{t+2}^T\mathbf{K}_{t+1}\right\|_u \\ &\leq \sum_{s=t+2}^{T-1} D_{s+1}c_1 d\Big(\prod_{n=t+1}^{s-1}(A_n + c_1 d)\Big) + D_{t+2}c_1 d \\ &=: c_3 < \infty,\end{aligned}$$

and

$$
\begin{aligned}
&\left\|\mathrm{Cov}_t\big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\mathbf{B}_{t+1}\big]\right\|_u \\
&= \left\|\mathbb{E}_t\left[\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T\,\mathbb{E}_{t+1}[M_{t+1}]\right] - \mathbb{E}_t\left[\mathbf{B}_{t+1}\right]\mathbb{E}_t\left[\mathbf{B}_{t+1}^T\,\mathbb{E}_{t+1}[M_{t+1}]\right]^T\right\|_u \\
&\leq \left\|\mathbb{E}_t\left[\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T\,\mathbb{E}_{t+1}[M_{t+1}]\right]\right\|_u + \left\|\mathbb{E}_t\left[\mathbf{B}_{t+1}\right]\mathbb{E}_t\left[\mathbf{B}_{t+1}^T\,\mathbb{E}_{t+1}[M_{t+1}]\right]^T\right\|_u \\
&\leq 2c_1^2c_3 =: c_4.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\left\|\frac{1}{\gamma D_{t+1}}\Sigma_t^{-1}\mu_t - \frac{A_t}{D_{t+1}}\Sigma_t^{-1}\mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big]\right\|_u \\
&\leq \left\|\frac{1}{\gamma D_{t+1}}\Sigma_t^{-1}\mu_t\right\|_u + \left\|\frac{A_t}{D_{t+1}}\Sigma_t^{-1}\mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big]\right\|_u \\
&\leq \frac{1}{\gamma D_{t+1}}\left\|\Sigma_t^{-1}\right\|_u\left\|\mu_t\right\|_u + \frac{A_t}{D_{t+1}}\left\|\Sigma_t^{-1}\right\|_u\left\|\mathrm{Cov}_t\Big[\mathbf{B}_{t+1}, \mathbb{E}_{t+1}[M_{t+1}]\Big]\right\|_u \\
&\leq \frac{c_1c_2}{\gamma D_{t+1}} + \frac{A_tc_2c_4}{D_{t+1}}.
\end{aligned}
$$

This yields

$$
\left\|\mathbf{K}_t^T\mathbf{K}_t\right\|_u \leq \frac{1}{\lambda_1^2}\Big(\frac{c_1c_2}{\gamma D_{t+1}} + \frac{A_tc_2c_4}{D_{t+1}}\Big)^2 < \infty.
$$

Hence, we have shown that $\mathbf{K}_t$ is bounded for any $t \in \{0,\ldots,T-1\}$. This implies the control u_t^* is admissible if and only if the wealth level $X_t^{\bar{\mathbf{u}}^*}$ is square-integrable. Since the initial wealth of investor entering the market, the starting wealth is a constant at that point. Then, this implies the control u_0 is admissible. As a consequence, X_1 is square-integrable under the admissible control u_0. Follow such procedure, at maturity time, we obtain an admissible equilibrium control sequence $\{\mathbf{u}_0^*,\ldots,\mathbf{u}_{T-1}^*\}$. □

Although formula (3.6) can be expressed in the feedback form, a more convenient formulation can be obtained using a slightly different decomposition of wealth than in the proof of Theorem 3.3.2.

To simplify notation in the Markovian case, we write $\mathbb{E}_{tsx}[\cdot]$ for $\mathbb{E}[\cdot|S_t = s, X_t = x]$ and analogously for the variance and covariance operator. When the conditioning is on the state variables S_t, the state variable x is omitted in the notation, for example, $\mathbb{E}_{ts}[\cdot] = \mathbb{E}[\cdot|S_t = s]$.

Theorem 3.3.5. *In the Markovian setting (i.e. the stock price process admits strong Markov property), an open-loop equilibrium strategy* $\bar{\mathbf{u}}^*$ *satisfies*

$$\mathbf{u}_t^* = \alpha_t(S_t)\ X_t,$$

where

$$\alpha_t(s) = \big(\mathrm{Cov}_{ts}\big[h_{t+1}(sB_{t+1})B_{t+1}, B_{t+1}\big]\big)^{-1}\Big(\frac{1}{\gamma}\,\mathbb{E}_{ts}[B_{t+1}] - A_t\mathrm{Cov}_{ts}\big[h_{t+1}(sB_{t+1}), B_{t+1}\big]\Big)$$

and the function h is given by the following recursive equation:

$$h_t(s) = \mathbb{E}_{ts}[h_{t+1}(sB_{t+1})(A_t + B_{t+1}^T\alpha_t(s))], \qquad t \in \{0, 1, \ldots, T-1\}$$

$$h_T(s) = 1,$$

with $\mathbb{E}_{tsx}[X_T^{\bar{\mathbf{u}}^*}] = \mathbb{E}[X_T^{\bar{\mathbf{u}}^*}|S_t = s, X_t = x] = h_t(s)x$.

Proof. In Markovian setting, the We will prove by backward induction that in the Markovian setting, for each t, there are functions $\alpha_t, h_t : (0,\infty)^n \to \mathbb{R}$ such that an equilibrium strategy is given by

$$\mathbf{u}_t^* = \alpha_t(S_t)X_t, \tag{3.9}$$

$$\mathbb{E}_{tsx}[X_T^{\bar{\mathbf{u}}^*}] = h_t(s)x. \tag{3.10}$$

We will use the following notation:

$$\hat{\Gamma}_t(s) = \mathrm{Cov}_{ts}[h_{t+1}(sB_{t+1})B_{t+1}, B_{t+1}],$$

$$\hat{\mu}_t(s) = \mathbb{E}_{ts}[B_{t+1}],$$

$$\hat{\beta}_t(s) = \mathrm{Cov}_{ts}[h_{t+1}(sB_{t+1}), B_{t+1}].$$

Note that $\hat{\Gamma}_t$ is an $n \times n$ matrix while the other two objects are column n-vectors.

Assume that the conditional expectation of the terminal wealth agrees the representation (3.10) for time $t+1$ (it trivially holds for T with $h_T(s) = 1$). The proof uses the characterisation (3.8) of an open-loop equilibrium. For that we have to compute β_t in terms of the quantities defined above. Notice that $X_T = D_t X_t + V_t$, hence $E_{tsx}[V_t] = (h_t(s) - D_t)x$. Therefore, $\beta_t = \mathrm{Cov}_t[\mathbb{E}_{t+1} V_{t+1}, \mathbf{B}_{t+1}]$ has the following representation

$$\begin{aligned}\beta_t &= \mathrm{Cov}_{tsx}\Big[\big(h_{t+1}(sB_{t+1}) - D_{t+1}\big)(Ax + B_{t+1}^T \mathbf{u}_t^*),\, \mathbf{B}_{t+1}\Big] \\ &= \hat{\Gamma}_t(s)\mathbf{u}_t^* + Ax\hat{\beta}_t(s) - D_{t+1}\Sigma_t \mathbf{u}_t^*.\end{aligned}$$

Inserting it into (3.8) and making use of the Markovian property of the price process gives

$$\mathbf{u}_t^* = \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \Big(\hat{\mu}_t(S_t) X_t - \gamma \hat{\Gamma}_t(S_t) \mathbf{u}_t^* - \gamma A \hat{\beta}_t(S_t) X_t + \gamma D_{t+1} \Sigma_t \mathbf{u}_t^* \Big).$$

This simplifies to

$$\mathbf{0} = \frac{1}{\gamma D_{t+1}} \Sigma_t^{-1} \Big(\hat{\mu}_t(S_t) X_t - \gamma \hat{\Gamma}_t(S_t) \mathbf{u}_t^* - \gamma A \hat{\beta}_t(S_t) X_t \Big).$$

Due to positive-definiteness of Σ_t, we conclude that the expression inside of the bracket has to be zero. Hence $\mathbf{u}_t^*$ satisfies

$$\mathbf{u}_t^* = \hat{\Gamma}_t^{-1}(S_t) \Big(\frac{1}{\gamma} \hat{\mu}_t(S_t) - A \hat{\beta}_t(S_t) \Big) X_t.$$

Therefore, $\mathbf{u}_t^*$ can be represented as a function of s and x: $\mathbf{u}_t^*(s,x) = \alpha_t(s)x$, where

$$\alpha_t(s) = \hat{\Gamma}_t^{-1}(s)\Big(\frac{1}{\gamma}\hat{\mu}_t(s) - A\hat{\beta}_t(s)\Big)$$

which coincides with the expression in the statement of Theorem 3.3.5.

It remains to prove the recursive formula for $h_t(s)$. Substituting the open-loop equilibrium control into the functional at time t, we obtain

$$\begin{aligned}
\mathbb{E}_{tsx}[X_T^{\bar{\mathbf{u}}^*}] &= \mathbb{E}_{tsx}[\mathbb{E}[X_T^{\bar{\mathbf{u}}^*}|S_{t+1} = sB_{t+1}, X_{t+1} = A_t x + B_{t+1}^T u_t^*]] \\
&= \mathbb{E}_{tsx}\big[h_{t+1}(s_t B_{t+1})(A_t x + B_{t+1}^T u_t^*)\big] \\
&= \mathbb{E}_{tsx}\big[h_{t+1}(s_t B_{t+1})(A_t + B_{t+1}^T \alpha_t(s))x\big] \\
&= h_t(s)x.
\end{aligned}$$

□

Corollary 3.3.6. *Assume B_t, $t \in \{1,\ldots,T\}$, are independent random variables with the mean μ_t and the covariance matrix Σ_t. The unique open-loop equilibrium strategy $\bar{\mathbf{u}}^*$ is given by*

$$\mathbf{u}_t^* = \frac{1}{\gamma h_{t+1}}\Sigma_t^{-1}\mu_t\, X_t, \tag{3.11}$$

where the sequence of real numbers h_t, $t = 1,\ldots,T$, satisfies

$$h_t = h_{t+1}A_t + \frac{1}{\gamma}\mu_t^T\Sigma_t^{-1}\mu_t.$$

Proof. Under the independence assumption, the result of Corollary 3.3.6 can be easily shown by Theorem 3.3.5. Hence, we only need to show that the equilibrium control found by Corollary 3.3.6 is admissible, that is $\mathbf{B}_t^T\mathbf{u}_t \in \mathcal{L}^2$ for any $t \in \{0,\ldots T-1\}$, i.e.

$$\mathbb{E}\left[(\mathbf{B}_{t+1}^T\mathbf{u}_t^*)^2\right] = \mathbb{E}\left[(\mathbf{B}_{t+1}^T\frac{h_{t+1}}{\gamma}\Sigma_t^{-1}\mu_t X_t^{\bar{\mathbf{u}}^*})^2\right]$$

Recall that

$$h_T = 1,$$

and

$$h_{T-1} = A_{T-1} + \frac{1}{\gamma}\mu_{T-1}^T \Sigma_{T-1}^{-1}\mu_{T-1}.$$

then

$$h_{t+1} = D_{t+1} + \frac{1}{\gamma}\sum_{i=t+1}^{T-1}\frac{D_{t+1}}{D_{i+1}}\mu_i^T\Sigma_i^{-1}\mu_i < \infty.$$

Under the independence assumption, Σ_t and μ_t are constant, and B_{t+1} is independent from the filtration $\mathcal{F}_t$, so is independent from X_t.

$$\begin{aligned}
\mathbb{E}\left[(\mathbf{B}_{t+1}^T\mathbf{u}_t^*)^2\right] &= \mathbb{E}\left[(\mathbf{u}_t^{*T}\mathbf{B}_{t+1})^2\right] \\
&= \big(\frac{h_{t+1}}{\gamma}\big)^2\,\mathbb{E}\left[\Sigma_t^{-1}\mu_t^T\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T\Sigma_t^{-1}\mu_t(X_t^{\bar{\mathbf{u}}^*})^2\right] \\
&= \big(\frac{h_{t+1}}{\gamma}\big)^2\Sigma_t^{-1}\mu_t^T\,\mathbb{E}\left[\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T(X_t^{\bar{\mathbf{u}}^*})^2\right]\Sigma_t^{-1}\mu_t \\
&= \big(\frac{h_{t+1}}{\gamma}\big)^2\Sigma_t^{-1}\mu_t^T\,\mathbb{E}\left[\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T\right]\Sigma_t^{-1}\mu_t\,\mathbb{E}\left[(X_t^{\bar{\mathbf{u}}^*})^2\right]
\end{aligned}$$

Since the mean and variance of $\mathbf{B}_t$ is well-defined, then the second moment is finite and

$$\big(\frac{h_{t+1}}{\gamma}\big)^2\Sigma_t^{-1}\mu_t^T\,\mathbb{E}\left[\mathbf{B}_{t+1}\mathbf{B}_{t+1}^T\right]\Sigma_t^{-1}\mu_t < \infty$$

Hence, we start with a constant initial wealth $X_0 = x_0$, it implies that the equilibrium control u_0 is admissible. As a result, the process X_1 will be square-integrable and $\mathbb{E}\left[(\mathbf{B}_2^T\mathbf{u}_1^*)^2\right] < \infty$. Again, this implies that u_1 is admissible. Therefore, we follow this procedure until the maturity T and obtain the admissible equilibrium control law $\{u_0, u_1, \ldots, u_{T-1}\}$. □

3.4 Convergence of discrete open-loop control for Black-Scholes model

In this section, we demonstrate the method for finding the open-loop equilibrium control in continuous setting through discrete approximation. For the sake of simplicity, we assume the risk-free rate is 0 and the dynamic of risky assets follows:

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W_t is a Brownian motion and $t \in [0, T]$. Denote u_t as the amount of money invested in stock market at time t, then the wealth process $(X_t)_{t\geq 0}$ follows:

$$dX_t = \frac{u_t}{S_t} dS_t.$$

Assume the investor aims to maximise the functional:

$$J(t, x_t; u) = x_t \, \mathbb{E}_t[X_T^u] - \frac{\gamma}{2} Var_t(X_T^u),$$

where γ is a positive constant.

We can discretise the time horizon into N partitions with equal sub-interval Δt. Denote $\Delta t = \frac{T}{N}$, then for $i \in \{0, \ldots, N-1\}$, the discretised dynamics of the stock price and wealth process follow:

$$S_{i\Delta t} = S_0 \exp\big((\mu - \frac{\sigma^2}{2}) i\Delta t + \sigma W_{i\Delta t}\big),$$

and

$$X_{(i+1)\Delta t} = X_{i\Delta t} + (\frac{S_{(i+1)\Delta t}}{S_{i\Delta t}} - 1) u_i.$$

We know $A = e^0 = 1$, and for $i \in \{0, \ldots, N-1\}$, we have

$$\begin{aligned} B_{i+1}^{\Delta t} = \frac{S_{(i+1)\Delta t}}{S_{i\Delta t}} - 1 &= \exp\big((\mu - \frac{\sigma^2}{2})\Delta t + \sigma(W_{(i+1)\Delta t} - W_{i\Delta t})\big) - 1 \\ &= \exp\big((\mu - \frac{\sigma^2}{2})\Delta t + \sigma \eta_{\Delta t}\big) - 1 \end{aligned}$$

where $\eta_{\Delta t}$ is normal random variable with mean 0 and variance Δt. We notice that $B_i^{\Delta t}$ is indeed identically distributed, then we will omit the index i. From the results Corollary 3.3.6 from previous section, the open-loop equilibrium strategy is given by

$$u_i^* = \frac{\mathbb{E}[B^{\Delta t}]}{\gamma h_{i+1} \mathrm{Var}(B^{\Delta t})} x_i \tag{3.12}$$

and

$$h_i = h_{i+1} + \frac{\mathbb{E}[B^{\Delta t}]^2}{\gamma Var(B^{\Delta t})}, \quad \text{for } i \in \{0, \dots, N-1\}.$$

Denote that $\xi_{\Delta t} = \frac{\mathbb{E}[B^{\Delta t}]^2}{\gamma Var(B^{\Delta t})}$, then

$$\begin{aligned} h_i &= h_{i+1} + \xi_{\Delta t} \\ &= h_T + (N-i)\xi_{\Delta t} \\ &= 1 + (N-i)\xi_{\Delta t} \end{aligned}$$

for $i \in \{0, \dots, T-1\}$, and

$$\begin{aligned} \xi_{\Delta t} &= \frac{\mathbb{E}[\frac{S_{(i+1)\Delta t}}{S_{i\Delta t}} - 1]^2}{\gamma Var(\frac{S_{(i+1)\Delta t}}{S_{i\Delta t}} - 1)} \\ &= \frac{\mathbb{E}\big[\exp((\mu - \frac{\sigma^2}{2}) + \sigma\eta_{\Delta t}) - 1\big]^2}{\gamma Var\big(\exp((\mu - \frac{\sigma^2}{2}) + \sigma\eta_{\Delta t})\big)} \\ &= \frac{(e^{\mu\Delta t} - 1)^2}{\gamma e^{2\mu\Delta t}\{e^{\sigma^2\Delta t} - 1\}}. \end{aligned}$$

For any $t \in [0, T]$, denote $k = \lfloor \frac{t}{\Delta t} \rfloor$, then $k \cdot \Delta t \leq t < (k+1) \cdot \Delta t$. Next, we will show that as N tends to infinity, h_k and h_{k+1} will tend to same limit,

$$\begin{aligned} \lim_{N\to\infty} h_k &= \lim_{\Delta t\to 0} h_k \\ &= 1 + \lim_{\Delta t\to 0} \xi_{\Delta t}(N-k) \\ &= 1 + \frac{1}{\gamma} \lim_{\Delta t\to 0} \frac{\mathbb{E}[B^{\Delta t}]^2}{Var(B^{\Delta t})} \cdot \big(\frac{T}{\Delta t} - k\big) \\ &= 1 + \frac{1}{\gamma} \lim_{\Delta t\to 0} \frac{\mathbb{E}[B^{\Delta t}]^2}{\Delta t \cdot Var(B^{\Delta t})} \cdot \big(T - k \cdot \Delta t\big) \\ &= 1 + \frac{1}{\gamma} \lim_{\Delta t\to 0} \frac{(e^{\mu\Delta t} - 1)^2}{\Delta t \cdot e^{2\mu\Delta t}\{e^{\sigma^2\Delta t} - 1\}} \cdot \big(T - k \cdot \Delta t\big). \end{aligned}$$

By L'Hôpital's rule, we have

$$\begin{aligned}
\lim_{\Delta t \to 0} \frac{(e^{\mu\Delta t}-1)^2}{\Delta t e^{2\mu\Delta t}\{e^{\sigma^2\Delta t}-1\}} &= \lim_{\Delta t \to 0} \frac{1-2e^{-\mu\Delta t}+e^{-2\mu\Delta t}}{(e^{\sigma^2\Delta t}-1)\cdot\Delta t} \\
&\overset{\text{L'Hôpital}}{=} \lim_{\Delta t \to 0} \frac{2\mu e^{-\mu\Delta t}-2\mu e^{-2\mu\Delta t}}{\sigma^2 e^{\sigma^2\Delta t}\Delta t + e^{\sigma^2\Delta t}-1} \\
&\overset{\text{L'Hôpital}}{=} \lim_{\Delta t \to 0} \frac{-2\mu^2 e^{-\mu\Delta t}+4\mu^2 e^{-2\mu\Delta t}}{\Delta t\cdot\sigma^4 e^{\sigma^2\Delta t}+\sigma^2 e^{\sigma^2\Delta t}+\sigma^2 e^{\sigma^2\Delta t}} \\
&= \frac{\mu^2}{\sigma^2}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\lim_{N\to\infty} h_k^{\Delta t} &= 1 + \frac{\mu^2}{\gamma\sigma^2}\lim_{\Delta t\to 0}(T - k\cdot\Delta t) \\
&= 1 + \frac{\mu^2}{\gamma\sigma^2}\lim_{\Delta t\to 0}\big(T - \lfloor\frac{t}{\Delta t}\rfloor\cdot\Delta t\big) \\
&= 1 + \frac{\mu^2}{\gamma\sigma^2}(T-t).
\end{aligned}$$

Similarly, since $\xi_{\Delta t}$ does not depend on the index k, then

$$\begin{aligned}
\lim_{N\to\infty} h_{k+1} &= 1 + \frac{\mu^2}{\gamma\sigma^2}\lim_{\Delta t\to 0}\big(T-(k+1)\cdot\Delta t\big) \\
&= 1 + \frac{\mu^2}{\gamma\sigma^2}(T-t).
\end{aligned}$$

As N goes to infinity, the mapping $t \to h_t$ is now in continuous time, then

$$h_t = 1 + \frac{\mu^2}{\gamma\sigma^2}(T-t),$$

and

$$\begin{aligned}
u_t^* &= \frac{\mu}{\gamma\cdot(1+\frac{\mu^2}{\gamma\sigma^2}(T-t))\sigma^2}X_t \\
&= \frac{\mu}{\gamma\sigma^2+\mu^2(T-t)}X_t. \qquad (3.13)
\end{aligned}$$

Remark. In (Hu *et al.*, 2012, Section 5.4.1), the continuous open-loop equilibrium control for the same problem is found to be:

$$u_t^* = \frac{\mu}{\gamma\, M_t\, \sigma^2} X_t \tag{3.14}$$

where

$$M_t = 1 + \frac{\mu^2}{\gamma\sigma^2}(T - t).$$

which is the limit of the control for discrete framework as shown above.

3.5 Conclusion

This chapter has introduced a definition of discrete time open-loop equilibrium strategy. Under the discrete time open-loop definition 3.3.1, the relationship perturbed wealth process and unperturbed wealth process has been identified by equation (3.3). From the equation (3.3) the rationale of the open-loop equilibrium is that the investor tests a perturbed strategy against current investment plan by deviating a small amount of stock investment ε over a single-period. In particular, the investor has two objectives: short-term investment and long-term investment plans. The short-term investment plan (e.g. the term $D_{t+1}\mathbf{B}_{t+1}^T\varepsilon$ in equation (3.3)) for open-loop equilibrium investor is to invest in stock market for one period. After one period, the investor will withdraw this investment from stock market and deposit it into the bank account until the maturity.

The second focus of the research is to obtain the necessary and sufficient condition (Theorem 3.3.2 and Theorem 3.3.3) for open-loop equilibrium strategy. It has been shown that the type of perturbation does not affect the resulting equilibrium strategy in discrete time. Moreover, given the specific matrix inversion is well-defined,Theorem 3.3.3 shows that the open-loop equilibrium strategy is of linear feedback type and is unique. Therefore, there exists only one open-loop strategy for multiplicative wealth dynamic. The feasibility of the solution has been provided by Corollary 3.3.4 with bounded excess return factor $\mathbf{B}_t$.

The condition obtained for open-loop can be used for generating a numerical algorithm for constructing an open-loop equilibrium control. It has been shown that the open-loop equilibrium strategy we developed in this chapter approximates the continuous version of open-loop equilibrium strategy by Hu *et al.* (2012) in Black-Scholes model. Providing the uniqueness result for continuous open-loop equilibrium strategy from Hu *et al.* (2015), it verifies that our interpretation of the open-loop investor's rationale coincides with the original open-loop equilibrium definition. However, for the Black-Scholes example presenting in this chapter, it has be noticed that the value of M_t strictly increases as $(T - t)$ increases. This implies that the open-loop equilibrium strategy u_t^* in equation (3.14) decays as time t goes backward. We will discuss these issues in Chapter 6.2.

Chapter 4

Closed-loop equilibrium strategy in discrete time

4.1 Introduction

In Bjork & Murgoci (2014), the authors study the closed-loop equilibrium strategy in discrete time. Since the closed-loop was first developed for Markovian system, the perturbation is defined as the deterministic function of current time and wealth state. It is expected that, when the investor deviates from the strategy slightly, there exists a chain reaction on the future perturbations. Based on such hypothesis, we will present a similar perturbation method as shown in Chapter 3. As mentioned in Chapter 3, the investment amount of initial open-loop equilibrium strategy goes to 0 as the investment length goes to infinity. A similar phenomenon has been observed for closed-loop equilibrium strategy in all existing studies. Therefore, this chapter will study the asymptotic behaviour of closed-loop equilibrium strategy as investment length goes to infinity.

4.2 Formulation of the problem and preliminaries

In this chapter, we will study the same problem as in Chapter 3.2. Recall the aim of investor from equation 3.2:

$$\operatorname*{ess\,sup}_{\bar{\mathbf{u}}} \left\{ X_t \, \mathbb{E}_t[X_T^{\bar{\mathbf{u}}}] - \frac{\gamma}{2} \mathrm{Var}_t[X_T^{\bar{\mathbf{u}}}] \right\}, \tag{4.1}$$

subject to the wealth dynamic

$$X_{t+1}^{\bar{\mathbf{u}}} = A_t X_t^{\bar{\mathbf{u}}} + \mathbf{B}_{t+1}^T \mathbf{u}_t, \tag{4.2}$$

where $A_t = e^{r_t^0}$ and $\mathbf{B}_{t+1} := [B_{t+1}^1, B_{t+1}^2, \dots, B_{t+1}^n]^T = [(b_{t+1}^1 - e^{r_t^0}), (b_{t+1}^2 - e^{r_t^0}), \dots, (b_{t+1}^n - e^{r_t^0})]^T$. However, compared to the previous chapter, we will study the problem within the Markovian framework. Markovian framework implies the stock price admitting a strong Markov property. Since under the strong Markovian property, the stock price at time t is independent of the prices at time s with $s < t$. This suggests us to concentrate on the feedback type of the strategies. Now we will introduce the definition of feedback strategy and the equilibrium condition. Also, we will use the upper case $\bar{\mathbf{U}}$ to denote the strategy of feedback type in Markovian setting to distinguish it from the notation of open-loop strategy in general setting. Recall that $(\mathcal{F}_t)$ is the filtration generated by stock prices $(S_k)_{k\in 0,\dots,t}$.

Definition 4.2.1. A sequence $\bar{\mathbf{U}}_t \equiv (U_t, \dots, U_{T-1})$ of random functions: $\Omega \times \mathbb{R} \to \mathbb{R}^n$ in is a admissible feedback strategy for an initial wealth $x_t \in \mathcal{L}^2(\mathcal{F}_t)$ if

i U_s is $\frac{\mathcal{F}_s \otimes \mathbb{B}(\mathbb{R})}{\mathbb{B}(\mathbb{R}^n)}$-measurable for $s = t, \dots, T-1$.

ii the wealth process defined as $X_{s+1}^{\bar{\mathbf{U}}} = A_s X_s^{\bar{\mathbf{U}}} + B_{s+1}^T U_s(X_s^{\bar{\mathbf{U}}})$ satisfies $X_s^{\bar{\mathbf{U}}} \in \mathcal{L}^2(\mathcal{F}_s)$ for $s = t, \dots, T-1$.

If stock price is a Markov process, then we introduce an obvious specialisation of the above definition. For the sake of simplicity, notations remain the same as previously assigned.

Definition 4.2.2. A sequence $\bar{\mathbf{U}}_t \equiv (U_t, \dots, U_{T-1})$ of function $\mathcal{R} \times (0, \infty)^n \to \mathcal{R}^n$ is a admissible Markovian feedback strategy for an initial wealth $X_t \in \mathcal{L}^2(\mathcal{F}_t)$ if

i U_s is $\frac{\mathbb{B}(\mathbb{R}) \otimes \mathbb{B}(\mathbb{R}^n)}{\mathbb{B}(\mathbb{R}^n)}$-measurable for $s = t, \dots, T-1$.

ii the wealth process defined as $X_{s+1}^{\bar{\mathbf{U}}} = A_s X_s^{\bar{\mathbf{U}}} + B_{s+1}^T U_s(X_s^{\bar{\mathbf{U}}})$ satisfies $X_s^{\bar{\mathbf{U}}} \in \mathcal{L}^2(\mathcal{F}_s)$ for $s = t, \dots, T-1$.

Definition 4.2.3. The wealth process $(X_s^{\bar{\mathbf{U}}_{t,x_t}})_{s=t,\dots,T}$ for a feedback strategy $\bar{\mathbf{U}}_t$ and initial wealth $X_t \in \mathcal{L}^2(\mathcal{F}_t)$ is defined as follows:

$$X_t^{\bar{\mathbf{U}}_{t,x_t}} = X_t,$$

$$X_{s+1}^{\bar{\mathbf{U}}_{t,x_t}} = A_s X_s^{\bar{\mathbf{U}}_{t,x_t}} + B_{s+1}^T U_s(X_s^{\bar{\mathbf{U}}_{t,x_t}}), \quad s = t, \dots, T-1.$$

For the simpilicity of notation, we will often omit the dependence on X_t and $\bar{\mathbf{U}}_t$.

Definition 4.2.4. For an admissible feedback strategy $\bar{\mathbf{U}}_t$ and an intial wealth x_t, its performance is defined as

$$J(t, x_t; \bar{\mathbf{U}}_t) = X_t \, \mathbb{E}_t[X_T^{\bar{\mathbf{U}}_{t,x_t}}] - \frac{\gamma}{2} \mathrm{Var}_t[X_T^{\bar{\mathbf{U}}_{t,x_t}}].$$

Definition 4.2.5. An admissible feedback strategy $\bar{\mathbf{U}}^* \equiv \bar{\mathbf{U}}_0^*$ for the initial wealth x_0 is a closed-loop equilibrium control if for any $t \in \{0, 1, \dots, T-1\}$ and for any bounded $\mathbb{R}^n$-valued $\mathcal{F}_t$-measurable random variable ε, we have

$$J(t, x_t^{\bar{\mathbf{U}}^*}; \bar{\mathbf{U}}_t^{\varepsilon}) - J(t, x_t^{\bar{\mathbf{U}}^*}; \bar{\mathbf{U}}_t^*) \leq 0 \quad a.s.,$$

where $\bar{\mathbf{U}}^* = (U_t^*, U_{t+1}^* \dots, U_{T-1}^*)$ and $\bar{\mathbf{U}}^{\varepsilon} = (U_t^* + \varepsilon, U_{t+1}^*, \dots, U_{T-1}^*)$.

4.3 Closed-loop strategy for mean-variance problem

Note that the definition of closed-loop equilibrium stands in sharp contrast with open-loop equilibrium from the perturbed strategy. The deviation term ε at time t will affect the future strategies $\{\mathbf{U}_s\}$ for $s = t+1, \dots, T-1$. An explanation for this difference is: when the open-loop investor decides to test the strategy by a perturbation at time t, due to the nature of general dynamic system, he could not examine the future influence arising by this perturbation after time $t+1$. Therefore, he would like to have a "blind" test against the present control strategy. However, in Markovian system, the future dynamic generated by the perturbation can be tracked and the investor can adjust the future strategies corresponding to this perturbation.

Theorem 4.3.1. *Let control sequence $\{\bar{\mathbf{U}}_s(X_s^{\bar{\mathbf{U}}}, S_s)\}_{s=0,\dots,T-1}$ be a closed-loop equilibrium control for problem (4.1) under the wealth dynamic (4.2). Then the equilibrium control $\bar{\mathbf{U}}$ admits a separable form of*

$$\bar{\mathbf{U}}_t = \mathbf{K}(t)X_t, \qquad \text{for } t = 0, \dots, T-1.$$

Furthermore, let

$$K_{T-1} = \frac{1}{\gamma}\,\mathbb{E}_{T-1}[B_T]\mathrm{Cov}_{T-1}^{-1}(B_T),$$

and, for $t = 0, 1, \dots, T-2$,

$$K_t = \mathrm{Cov}_t^{-1}(G_{t+1}B_{t+1})\big(\frac{1}{\gamma}\,\mathbb{E}_t[B_{t+1}G_{t+1}] - A_t\mathrm{Cov}_t\big)$$

where

$$G_t = \prod_{i=t}^{T-1}(A_i + \mathbf{B}_{i+1}^T\mathbf{K}_i), \quad \text{for } t = 1, 2, \dots, T-1. \tag{4.3}$$

Assume that (K_t) is well-defined, i.e., $\mathrm{Cov}_{T-1}(B_T)$ and $\mathrm{Cov}_t(G_{t+1}B_{t+1})$ are a.s. invertible and all the expectations exist. Define a feedback control $\bar{\mathbf{U}}^* = (U_0^*, U_1^*, \dots, U_{T-1}^*)$ by

$$U_t^*(x) = K_t x.$$

i If $X_T^{\bar{\mathbf{U}}^{*,1}} \in \mathcal{L}^2(\mathcal{F}_T)$, *then* $\bar{\mathbf{U}}^*$ *is a unique equilibrium closed-loop feedback control for any initial wealth* $x_0 \in \mathbb{R} \setminus \{0\}$.

ii If $X_T^{\bar{\mathbf{U}}^{*,1}} \notin \mathcal{L}^2(\mathcal{F}_T)$, *then there is no equilibrium closed-loop feedback control for any* $x_0 \in \mathbb{R} \setminus \{0\}$.

Proof. We will prove the above result using mathematical induction. At time $T-1$, recall $D_T = 1$, the closed-loop equilibrium strategy coincides with the single-period optimal control, which is given by

$$\mathbf{U}_{T-1} = \frac{\mu_{T-1}}{\gamma} \Sigma_{T-1}^{-1} X_{T-1},$$

where $\Sigma_{T-1} = \mathrm{Cov}_{T-1}[\mathbf{B}_T]$ and $\mu_{T-1} = \mathbb{E}_{T-1}[\mathbf{B}_T]$. Therefore, at time $T-1$, the closed-loop equilibrium control is in the form:

$$\mathbf{U}_{T-1} = \mathbf{K}_{T-1} X_{T-1},$$

where $\mathbf{K}_{T-1} = \frac{1}{\gamma}\mu_{T-1}\Sigma_{T-1}^{-1}$. Assume $\mathbf{U}_s$ follow the expression $\mathbf{K}_s X_s$ for all $s \in \{t+1, \dots, T-1\}$, then the dynamics (4.2) allow for explicit expression for the terminal wealth. For convenience, we also set $D_t = \prod_{i=1}^{T-1} A_i$ and recall that is a deterministic quantity. Given $X_t \in L^2(\mathcal{F}_t)$ at time t and a separable admissible control $\bar{\mathbf{U}}_t$, for a perturbed control $\bar{\mathbf{U}}_t^{\varepsilon}$ we have

$$\begin{aligned}
X_T^{\bar{\mathbf{U}}_T^{\varepsilon}} &= A_{T-1} X_{T-1}^{\bar{\mathbf{U}}^{\varepsilon}} + \mathbf{B}_T^T \mathbf{U}_{T-1}^{\varepsilon}(X_{T-1}^{\bar{\mathbf{U}}^{\varepsilon}}) = (A_{T-1} + \mathbf{B}_T^T \mathbf{K}_{T-1}) X_{T-1}^{\bar{\mathbf{U}}^{\varepsilon}} \\
&= \prod_{i=t+1}^{T-1} (A_i + \mathbf{B}_{i+1}^T \mathbf{K}_i) X_{t+1}^{\bar{\mathbf{U}}^{\varepsilon}} \\
&= \prod_{i=t+1}^{T-1} (A_i + \mathbf{B}_{i+1}^T \mathbf{K}_i)(A_t X_t + \mathbf{B}_{t+1}^T \mathbf{U}_t^{\varepsilon}) \\
&= \prod_{i=t+1}^{T-1} (A_i + \mathbf{B}_{i+1}^T \mathbf{K}_i)\big(A_t X_t + \mathbf{B}_{t+1}^T (\mathbf{U}_t + \varepsilon)\big) \\
&= X_T^{\bar{\mathbf{U}}} + \mathbf{B}_{t+1}^T \varepsilon \prod_{i=t+1}^{T-1} (A_i + \mathbf{B}_{i+1}^T \mathbf{K}_i) \qquad (4.4)
\end{aligned}$$

Here,to obtain the expression (4.4), we use the induction hypothesis that the strategies $(U_{t+1}, \dots, U_{T-1})$ admit the form $U_s = \mathbf{K}_s X_s$ for $s \in \{t+1, \dots, T-1\}$.

Next, we will prove the closed-loop equilibrium control at time t obtains the same expression. Fix an admissible strategy $\bar{\mathbf{U}}$ and its perturbed version $\bar{\mathbf{U}}^{\varepsilon}$ with a deterministic perturbation $\varepsilon \in \mathbb{R}^n$ applied at time t. Recalling expression (4.4)-(4.3), we have

$$E_t[X_T^{\bar{\mathbf{U}},\varepsilon}] = E_t[X_T^{\bar{\mathbf{U}}}] + E_t[\mathbf{B}_{t+1}^T G_{t+1}]\varepsilon.$$

For variance term, we have

$$\begin{aligned}\mathrm{Var}_t[X_T^{\bar{\mathbf{U}}^{\varepsilon}}] &= \mathrm{Var}_t[X_T^{\bar{\mathbf{U}}} + \mathbf{B}_{t+1}^T G_{t+1}\varepsilon] \\ &= \mathrm{Var}_t[X_T^{\bar{\mathbf{U}}}] + 2\varepsilon^T \mathrm{Cov}_t(X_T^{\bar{\mathbf{U}}}, \mathbf{B}_{t+1}^T G_{t+1}) + \varepsilon^T \mathrm{Cov}_t[\mathbf{B}_{t+1} G_{t+1}]\varepsilon\end{aligned}$$

By the definition 4.2.5, we have

$$\begin{aligned}&J(t, X_t^{\bar{\mathbf{U}}}; \bar{\mathbf{U}}^{\varepsilon}) - J(t, X_t^{\bar{\mathbf{U}}}; \bar{\mathbf{U}}) \\ &= X_t E_t[\mathbf{B}_{t+1}^T G_{t+1}]\varepsilon - \gamma\varepsilon^T \mathrm{Cov}_t(X_T^{\bar{\mathbf{U}}}, \mathbf{B}_{t+1}^T G_{t+1}) - \frac{\gamma}{2}\varepsilon^T \mathrm{Cov}_t[\mathbf{B}_{t+1} G_{t+1}]\varepsilon. \quad (4.5)\end{aligned}$$

For $\bar{\mathbf{U}}$ to be a closed-loop equilibrium strategy, the above expression has to be a.s. non-positive for any ε. We assume $\mathrm{Cov}_{T-1}(B_T)$ and $\mathrm{Cov}_t(G_{t+1}B_{t+1})$ are a.s. invertible and all the expectations exist. The above expression is strictly concave in ε. The condition of non-positivity is satisfied if and only if the maximum is at $\varepsilon = \mathbf{0}$. Due to the concavity, the first order condition is necessary and sufficient. The derivative with respect to ε evaluated at $\varepsilon = 0$ reads

$$X_t E_t[\mathbf{B}_{t+1} G_{t+1}] - \gamma \mathrm{Cov}_t(X_T^{\bar{\mathbf{U}}}, \mathbf{B}_{t+1}^T G_{t+1}) = \mathbf{0} \quad (4.6)$$

Recall that the structure of separable control $\mathbf{U}_s = \mathbf{K}_s X_s$ for $s \in \{t+1, \dots, T-1\}$, then given $X_t \in L^2(\mathcal{F}_t)$ at time t and an admissible control $\bar{\mathbf{U}}_t$ we have

$$X_T = G_{t+1}(A_t X_t + \mathbf{B}_{t+1}^T \mathbf{U}_t)$$

By the first order condition (4.6), we have

$$\begin{aligned}\mathbf{0} &= X_t^{\bar{\mathbf{U}}}\,\mathbb{E}_t[\mathbf{B}_{t+1}G_{t+1}] - \gamma \mathrm{Cov}_t\big(G_{t+1}(A_tX_t + \mathbf{B}_{t+1}^T\mathbf{U}_t), \mathbf{B}_{t+1}^T G_{t+1}\big)\\ &= X_t^{\bar{\mathbf{U}}}\,\mathbb{E}_t[\mathbf{B}_{t+1}G_{t+1}] - X_tA_t\gamma \mathrm{Cov}_t\big(G_{t+1}, \mathbf{B}_{t+1}^T G_{t+1}\big)\\ &\quad - \gamma \mathrm{Cov}_t\big(G_{t+1}\mathbf{B}_{t+1}^T, \mathbf{B}_{t+1}^T G_{t+1}\big)\mathbf{U}_t\end{aligned}$$

Since $\mathrm{Cov}_t\big(G_{t+1}\mathbf{B}_{t+1}^T, \mathbf{B}_{t+1}^T G_{t+1}\big)$ is positive definite, then $\mathbf{U}_t$ has to satisfy

$$\mathbf{U}_t = \mathrm{Cov}_t^{-1}\big(G_{t+1}\mathbf{B}_{t+1}\big)\big(\frac{1}{\gamma}\,\mathbb{E}_t[\mathbf{B}_{t+1}G_{t+1}] - A_t\mathrm{Cov}_t(G_{t+1}, \mathbf{B}_{t+1}^T G_{t+1})\big)X_t.$$

Inserting it into (4.5) gives

$$J(t, X_t^{\bar{\mathbf{U}}}; \bar{\mathbf{U}}_t^{\varepsilon}) - J(t, X_t^{\bar{\mathbf{U}}}; \bar{\mathbf{U}}_t) = -\frac{\gamma}{2}\varepsilon^T \mathrm{Var}_t[\mathbf{B}_{t+1}G_{t+1}]\varepsilon,$$

which is non-positive for all ε. Now, we relax ε to be a bounded $\mathcal{F}_t$-measurable random variable. Due to positive definiteness of $\mathrm{Var}_t[\mathbf{B}_{t+1}G_{t+1}]$, the above expression is non-positive if and only if $\varepsilon \equiv \mathbf{0}$.

It can be seen that, the closed-loop equilibrium control $\mathbf{U}_t = \mathbf{K}_tX_t$ where

$$\mathbf{K}_t = \mathrm{Cov}_t^{-1}\big(G_{t+1}\mathbf{B}_{t+1}\big)\big(\frac{1}{\gamma}\,\mathbb{E}_t[\mathbf{B}_{t+1}G_{t+1}] - A_t\mathrm{Cov}_t(G_{t+1}, \mathbf{B}_{t+1}^T G_{t+1})\big).$$

Therefore, by the mathematical induction, the equilibrium strategy $\bar{\mathbf{U}}$ admits a separable form of

$$\bar{\mathbf{U}}_t = \mathbf{K}(t)X_t, \qquad \text{for } t = 0, \dots, T-1.$$

□

Remark. It can be seen in Theorem 4.3.1 that the closed-loop equilibrium control at time t depends on the entire future choices for $\{t+1, \dots, T-1\}$ via $\mathbf{K}$. However, the investor decides the closed-loop control recursively. Compared with open-loop equilibrium control, the separable structure and the recursive construction ensure the uniqueness of the solution. Therefore, the following result can be concluded.

Corollary 4.3.2. *There exists a unique closed-loop equilibrium control* $\bar{\mathbf{U}}_t^*$ *satisfying*

$$\mathbf{U}_t^* = \mathrm{Cov}_t(G_{t+1}\mathbf{B}^T, G_{t+1})^{-1}\big(\frac{1}{\gamma}\,\mathbb{E}_t[\mathbf{B}_{t+1}G_{t+1}] - A_t\mathrm{Cov}_t(G_{t+1})\big)X_t^{\bar{\mathbf{U}}},$$

where

$$G_t = \prod_{i=t}^{T-1}(A_i + \mathbf{B}_{i+1}^T\mathbf{K}_i).$$

Corollary 4.3.3. *Assume* $\mathbf{B}_t$*,* $t \in \{1,\ldots,T\}$ *are independent random variables with the mean* μ_t *and the covariance matrix* Σ_t*. A unique closed-loop equilibrium strategy* $\bar{\mathbf{U}}^*$ *is given by*

$$\mathbf{U}_t^* = \Big(\beta_{t+1}(\Sigma_t + \mu_t\mu_t^T) - \alpha_{t+1}^2\mu_t\mu_t^T\Big)^{-1}\Big(\frac{1}{\gamma}\alpha_{t+1}\mu_t - A_t(\beta_{t+1} - \alpha_{t+1}^2)\mu_t\Big)X_t^{\bar{\mathbf{U}}} \tag{4.7}$$

where the function α *and* β *are given by the following recursive equations*

$$\alpha_t = (A_t + \mu_t^T\mathbf{K}_t)\alpha_{t+1}, \quad \alpha_T = 1,$$

and

$$\beta_t = (A_t^2 + \mathbf{K}_t^T(\Sigma_t + \mu_t\mu_t^T)\mathbf{K}_t + 2A_t\mu_t^T\mathbf{K}_t)\beta_{t+1}, \quad \beta_T = 1.$$

Proof. If $\mathbf{B}_i$'s are independent, then

$$\mathbb{E}_t[G_{t+1}\mathbf{B}_{t+1}] = \prod_{i=t+1}^{T-1}(A_i + \mathbb{E}[\mathbf{B}_{i+1}]^T\mathbf{K}_i])E[\mathbf{B}_{t+1}]$$

and

$$\text{Cov}_t(G_{t+1}, G_{t+1}\mathbf{B}_{t+1}) = E[G_{t+1}^2]E[\mathbf{B}_{t+1}] - (E[G_{t+1}])^2 E[\mathbf{B}_{t+1}].$$

Denote $\alpha_{t+1} := E[G_{t+1}]$ and $\beta_{t+1} := E[G_{t+1}^2]$, then

$$\alpha_t = E[G_t] = E[(A_t + \mathbf{B}_{t+1}^T \mathbf{K}_t)G_{t+1}] = (A_t + \mu_t^T \mathbf{K}_t)\alpha_{t+1}$$

and

$$\begin{aligned}\beta_t = E[G_t^2] &= E[(A_t + \mathbf{B}_{t+1}\mathbf{K}_t)^2 G_{t+1}^2] \\ &= (A_t^2 + \mathbf{K}_t^T(\Sigma_t + \mu_t\mu_t^T)\mathbf{K}_t + 2A_t\mu_t^T\mathbf{K}_t)\beta_{t+1}\end{aligned}$$

with

$$\alpha_T = 1 \quad \text{and} \quad \beta_T = 1.$$

Hence, the closed-loop equilibrium control becomes

$$\mathbf{U}_t^* = \Big(\beta_{t+1}(\Sigma_t + \mu_t\mu_t^T) - \alpha_{t+1}^2\mu_t\mu_t^T\Big)^{-1}\Big(\frac{1}{\gamma}\alpha_{t+1}\mu_t - A_t(\beta_{t+1} - \alpha_{t+1}^2)\mu_t\Big)X_t^{\bar{\mathbf{U}}}$$

□

Remark 4.3.4. Suppose there is one risky asset with independent and identical increment, then 4.3.3 is the same as the construction in Bjork & Murgoci (2014).

4.4 Asymptotic behaviour of closed-loop control

Corollary 4.3.3 provides a recursive way to construct a closed-loop equilibrium control. Such construction allows us to determine the control for infinite time backwards. Since the equilibrium strategy is designed for long-term investors, it is important to investigate the behaviour as time goes back. For the sake of simplicity, we will assume that there is only one risky asset where the log returns are independently and identically distributed. Meanwhile, the risk-free asset has a constant return which implies $A_t = e^{r_t^0} = A = \text{constant} > 1$.

Recall $\mathbf{U}_t$ is a closed-loop equilibrium strategy in the form of

$$U_t = K_t\, X_t.$$

where t denote the time to maturity T. In the independent and identical case,

$$K_{t+1} = \frac{\frac{1}{\gamma}\alpha_t\mu - A\big\{\beta_t\mu - \alpha_t^2\mu\big\}}{\beta_t(\sigma^2+\mu^2) - \alpha_t^2\mu^2}$$

where

$$\alpha_{t+1} = (A + \mu K_{t+1})\alpha_t, \quad \alpha_0 = 1$$

and

$$\beta_{t+1} = \Big(A^2 + (\sigma^2+\mu^2)K_{t+1}^2 + 2A\mu K_{t+1}\Big)\beta_t = \Big((A+\mu K_{t+1})^2 + \sigma^2 K_{t+1}^2\Big)\beta_t, \quad \beta_0 = 1.$$

From here, in order to simplify the notations, we can rewrite K_{t+1} by dividing both the numerator and denominator by μ^2,

$$K_{t+1} = \frac{1}{\mu}\cdot\frac{\frac{1}{\gamma}\alpha_t - A\big\{\beta_t - \alpha_t^2\big\}}{\beta_t(\frac{\sigma^2}{\mu^2}+1) - \alpha_t^2},$$

and also write α and β as,

$$\alpha_{t+1} = \Big(A + \frac{\frac{1}{\gamma}\alpha_t - A\big\{\beta_t - \alpha_t^2\big\}}{\beta_t(\frac{\sigma^2}{\mu^2}+1) - \alpha_t^2}\Big)\alpha_t,$$

and

$$\beta_{t+1} = \Big(\big(A + \frac{\frac{1}{\gamma}\alpha_t - A\big\{\beta_t - \alpha_t^2\big\}}{\beta_t(\frac{\sigma^2}{\mu^2}+1) - \alpha_t^2}\big)^2 + \frac{\sigma^2}{\mu^2}\big(\frac{\frac{1}{\gamma}\alpha_t - A\big\{\beta_t - \alpha_t^2\big\}}{\beta_t(\frac{\sigma^2}{\mu^2}+1) - \alpha_t^2}\big)^2\Big)\beta_t.$$

Denote

$$\phi = \frac{\sigma^2}{\mu^2}, \tag{4.8}$$

then

$$K_{t+1} = \frac{d_{t+1}}{\mu}, \tag{4.9}$$

where

$$d_{t+1} = \frac{\frac{1}{\gamma}\alpha_t - A\{\beta_t - \alpha_t^2\}}{\beta_t(\phi+1) - \alpha_t^2}, \tag{4.10}$$

then

$$\alpha_{t+1} = (A + d_{t+1})\alpha_t \tag{4.11}$$

and

$$\beta_{t+1} = \Big((A + d_{t+1})^2 + \phi\, d_{t+1}^2\Big)\beta_t \tag{4.12}$$

Note that the behaviour of closed-loop equilibrium control K_{t+1} is the same as the behaviour of d_{t+1} by a constant scaling factor. The main difficulty is the nested dependence between the variables d, α, β and a constant ϕ. It can be seen that ϕ is the inverse value of Sharpe ratio square. Therefore, the behaviour of closed-loop control depends on the value of the Sharpe ratio. Note that the expression (4.10) can be also written as:

$$d_{t+1} = \frac{\frac{1}{\gamma}\frac{\alpha_t}{\beta_t} - A\{1 - \frac{\alpha_t^2}{\beta_t}\}}{(\phi+1) - \frac{\alpha_t^2}{\beta_t}}. \tag{4.13}$$

Lemma 4.4.1. *$\frac{\alpha_t^2}{\beta_t}$ is monotonically decreasing as t increases, that is $\frac{\alpha_{t+1}^2}{\beta_{t+1}} \leq \frac{\alpha_t^2}{\beta_t}$ for all t and the equality holds iff $d_1 = 0$. Also, there exists a constant $c_1 < 1$*

such that

$$\frac{\alpha_t^2}{\beta_t} \leq \frac{\alpha_1^2}{\beta_1} = c_1 < 1.$$

Proof. Since $\phi d_t^2 \geq 0$, then

$$\frac{\alpha_{t+1}^2}{\beta_{t+1}} = \frac{(A+d_{t+1})^2}{(A+d_{t+1})^2 + \phi d_{t+1}^2} \frac{\alpha_t^2}{\beta_t} \leq \frac{\alpha_t^2}{\beta_t}$$

and the above equality holds only if $d_{t+1} = 0$. It is easy to verify that the base value

$$d_1 = \frac{1}{\gamma \phi} \neq 0.$$

So,

$$\frac{\alpha_1^2}{\beta_1} = \frac{(A+d_1)^2}{(A+d_1)^2 + \phi d_1^2} \frac{\alpha_0^2}{\beta_0} < \frac{\alpha_0^2}{\beta_0} = 1.$$

□

Remark 4.4.1. Lemma 4.4.1 shows the ratio of $\frac{\alpha_t^2}{\beta_t}$ is strictly less than 1 for all t and will always decrease unless the value $d_{t+1} = 0$. Therefore, the denominator of equation (4.13) is strictly positive. The sign of d_t only depends on the numerator of equation (4.13)

Lemma 4.4.2. *There exists a t^* such that $d_t < 0$ for all $t > t^*$.*

Proof. Assume that $d_t \geq 0$ for all t, then

$$\frac{\alpha_{t+1}}{\beta_{t+1}} = \frac{A+d_{t+1}}{(A+d_{t+1})^2 + \phi\, d_{t+1}^2} \cdot \frac{\alpha_t}{\beta_t} = \frac{1}{(A+d_{t+1}) + \frac{\phi\, d_{t+1}^2}{A+d_{t+1}}} \cdot \frac{\alpha_t}{\beta_t} \leq \frac{1}{A} \frac{\alpha_t}{\beta_t}$$

Next, we will show that $\frac{\alpha_t}{\beta_t}$ tends to 0 as t increases. Since all $d_t \geq 0$ for all t, the above denominator

$$(A+d_t) + \frac{\phi d_t^2}{A+d_t} > 1.$$

Therefore, there exists a $\delta \neq 0$ such that

$$1 \geq \frac{1}{(A+d_t) + \frac{\phi\, d_t^2}{A+d_t}} + \frac{\delta}{(A+d_t) + \frac{\phi\, d_t^2}{A+d_t}}.$$

Since all $d_t \geq 0$ for all t, then the above denominator

$$(A + d_t) + \frac{\phi d_t^2}{A + d_t} > 1.$$

Since $\frac{\alpha_t}{\beta_t} \leq 1$ and $0 < \frac{\alpha_t^2}{\beta_t^2} < 1$, then

$$\begin{aligned} d_t &= \frac{\frac{1}{\gamma} - A + A\frac{\alpha_t^2}{\beta_t}}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} \\ &\leq \frac{\frac{1}{\gamma}}{\phi} = d_1. \end{aligned}$$

Therefore, d_t is bounded above by $d_1 = \frac{1}{\gamma\phi}$, and $\frac{\delta}{(A+d_t)+\frac{\phi\, d_t^2}{A+d_t}}$ is strictly bounded away from 0. Hence, $\lim_{t\to\infty} \frac{\alpha_t}{\beta_t} = 0$. Now,

$$\begin{aligned} \lim_{t\to\infty} d_t = \lim_{t\to\infty} \frac{\frac{1}{\gamma}\cdot\frac{\alpha_t}{\beta_t} - A(1 - \frac{\alpha_t^2}{\beta_t})}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} &\leq \lim_{t\to\infty} \frac{\frac{1}{\gamma}\cdot\frac{\alpha_t}{\beta_t} - A(1 - c_1)}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} \\ &= \lim_{t\to\infty} \frac{\frac{1}{\gamma}\cdot\frac{\alpha_t}{\beta_t}}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} - \lim_{t\to\infty} \frac{A(1 - c_1)}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} \\ &\leq -\frac{A\,(1 - c_1)}{\phi + 1} < 0. \end{aligned}$$

This contradicts the positivity of d_t for all t. □

Remark 4.4.2. Lemma 4.4.2 shows the closed-loop equilibrium control can attain a negative values when t is sufficient large. The following lemma will state a lower bound for the equilibrium control.

Lemma 4.4.3. *There exists a lower bound, $-A$, such that $d_t > -A$ and $\alpha_t, \beta_t > 0$ for all t.*

Proof. By Lemma 4.4.1, in the formula (4.10), the denominator is positive. Hence $d_{t+1} > -A$ is equivalent to

$$\frac{1}{\gamma}\frac{\alpha_t}{\beta_t} - A\,(1 - \frac{\alpha_t^2}{\beta_t}) > -A\,(\phi + 1 - \frac{\alpha_t^2}{\beta_t}),$$

which is

$$\frac{\alpha_t}{\gamma\beta_t} > -A\phi.$$

As β_t is strictly positive for all t, then it is sufficient to show $\alpha_t > 0$. This is true because $\alpha_0 = 1$ implies $d_1 > -A$ and the relation (4.11) states for all t

$$\alpha_{t+1} = (A + d_t)\alpha_t.$$

This ensures $\alpha_1 > 0$ and by using induction, α_t is positive for all t, which completes the proof. □

As we mentioned, the limit of closed-loop equilibrium control depends on the Sharpe ratio. For a specific range of Sharpe ratio, the limit of equilibrium control can be provided.

Lemma 4.4.4. *If $\phi > \frac{1}{4A(A-1)}$, then $\Theta_t := \frac{1}{\gamma}\frac{\alpha_t}{\beta_t} + A\frac{\alpha_t^2}{\beta_t} \to 0$ and $d_t \to \frac{-A}{\phi+1}$ as $t \to \infty$.*

Proof. Firstly, recall the expression (4.13) of d_t

$$d_{t+1} = \frac{\frac{1}{\gamma}\frac{\alpha_t}{\beta_t} - A\left\{1 - \frac{\alpha_t^2}{\beta_t}\right\}}{(\phi+1) - \frac{\alpha_t^2}{\beta_t}}, \tag{4.14}$$

showing the limit of d_t is equivalent to study the limit of $\frac{\alpha_t}{\beta_t}$ and $\frac{\alpha_t^2}{\beta_t}$ We will show that,

$$\frac{\alpha_t}{\beta_t} \to 0, \quad \text{as } t \to \infty.$$

Then, for $\phi > \frac{1}{4A(A-1)}$,

$$(A + d_t)^2 + \phi d_t^2 - (A + d_t) > 0 \quad \text{for all real value } d_t.$$

Providing the value of ϕ, there exists a $\delta > 0$ such that

$$(A + d_t)^2 + \phi d_t^2 - (A + d_t) \geq \delta.$$

Recall the equation (4.11) and (4.12) we have

$$\frac{\alpha_{t+1}}{\beta_{t+1}} = \frac{(A+d_t)}{(A+d_t)^2 + \phi d_t^2} \frac{\alpha_t}{\beta_t}.$$

Then, we have

$$\frac{(A+d_t)}{(A+d_t)^2 + \phi d_t^2} \leq 1 - \frac{\delta}{(A+d_t)^2 + \phi d_t^2}.$$

To show the limit of $\frac{\alpha_t}{\beta_t}$ is equivalent to have $\frac{(A+d_t)}{(A+d_t)^2+\phi d_t^2}$ being strictly bounded away from 1 by a constant. However, it is not clear now as d_t can be arbitrarily large. Hence, we need to prove that the limit of supremum of d_t is bounded. Since Lemma 4.4.3 shows $\frac{\alpha_t}{\beta_t} \leq \frac{1}{A}\frac{\alpha_0}{\beta_0} \leq 1$ and Lemma 4.4.1 shows $0 < \frac{\alpha_t^2}{\beta_t} \leq 1$, then

$$d_t = \frac{\frac{1}{\gamma}\frac{\alpha_t}{\beta_t} - A + A\frac{\alpha_t^2}{\beta_t}}{\phi + 1 - \frac{\alpha_t^2}{\beta_t}} \leq \frac{\frac{1}{\gamma}}{\phi} = d_0 < \infty \quad \text{for all } t.$$

Therefore, there exists a $\tilde{\delta} > 0$ such that

$$1 - \tilde{\delta} \geq \frac{A + d_t}{(A+d_t)^2 + \phi d_t^2},$$

and

$$\lim_{t\to\infty} \frac{\alpha_t}{\beta_t} = 0.$$

Next, we will find the limit of $\frac{\alpha_t^2}{\beta_t}$. Lemma 4.4.1 and Lemma 4.4.3 show that $\frac{\alpha_t^2}{\beta_t}$ is a monotonic decreasing sequence which is bounded below by 0. Then there is $\kappa \in (0,1)$

$$\lim_{t\to\infty} \frac{\alpha_t^2}{\beta_t} = \kappa.$$

Then,

$$\limsup_{t\to\infty} \Theta_t = 0 + A\kappa < A.$$

This implies d_t is bounded away from 0 for sufficiently large t, since

$$\begin{aligned}\limsup_{t\to\infty} d_t &= \limsup_{t\to\infty} \frac{\frac{1}{\gamma}\frac{\alpha_t}{\beta_t} - A\big\{1 - \frac{\alpha_t^2}{\beta_t}\big\}}{(\phi+1) - \frac{\alpha_t^2}{\beta_t}} \\ &= \frac{1}{\phi + 1 - \kappa}\big(\limsup_{t\to\infty} \Theta_t - A\big) \\ &< 0.\end{aligned}$$

Then, from Lemma 4.4.1, we have $\kappa = 0$. For sufficient large t, there is an $\eta < 1$ such that

$$\frac{\alpha_{t+1}^2}{\beta_{t+1}} = \frac{1}{1 + \frac{\phi d_t^2}{(A+d_t)^2}} \frac{\alpha_t^2}{\beta_t} \le \eta \frac{\alpha_t^2}{\beta_t}.$$

Therefore,

$$\lim_{t\to\infty} \Theta_t \to 0$$

and

$$\lim_{t\to\infty} d_t = \frac{-A}{1+\phi}.$$

□

Note that it is also important to explore the case when Sharpe ratio is outside of above range. By analysing the different range of Sharpe ratio, the different investment behaviour can be studied. Recall that α_t represents the dynamics of the first moment of wealth process, then $d_t + A$ shows the expected investment return or loss over a single period. Therefore, we will focus on the cases when $d_t + A > 1$ and $d_t + A < 1$.

Lemma 4.4.5. *Assume* $A + d_t \geq 1$, *then* $d_{t+1} < d_t$.

Proof. Equation (4.10) is equivalent to

$$\frac{1}{\gamma}\alpha_t + (A + d_t)\alpha_t^2 = (d_t\phi + d_t + A)\beta_t. \tag{4.15}$$

To show $d_{t+1} < d_t$, it is equivalent to show

$$\frac{\frac{1}{\gamma}\alpha_{t+1} - A(\beta_{t+1} - \alpha_{t+1}^2)}{\beta_{t+1}(\phi + 1) - \alpha_{t+1}^2} \leq d_t. \tag{4.16}$$

Lemma 4.4.1 provides the denominator is positive, then the above inequality (4.15) can be rearranged as:

$$\begin{aligned}
&\text{LHS of (4.15)} = \frac{1}{\gamma}\alpha_{t+1} + A\alpha_{t+1}^2 + d_t\alpha_{t+1}^2 \\
&= \frac{1}{\gamma}(A + d_t)\alpha_t + A(A + d_t)^2\alpha_t^2 + d_t(A + d_t)^2\alpha_t^2 \quad \text{(by the relation (4.11))} \\
&= \frac{1}{\gamma}(A + d_t)\alpha_t + (A + d_t)^3\alpha_t^2 \\
&< \text{RHS of (4.15)} \\
&= (d_t\phi + d_t + A)\ \beta_{t+1} \\
&= (d_t\phi + d_t + A)\left((A + d_t)^2 + \phi\ d_t^2\right)\beta_t \quad \text{(by the relation (4.12))} \\
&= \left((A + d_t)^2 + \phi\ d_t^2\right)\left(\frac{1}{\gamma}\alpha_t + (A + d_t)\alpha_t^2\right) \quad \text{(by equation (4.15))} \\
&= \frac{1}{\gamma}(A + d_t)^2\alpha_t + (A + d_t)^3\alpha_t^2 + \frac{1}{\gamma}\phi d_t^2\alpha_t + \phi d_t^2(A + d_t)\alpha_t^2
\end{aligned}$$

Therefore, cancelling the above term $(A + d_t)^3\alpha_t^2$, it suffices to show that

$$\frac{1}{\gamma}(A + d_t) < \frac{1}{\gamma}(A + d_t)^2 + \frac{1}{\gamma}\phi d_t^2\alpha_t + \phi d_t^2(A + d_t)\alpha_t^2$$

When $(A + d_t) \geq 1$, then the above inequality is trivial. $\square$

Remark 4.4.3. Under the assumption of Lemma 4.4.4 $\phi \geq \frac{1}{A-1}$, we have the limit of d_t is equal to $\frac{-A}{1+\phi}$. As a result,

$$\frac{-A}{1+\phi} + A = \frac{A}{1+\frac{1}{\phi}} \geq 1,$$

which is sufficient to show

$$A > 1 + \frac{1}{\phi} \quad \Leftrightarrow \phi > \frac{1}{A-1}.$$

This is fulfilled by the assumption of Lemma 4.4.4. Providing Lemma 4.4.5, we have a monotonic decreasing sequence d_t converges to the limit $\frac{-A}{1+\phi}$. We decide to prove the behaviour of the closed-loop control for full generality ($\phi < \frac{1}{A-1}$) in the next lemma. The proof of Lemma 4.4.6 will be needed in the remaining results.

Lemma 4.4.6. *Suppose* $\phi \geq \frac{1}{(A-1)}$. *If* $d_t + A \geq 1$, *then*

$$d_{t+1} + A \geq 1.$$

Proof. We need to show

$$d_{t+1} = \frac{\frac{1}{\gamma}\alpha_{t+1} - A(\beta_{t+1} - \alpha_{t+1}^2)}{\beta_{t+1}(\phi+1) - \alpha_{t+1}^2} \geq 1 - A.$$

Define

$$\Lambda(\alpha_t, \beta_t) := -\frac{1}{\gamma}\alpha_{t+1} + A(\beta_{t+1} - \alpha_{t+1}^2) + (1-A)\Big(\beta_{t+1}(\phi+1) - \alpha_{t+1}^2\Big).$$

Then, we have

$$\Lambda(\alpha_t, \beta_t) = -\frac{1}{\gamma}\alpha_{t+1} - \alpha_{t+1}^2 + \big(\phi + 1 - A\phi\big)\beta_{t+1}.$$

From (4.11),

$$\begin{aligned}\frac{1}{\gamma}\alpha_{t+1}+\alpha_{t+1}^2 &= \frac{1}{\gamma}(A+d_t)\alpha_t+(A+d_t)^2\alpha_t^2\\ &=(A+d_t)\Big(\frac{1}{\gamma}\alpha_t+(A+d_t)\alpha_t^2\Big)\\ &=(A+d_t)(d_t\phi+d_t+A)\beta_t \qquad \text{by equation (4.15).}\end{aligned}$$

Then,

$$\Lambda(\alpha_t,\beta_t)=-(A+d_t)(d_t\phi+d_t+A)\beta_t+\big(\phi+1-A\phi\big)\Big((A+d_t)^2+\phi d_t^2\Big)\beta_t.$$

By Lemma 4.4.3, we have $\beta_t>0$, then

$$d_{t+1}+A\geq 1 \qquad \Longleftrightarrow \qquad \frac{\Lambda(d_t)}{\beta_t}\leq 0.$$

We can expand the bracket and simplify the terms to obtain

$$\frac{\Lambda}{\beta_t}(d_t)=d_t^2\phi(\phi+1)(1-A)+d_t\phi A(1-2A)+\phi A^2(1-A)$$

Since $A>1$, then $\frac{\Lambda(d_t)}{\beta_t}$ is a quadratic function in terms of d_t with all the coefficients being negative. It is trivial that $\frac{\Lambda(d_t)}{\beta_t}\leq 0$ for all $d_t\geq 0$. Therefore, it suffices to show $\frac{\Lambda(d_t)}{\beta_t}\leq 0$ for $d_t\in[1-A,0)$. We look at the discriminant of $\frac{\Lambda(d_t)}{\beta_t}$:

$$\Delta := A^2\phi^2(1-2A)^2-4\phi^2A^2(1-A)^2(1+\phi).$$

If $\Delta\leq 0$, then $\frac{\Lambda(d_t)}{\beta_t}\leq 0$ for any d_t, that is

$$\Delta = A^2\phi^2(1-2A)^2-4\phi^2A^2(1-A)^2(1+\phi)\leq 0.$$

The above inequality is true if

$$\begin{aligned}\phi &\geq \frac{(1-2A)^2}{4(1-A)^2} - 1 \\ &= \frac{(1-2A)^2 - 4(1-A)^2}{4(1-A)^2} \\ &= \frac{4A-3}{4(1-A)^2} \\ &= \frac{1}{A-1} + \frac{1}{4(A-1)^2}.\end{aligned}$$

If $\Delta > 0$, in which $\frac{1}{A-1} + \frac{1}{4(A-1)^2} > \phi > \frac{1}{A-1}$, then

$$d_t^+ := \frac{-\phi A(1-2A) - \sqrt{\Delta}}{2\phi(1-A)(1+\phi)} \leq (1-A)$$

This is equivalent to

$$-\sqrt{\Delta} \geq \phi A(1-2A) + 2\phi(1-A)(1-A)(1+\phi).$$

The above inequality requires the right-hand side to be negative,

$$\phi A(1-2A) + 2\phi(1-A)^2(1+\phi) < 0,$$

that is

$$\begin{aligned}\phi &< \frac{A(2A-1)}{2(A-1)^2} - 1 \\ &= \frac{3}{2(A-1)} + \frac{1}{2(A-1)^2}\end{aligned}$$

This is satisfied because $\phi < \frac{1}{A-1} + \frac{1}{4(A-1)^2} < \frac{3}{2(A-1)} + \frac{1}{2(A-1)^2}$. Squaring both sides, we obtain

$$\Delta \leq \phi^2 A^2(1-2A)^2 + 4\phi^2(1-A)^2(1-A)^2(1+\phi)^2 + 4\phi^2 A(1-2A)(1-A)(1-A)(1+\phi).$$

Therefore,

$$
\begin{aligned}
& A^2\phi^2(1-2A)^2 - 4\phi^2A^2(1-A)^2(1+\phi) \\
& \leq \phi^2A^2(1-2A)^2 + 4\phi^2(1-A)^2(1-A)^2(1+\phi)^2 \\
& \quad + 4\phi^2A(1-2A)(1-A)(1-A)(1+\phi),
\end{aligned}
$$

which leads to

$$
\begin{aligned}
& -4\phi^2A^2(1-A)^2(1+\phi) \\
& \leq 4\phi^2(1-A)^2(1-A)^2(1+\phi)^2 + 4\phi^2A(1-2A)(1-A)(1-A)(1+\phi).
\end{aligned}
$$

Dividing both side by $4\phi^2(1+\phi)(1-A)$,

$$
-A^2(1-A) \geq (1-A)^2(1-A)(1+\phi) + A(1-2A)(1-A).
$$

Therefore,

$$
\begin{aligned}
\phi &\geq \frac{A^2(1-A) + A(1-2A)(1-A)}{(A-1)(1-A)^2} - 1 \\
&= \frac{-A^2 - A(1-2A)}{(A-1)^2} - 1 \\
&= \frac{A(A-1)}{(A-1)^2} - 1 \\
&= \frac{A}{A-1} - 1 \\
&= \frac{1}{A-1}.
\end{aligned}
$$

Recall that $\frac{\Lambda(d_t)}{\beta_t} \leq 0$ for all $d_t \geq 0$. Therefore, it suffices to show $\frac{\Lambda(d_t)}{\beta_t} \leq 0$ for $d_t \in [1-A, 0)$.In the case of $d_t^+ < 1-A$, we have $\frac{\Lambda(d_t)}{\beta_t} \leq 0$ for any $d_t \in [1-A, 0)$, which implies $d_{t+1} + A \geq 1$. □

Lemma 4.4.7. *Suppose* $\phi < (A-1)^{-1}$*, if* $d_t + A < 1$*, then*

$$d_{t+1} + A < 1.$$

Proof. By Lemma 4.4.3 and Lemma 4.4.6 , it is equivalent to show that

$$\frac{\Lambda(d_t)}{\beta_t} > 0, \quad \text{for} \quad d_t \in (-A, 1-A),$$

where

$$\frac{\Lambda(d_t)}{\beta_t} = d_t^2\phi(\phi+1)(1-A) + d_t\phi A(1-2A) + \phi A^2(1-A).$$

Since $\phi < (A-1)^{-1}$, the discriminant Δ is positive. We will show that

$$\frac{\Lambda(-A)}{\beta_t} \geq 0 \quad \text{and} \quad \frac{\Lambda(1-A)}{\beta_t} \geq 0.$$

Since

$$\frac{\Lambda(-A)}{\beta_t} = A^2\phi(\phi+1)(1-A) - A\phi A(1-2A) + \phi A^2(1-A) \geq 0$$

if

$$(\phi+1)(1-A) - (1-2A) + (1-A) \geq 0.$$

Meanwhile,

$$\frac{\Lambda(1-A)}{\beta_t} = (1-A)^2\phi(\phi+1)(1-A) + (1-A)\phi A(1-2A) + \phi A^2(1-A) \geq 0$$

if

$$(1-A)^2(\phi+1) + A(1-2A) + A^2 \leq 0.$$

Both inequalities are true since $\phi < \frac{1}{A-1}$, which completes the proof. □

Note that the lemma 4.4.5 and lemma 4.4.6 shows that, when the Sharp ratio satisfies the condition $\phi > (A-1)^{-1}$, the equilibrium strategy can be negative. The expect return of investment is positive over every single period. Therefore, we have a monotonic decreasing control sequence which is bounded below. However, when the Sharp ratio satisfies the condition $\phi < (A-1)^{-1}$, $d_s + A$ will be less than 1 at some time t and $d_s + A$ is also less than 1 for all $s > t$. The phenomenon shows the investor starts to lose money in the investment intentionally. This is because when the time is far from the terminal and the expect excess return is positive, the wealth is expected to be increased dramatically over a long time period. As a result, the investor is losing money to stabilise the wealth trajectory.

4.5 Conclusion

In this chapter, we define the definition of feedback type strategy by 4.2.3. Comparing to the work in Bjork & Murgoci (2014), the closed-loop equilibrium strategy has been derived from a different perspective. The difference between open-loop and closed-loop equilibrium investor lies in the short-term investment plan. Comparing to the open-loop short-term plan, which is $D_{t+1}\mathbf{B}_{t+1}^T\varepsilon$, the short-term plan of the closed-loop equilibrium $\mathbf{B}_{t+1}^T\prod_{i=t+1}^{T-1}(A_i + \mathbf{B}_{i+1}^T\mathbf{K}_i)\varepsilon$ is to invest a small amount ε into stock market and leave this investment until the maturity.

Taking the advantage of separable form (4.4) of terminal wealth, under the equilibrium definition 4.2.5 and the assumption of invertibility in Lemma 4.3.1, we obtain a necessary and sufficient condition for closed-loop equilibrium strategy and the uniqueness of the solution in Lemma **??** and Lemma 4.3.1. We then show such condition yields a linear feedback type strategy which coincides with the equilibrium strategy in Bjork & Murgoci (2014) in one risky asset case.

We discuss the asymptotic behaviour of using the closed-loop equilibrium strategy for mean-variance optimisation problem. The relation between the sharp ratio and the asymptotic limit of closed-loop equilibrium strategy has been studied. Firstly, Lemma 4.4.5 addresses the fact that the closed-loop equilibrium strategy reduces the investment amount in risky asset as time goes backwards, which is similar to the open-loop equilibrium strategy. Next, it has been shown in

Lemma 4.4.2 that, for any mean and variance of excess return of risky asset, the closed-loop will always short the stocks at some point of investment time period. In particular, when the parameter $\phi = \frac{\sigma^2}{\mu^2} > \frac{1}{4A(A-1)}$, Lemma 4.4.4 proves that the asymptotic limit of closed-loop equilibrium strategy tends to $-\frac{-A}{\phi+1}$. It can be seen that the above asymptotic limit is negative. Although the closed-loop tends to such limit for all $\phi > \frac{1}{4A(A-1)}$, there exists a difference in the investment behaviour with respect to the different values of ϕ. For $\phi \geq \frac{1}{A-1}$, Lemma 4.4.6 shows the expected excess return factor of the investment portfolio $(A + d_t)$ is greater or equal to 1 in a single-period. This implies a positive excess return rate of the investment portfolio. However, when $\frac{1}{4A(A-1)} < \phi < \frac{1}{A-1}$, Lemma 4.4.7 describes a situation where the investor has negative portfolio return.

Chapter 5

Open-loop strategy in continuous time setting

5.1 Introduction

In this chapter, we study the open-loop equilibrium for the linear-quadratic problem in continuous time, where the objective functional includes both a quadratic term of the expected state and a state-dependent term. In previous chapters, it has been shown that the perturbation plays an important role in determining the type of equilibrium strategy. Therefore, the aim of this chapter is to explore the equilibrium strategy in continuous time by choosing a different perturbation.

In Bjork & Murgoci (2010), the closed-loop equilibrium is obtained by setting the perturbation as a deterministic function. In contrast, the open-loop equilibrium is obtained by setting the perturbation as a square-integrable random variable. In this chapter, we define an adapted process as the perturbation. The goal is to study the impact of perturbation in continuous time setting.

5.2 Theoretical framework

Let $T > 0$ be a finite time horizon and $(W_t)_{0\leq t\leq T}$ be a d-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The natural filtration of (W_t) is denoted by $\mathcal{F}_t = \{\sigma(W_s), 0 \leq s \leq t\}$.

For any time $t \in [0, T]$ and $X_t = x_t$, we consider a stochastic linearly controlled system:

$$dX_s = \left[A_s X_s + B_s^T u_s + b_s\right] ds + \sum_{i=1}^{d} \left[C_s^i X_s + D_s^i u_s + \sigma_s^i\right] dW_s^i; \quad \forall s \geq t, \tag{5.1}$$

where $A : [0, T] \to \mathbb{R}^{n\times n}$ is a deterministic function. The coefficients B, C^i and D^i are essentially bounded adapted process on $\Omega\times[0, T]$ with values in $\mathbb{R}^{l\times n}$, $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{n\times l}$, respectively. The other processes b and σ^i are in $\mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^n)$, where $\mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^n)$ is the space of $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes $\xi = \{\xi_s; t \leq s \leq T\}$ such that $\mathbb{E}\left[\int_t^T |\xi_s|^2 ds\right] < \infty$. The process $u \in \mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^l)$ is called an admissible control and its role is to control the dynamics of the diffusion $(X_s)_{s\in[t,T]}$. For any x, y, we assume

$$\left|A_s x + B_s^T u_s + b_s - A_s y - B_s^T u_s - b_s\right| \leq \operatorname*{ess\,sup}_{\omega\in\Omega, s\in[t,T]} |A_s|\, |x - y|$$

and

$$|C_s x + D_s u_s + \sigma_s - C_s y - D_s u_s - \sigma_s| \leq \operatorname*{ess\,sup}_{\omega\in\Omega, s\in[t,T]} |C_s|\, |x - y|$$

By Theorem 7, Section 2.5 in Krylov (2008), there exists a unique strong solution $X_s \in \mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^n)$ of (5.1).

For any time $t \in [0, T]$ with the state $X_t = x_t$ with $x_t \in \mathcal{L}^2_{\mathcal{F}}(t, T; \mathbb{R}^n)$, we aim to minimise the functional

$$J(t, x_t; u) = \frac{1}{2}\mathbb{E}_t \int_t^T [Q_s X_s^2 + R_s u_s^2] ds + \frac{G}{2}\mathbb{E}_t[X_T^2] - \frac{h}{2}(\mathbb{E}_t[X_T])^2 - (\mu_1 x_t + \mu_2)\mathbb{E}_t[X_T] \tag{5.2}$$

over the set of admissible controls, where X is the state process corresponding to the control u and $\mathbb{E}_t[\cdot]$ denotes $\mathbb{E}[\cdot|\mathcal{F}_t]$. Here Q and R are both non-negative essentially bounded adapted processes and G, h, μ_1, μ_2 are constants with G being non-negative.

We notice that, this linear-quadratic functional can represent many different types of problems. When Q and R are 0, the problem becomes a mean- variance optimisation with state-dependent risk aversion. When h, μ_1 and μ_2 are 0, the problem becomes an inventory problem with the stage cost and final stage cost.

Definition 5.2.1. Let u^* be an admissible control and $v \in \mathcal{L}^\infty_\mathcal{F}(0,T;\mathbb{R}^l)$, where $\mathcal{L}^\infty_\mathcal{F}(t,T;\mathbb{R}^l)$ is the space of $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes $\xi = \{\xi_s; t \leq s \leq T\}$ such that $\operatorname{ess\,sup}_{\omega\in\Omega,s\in[t,T]} |\xi_s(\omega)| < \infty$. For any $t \in [0,T]$ and $\varepsilon > 0$, define the perturbed control

$$u^{t,\varepsilon,v}_s = u^*_s + v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}. \tag{5.3}$$

Let $X^{t,\varepsilon,v}$ and X^* be the state processes corresponding to $u^{t,\varepsilon,v}$ and u^*, respectively.

Definition 5.2.2. An admissible u^* is called an open-loop equilibrium control if

$$\liminf_{\varepsilon\downarrow 0} \frac{J(t,X^*_t;u^{t,\varepsilon,v}) - J(t,X^*_t;u^*)}{\varepsilon} \geq 0, \quad \text{a.s.} \tag{5.4}$$

for any $t \in [0,T)$ and $v_s \in \mathcal{L}^\infty_\mathcal{F}(0,T;\mathbb{R}^l)$.

5.3 Necessary and Sufficient Condition for Equilibrium Controls

We will prove the above necessary and sufficient condition for $u^* \in \mathcal{L}^2_\mathcal{F}$ to be an equilibrium control. Throughout this section, we will denote by X^* the state process corresponding to the control u^*.

Let us introduce notation used in this section. We will denote by $\mathcal{L}^2_\mathcal{F}(\Omega;C(t,T;\mathbb{R}))$ the space of continuous $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes $(\xi_s)_{t\leq s\leq T}$ such that $\mathbb{E}\left\{\sup_{s\in[t,T]}|\xi_s|^2\right\} < \infty$. We also extend standard order notation to random variables. For a sequence of random variables ξ^ε parametrised by $\varepsilon > 0$, we write $\xi^\varepsilon = o(\varepsilon^k)$ if $\lim_{\varepsilon\to 0}\operatorname{ess\,sup}\left|\frac{\xi^\varepsilon}{\varepsilon^k}\right| = 0$; and $\xi^\varepsilon = O(\varepsilon^k)$ if $\lim_{\varepsilon\to 0}\operatorname{ess\,sup}\left|\frac{\xi^\varepsilon}{\varepsilon^k}\right| < c$ for some finite constant c.

For any $t \in [0,T]$, consider the following backward stochastic differential equations:

$$\begin{cases} dp(s;t) = -[A_s p(s;t) + C_s k(s;t) + Q_s X_s^*]ds + k(s;t)^T dW_s, \quad s \in [t,T], \\ p(T;t) = GX_T^* - h\mathbb{E}_t[X_T^*] - \mu_1 X_t^* - \mu_2; \end{cases} \tag{5.5}$$

$$\begin{cases} dP_s = -[2A_s P_s + C_s^2 P_s + 2C_s K_s + Q_s]ds + K_s^T dW_s, \quad s \in [t,T], \\ P_T = G. \end{cases} \tag{5.6}$$

Lemma 5.3.1. *For any admissible control u^*, equations (5.5) and (5.6) admit unique solutions $(p(\cdot;t), k(\cdot;t)) \in \mathcal{L}^2_{\mathcal{F}}(\Omega; C(t,T;\mathbb{R})) \times \mathcal{L}^2_{\mathcal{F}}(t,T;\mathbb{R})$ and $(P(\cdot), K(\cdot)) \in \mathcal{L}^2_{\mathcal{F}}(\Omega; C(t,T;\mathbb{R})) \times \mathcal{L}^2_{\mathcal{F}}(t,T;\mathbb{R})$, respectively, where $\mathcal{L}^2_{\mathcal{F}}(\Omega; C(t,T;\mathbb{R}))$ is the space of continuous $\{\mathcal{F}_s\}_{s\in[t,T]}$-adapted processes $\xi = \{\xi_s; t \le s \le T\}$ such that $\mathbb{E}\left[\sup_{s\in[t,T]} |\xi_s|^2\right] < \infty$. Moreover, $P_s \ge 0$ a.s. for all $s \in [t,T]$.*

Proof. Denote $g^1(s,u,v) = A_s u + C_s v + Q_s X_s^*$ and $g^2(s,u,v) = 2A_s u + C_s^2 v + 2C_s y + Q_s$. For any $s \in [t,T]$, $g^1(s,0,0) \in L^2_{\mathcal{F}}(s,T;\mathbb{R})$ and $g^2(s,0,0) \in L^2_{\mathcal{F}}(s,T;\mathbb{R})$. Moreover, it can be shown that $g^1(s,u,v)$ and $g^2(s,u,v)$ are almost surely uniformly Lipschitz with respect to (u,v). Therefore, according to Theorem 1.1 El Karoui *et al.* (2008), there exist unique pairs of solutions $(p(\cdot;t), k(\cdot;t))$ and $(P(\cdot), K(\cdot))$ satisfying (5.5) and (5.6), respectively. Moreover, by Proposition 1.3 in El Karoui *et al.* (2008), we have $P_s \ge 0$ a.s. for all $s \in [t,T]$. □

Next, we will prove an important proposition which is very similar to the proposition in Hu *et al.* (2012)[Proposition 3.1]. The reason for us to provide all the details of the proof is that, firstly, in Hu *et al.* (2012), they refer to a theorem in Yong & Zhou (1999) to prove the conditional expectation of supremum of the processes Y and Z are in the $O(\varepsilon)$ and $O(\varepsilon^2)$, respectively. However, the theorem in Yong & Zhou (1999) does not have a conditional version of the proof, and it only specifies the order for the supremum of the expectation. Secondly, our definition of equilibrium control differs from the definition in Hu *et al.* (2012). Therefore, we need to ensure that the proposition in Hu *et al.* (2012) still holds in our case.

Proposition 5.3.1. *For any admissible control u^*, $t \in [0,T), \varepsilon > 0$ and $v_s \in \mathcal{L}^\infty_{\mathcal{F}}(t,T;\mathbb{R})$,*

$$J(t, X_t^*; u^{t,\varepsilon,v}) - J(t, X_t^*; u^*) = \mathbb{E}_t \int_t^{t+\varepsilon} \left\{ \Lambda(s;t) v_s + \frac{1}{2} H_s v_s^2 \right\} ds + o(\varepsilon), \quad (5.7)$$

where $u^{t,\varepsilon,v}$ is defined by (2.3); $\Lambda(s;t) := B_s p(s;t) + D_s k(s;t) + R_s u_s^*$ and $H_s := R_s + D_s^2 P_s$.

The proof of this proposition is comprised of a number of lemmas. The constant C in the proofs may differ from line to line.

Define the processes $Y_s \equiv Y_s^{t,\varepsilon,v}$ and $Z_s \equiv Z_s^{t,\varepsilon,v}$ that satisfy

$$\begin{cases} dY_s = A_s Y_s ds + [C_s Y_s + D_s v_s \mathbb{1}_{\{s \in [t,t+\varepsilon)\}}] dW_s, \quad s \in [t,T], \\ Y_t = 0; \end{cases}$$

$$\begin{cases} dZ_s = [A_s Z_s + B_s v_s \mathbb{1}_{\{s \in [t,t+\varepsilon)\}}] ds + C_s Z_s dW_s, \quad s \in [t,T], \\ Z_t = 0; \end{cases}$$

Lemma 5.3.2. *Let $X^{t,\varepsilon,v}$ be the state process corresponding to $u^{t,\varepsilon,v}$, then*

$$X_s^{t,\varepsilon,v} = X_s^* + Y_s + Z_s, \quad s \in [t,T].$$

Proof. Since $u_s^{t,\varepsilon,v} = u_s^* + v_s \mathbb{1}_{\{s \in [t,t+\varepsilon)\}}$, for $X^{t,\varepsilon,v}$,

$$dX^{t,\varepsilon,v} = \left[A_s X_s + B_s^T u_s^{t,\varepsilon,v} + b_s\right] ds + \left[C_s X_s + D_s u_s^{t,\varepsilon,v} + \sigma_s^i\right] dW_s,$$

Set $\eta_s = X_s^{t,\varepsilon,v} - X_s^* - Y_s - Z_s$, we have

$$\begin{aligned}
d\eta_s &= \left(dX_s^{t,\varepsilon,v} - dX_s^*\right) - dY_s - dZ_s \\
&= \left(A_s(X^{s,\varepsilon,v} - X_s) + B_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right)ds + \left(C_s(X^{s,\varepsilon,v} - X_s) D_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right)dW_s \\
&\quad - A_s Y_s ds - [C_s Y_s + D_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]dW_s - [A_s Z_s + B_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]ds - C_s Z_s dW_s \\
&= A_s(X^{s,\varepsilon,v} - X_s - Y_s - Z_s)ds + C_s(X^{s,\varepsilon,v} - X_s - Y_s - Z_s)dW_s \\
&= A_s\eta_s ds + C_s\eta_s dW_s.
\end{aligned}$$

Therefore,

$$\begin{cases}
d\eta_s = A_s\eta_s ds + C_s\eta_s dW_s, \quad s\in[t,T], \\
\eta_t = 0.
\end{cases}$$

This yields, $\eta_s = 0$ for all $s\in[t,T]$ □

Lemma 5.3.3. *Let $L = max\left(\operatorname{ess\,sup}|A_s|, \operatorname{ess\,sup}|C_s|, \operatorname{ess\,sup}|D_s|, \operatorname{ess\,sup}|v_s|\right) < \infty$ be the maximum of essential suprema of the processes A_s, C_s, D_s and v_s over all $\omega\in\Omega$ and $s\in[t,T]$. The processes $\mathbb{E}_t[Y_s^{2k}]$ and $\mathbb{E}_t[Z_s^{2k}]$ have continuous trajectories with respect to s for any $k\geq 1$. There exists a constant M depending on r and L such that, for all $t\in[0,T]$*

$$\sup_{s\in[t,T]} \mathbb{E}_t\left[Y_s^{2r}\right] \leq M\cdot\varepsilon^r, \quad \sup_{s\in[t,T]} \mathbb{E}_t\left[Z_s^{2r}\right] \leq M\cdot\varepsilon^{2r} \quad a.s.$$

Proof. First we prove that $\mathbb{E}_t[Y_s^{2r}]$ is finite and $s\mapsto\mathbb{E}_t[Y_s^{2r}]$ is continuous (the proof for $\mathbb{E}_t[Z_s^{2r}]$ is analogous). Since $|A_sY_s|\leq L|Y_s|$ and

$$\begin{aligned}
\left|C_sY_s + D_s v_s \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}\right|^2 &\leq |C_s|^2|Y_s|^2 + \left|D_s v_s \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}\right|^2 + 2\left|C_s D_s v_s \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}\right||Y_s| \\
&\leq L^2|Y_s|^2 + L^4 + \left|C_s D_s v_s \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}\right|^2 + |Y_s|^2 \\
&\leq (L^4 + L^6) + (1 + L^2)|Y_s|^2,
\end{aligned}$$

by Krylov (2008)[Corollary 10, Section 2.5], we have

$$\mathbb{E}\big[\sup_{u\in[t,T]}|Y_u|^{2r}\big] \le N(r,L)T^{r-1}e^{N(r,L)T}\mathbb{E}\Big[\int_t^T (L^4+L^6)^{2r}ds\Big], \tag{5.8}$$

where $N(r,L)$ is a constant depending only on r and L. This, in particular, implies that $\mathbb{E}_t[Y_s^{2r}] < \infty$. To prove continuity of $s \mapsto \mathbb{E}_t[Y_s^{2r}]$, it is sufficient to show that for any $t_1 \in [t,T]$

$$\lim_{t_2\to t_1}\big(\mathbb{E}_t[Y_{t_2}^{2r}] - \mathbb{E}_t[Y_{t_1}^{2r}]\big) = 0$$

over $t_2 \in [t,T]$. This follows by conditional dominated convergence theorem (Williams, 1991, Lemma 9.7)

$$\lim_{t_2\to t_1}\big(\mathbb{E}_t[Y_{t_2}^{2r}] - \mathbb{E}_t[Y_{t_1}^{2r}]\big) = \lim_{t_2\to t_1}\mathbb{E}_t\big[Y_{t_2}^{2r} - Y_{t_1}^{2r}\big] = \mathbb{E}_t\big[\lim_{t_2\to t_1}(Y_{t_2}^{2r} - Y_{t_1}^{2r})\big] = 0 \quad a.s.,$$

where the dominant

$$\big(Y_{t_2}^{2r} - Y_{t_1}^{2r}\big) \le 2\sup_{u\in[t,T]}|Y_u|^{2r} \quad a.s.$$

is integrable by (5.8).

The rest of the proof is concerned with establishing the bounds for superma of conditional expectations of Y^{2r} and Z^{2r}. Applying Itô's formula to the process Y_s^{2r}

$$\begin{aligned} dY_s^{2r} = {} & \big[2rA_s + r(2r-1)C_s^2\big]Y_s^{2r}ds + 2r(2r-1)C_sD_sv_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}Y_s^{2r-1}ds \\ & + r(2r-1)(D_sv_s)^2\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}Y_s^{2r-2}ds \\ & + 2r[C_sY_s + D_sv_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]Y_s^{2r-1}dW_s. \end{aligned}$$

Take s_1 and δ such that $t \le s_1 < s_1+\delta \le T$. For $s \in [s_1, s_1+\delta)$, integrate both sides of the inequality from s_1 to s and take the conditional expectation with respect to $\mathcal{F}_t$ to get

$$\begin{aligned}\mathbb{E}_t[Y_s^{2r}] \le \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\big\{\mathbb{E}_t[\int_{s_1}^{s} Y_u^{2r} du] \\ + \mathbb{E}_t[\int_{s_1}^{s} |Y_u|^{2r-1} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du] + \mathbb{E}_t[\int_{s_1}^{s} Y_u^{2r-2} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du]\big\} \\ + \mathbb{E}_t[\int_{s_1}^{s} 2r[C_u Y_u + D_u v_u \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}] Y_u^{2r-1} dW_u],\end{aligned}$$

where

$$\begin{aligned}\tilde{M} := \max\big\{2r \cdot \sup_{s\in[t,T]} |A_s| + r(2r-1) \operatorname*{ess\,sup}_{s\in[t,T]} |C_s|^2, \\ 2r(2r-1) \operatorname*{ess\,sup}_{s\in[t,T]} |C_s D_s v_s|, r(2r-1) \operatorname*{ess\,sup}_{s\in[t,T]} |D_s v_s|^2\big\}\end{aligned}$$

is a constant depending only on r and L. Since (Y_u^l) is square-integrable for all $l \ge 1$ (see the beginning of the proof) and (v_u) is bounded, then the process $2r[C_u Y_u + D_u v_u \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}] Y_u^{2r-1}$ is square-integrable. Therefore, the stochastic integral is a martingale and it vanishes under the conditional expectation. By Fubini's theorem for conditional expectation we have

$$\begin{aligned}\mathbb{E}_t[Y_s^{2r}] \le \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\Big\{\int_{s_1}^{s} \mathbb{E}_t[Y_u^{2r}] du + \int_{s_1}^{s} \mathbb{E}_t[|Y_u|^{2r-1}] \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du \\ + \int_{s_1}^{s} \mathbb{E}_t[Y_u^{2r-2}] \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du\Big\}.\end{aligned}$$

Using Hölder inequality for conditional expectations (Chen, 2006, p. 332), we express lower conditional moments of Y_u by the $2r$-th moment:

$$\mathbb{E}_t[|Y_u|^{2r-1}] \le \left(\mathbb{E}_t[Y_u^{2r}]\right)^{\frac{2r-1}{2r}},$$

$$\mathbb{E}_t[|Y_u|^{2r-2}] \le \left(\mathbb{E}_t[Y_u^{2r}]\right)^{\frac{2r-2}{2r}}.$$

Hence,

$$\mathbb{E}_t[Y_s^{2r}] \leq \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\{\int_{s_1}^{s} \mathbb{E}_t[Y_u^{2r}]du + \int_{s_1}^{s} (\mathbb{E}_t[Y_u^{2r}])^{\frac{2r-1}{2r}} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du + \int_{s_1}^{s} (\mathbb{E}_t[Y_u^{2r}])^{\frac{2r-2}{2r}} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du\}.$$

Since the process $\mathbb{E}_t[Y_s^{2r}]$ has continuous trajectories, $\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}]$ is well-defined and

$$\begin{aligned}\mathbb{E}_t[Y_s^{2r}] \leq \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\{&\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] \cdot (s - s_1)\\ &+ (\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}])^{\frac{2r-1}{2r}} \int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du\\ &+ (\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}])^{\frac{2r-2}{2r}} \int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du\}.\end{aligned}$$

We will now use a trick from the proof in Zhou & Li (2000)[Lemma 4.2, Chapter 3] to ensure that after application of Young's inequality we $(\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}])$ has a coefficient strictly smaller than 1:

$$\begin{aligned}\mathbb{E}_t[Y_s^{2r}] \leq \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\{&\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] \cdot (s - s_1)\\ &+ (\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}])^{\frac{2r-1}{2r}} \cdot (\frac{\delta}{2})^{\frac{2r-1}{2r}} \cdot (\frac{2}{\delta})^{\frac{2r-1}{2r}} \int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du\\ &+ (\sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}])^{\frac{2r-2}{2r}} \cdot (\frac{\delta}{2})^{\frac{2r-2}{2r}} \cdot (\frac{2}{\delta})^{\frac{2r-2}{2r}} \int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}}du\}\end{aligned}$$

Applying Young's inequality to the third term with $p = \frac{2r}{2r-1}$ and $q = 2r$, and to the last term with $p = \frac{2r}{2r-2}$ and $q = r$, then a.s.

$$\begin{aligned}\mathbb{E}_t[Y_s^{2r}] \le \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M}\Big\{ \sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] \cdot (s-s_1) + \frac{2r-1}{2r} \cdot \frac{\delta}{2} \cdot \sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] \\ + \frac{1}{2r} \cdot (\frac{2}{\delta})^{2r-1} \cdot (\int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du)^{2r} + \frac{2r-2}{2r} \cdot \frac{\delta}{2} \cdot \sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] \\ + \frac{1}{r} \cdot (\frac{2}{\delta})^{r-1} \cdot (\int_{s_1}^{s} \mathbb{1}_{\{u\in[t,t+\varepsilon)\}} du)^{r} \Big\}\end{aligned}$$

Since $s - s_1 < \delta$, $\frac{2r-1}{2r} < 1$ and $\frac{2r-2}{2r} < 1$, then

$$\mathbb{E}_t[Y_s^{2r}] \le \mathbb{E}_t[Y_{s_1}^{2r}] + \tilde{M} \cdot 2\delta \sup_{u\in[s_1,s)} \mathbb{E}_t[Y_u^{2r}] + \frac{\tilde{M}}{2r} \cdot (\frac{2}{\delta})^{2r-1} \cdot \varepsilon^{2r} + \frac{\tilde{M}}{k} \cdot (\frac{2}{\delta})^{r-1} \cdot \varepsilon^{r}$$

Fixing $\delta = 1/(4\tilde{M})$, taking the supremum for $s \in [s_1, s_1+\delta)$, for any $\varepsilon < 1$, we have the following estimate

$$\begin{aligned}\sup_{s\in[s_1,s_1+\delta)} \mathbb{E}_t[Y_r^{2r}] &\le 2\mathbb{E}_t[Y_{s_1}^{2r}] + C_1 \cdot \varepsilon^{2r} + C_2 \cdot \varepsilon^{r} \\ &\le 2\mathbb{E}_t[Y_{s_1}^{2r}] + C \cdot \varepsilon^{r},\end{aligned}$$

for constants $C_1 = C_1(r, L)$ and $C_2 = C_2(r, L)$ depending only on r and L, and $C = C_1 + C_2$. This estimate holds on intervals $[t, t+\delta), [t+\delta, t+2\delta), \ldots, [t+n\delta, t+\varepsilon)$ for $n := [\frac{\varepsilon}{\delta}]$. Recalling that $Y_t = 0$, we have

$$\sup_{s\in[t,T]} \mathbb{E}_t[Y_s^{2r}] \le \sum_{i=0}^{n} (2^{i+1} - 1) \cdot C(r, L) \cdot \varepsilon^{r}.$$

which proves the estimate for the process Y with the constant $M_1 = C(r, L) \sum_{i=0}^{[1/\delta]} (2^{i+1} - 1)$ for $\varepsilon < 1$.

For the process (Z_s), since the perturbation $D_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}$ exists only in the drift term, when we apply the Itô's formula for Z_s^{2r}, the term Z_s^{2r-2} does not appear.

Following similar lines of reasoning as above, we obtain

$$\sup_{s\in[s_1,s_1+\delta)} \mathbb{E}_t[Z_s^{2r}] \le 2\mathbb{E}_t[Z_{s_1}^{2r}] + \tilde{C}\cdot\varepsilon^{2r}.$$

where $\tilde{C} = \tilde{C}(r,L)$. Therefore, for $\varepsilon < 1$,

$$\sup_{s\in[t,T]} \mathbb{E}_t[Z_s^{2r}] \le M_2(r,L)\varepsilon^{2r},$$

with $M_2 = \tilde{C}(r,L)\sum_{i=0}^{[1/\delta]}(2^{i+1}-1)$. Then $M = \max\{M_1(r,L), M_2(r,L)\}$ is the constant in the statement of the theorem. □

Denote $\xi_s^{t,\varepsilon,v} = Y_s^{t,\varepsilon,v} + Z_s^{t,\varepsilon,v}$. Hence, the dynamics of $\xi_s^{t,\varepsilon,v}$ follow

$$\begin{cases} d\xi_s = [A_s\xi_s + B_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]ds + [C_s\xi_s + D_s v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]dW_s, \quad s\in[t,T], \\ \xi_t = 0. \end{cases} \tag{5.9}$$

Lemma 5.3.4. *The following estimate holds uniformly in $s\in[t,T]$*

$$\mathbb{E}_t[\xi_s^{t,\varepsilon,v}] = O(\varepsilon). \tag{5.10}$$

Proof. Since the diffusion term of (Y_s) is square integrable, it disappears under conditional expectation $\mathbb{E}_t[Y_s]$. Recalling that (A_s) is deterministic, we have $\mathbb{E}_t[A_sY_s] = A_s\,\mathbb{E}_t[Y_s]$ so we can write an ODE for $\theta_s = \mathbb{E}_t[Y_s]$:

$$\begin{cases} d\theta_s = A_s\theta_s ds, \quad s\in[t,T] \\ \theta_t = 0. \end{cases}$$

Therefore, $\theta_s = 0$ and $\mathbb{E}_t[\xi_s] = \mathbb{E}_t[Z_s]$ for $s\in[t,T]$. By Jensen's inequality and

Lemma 5.3.3, we obtain

$$(\mathbb{E}_t[Z_s])^2 \le \mathbb{E}_t[Z_s^2] = O(\varepsilon^2), \quad \forall s \in [t,T]. \tag{5.11}$$

□

Remark 5.3.2. The proof of Lemma 5.3.4 reveals why (A_s) is assumed to be deterministic as this allows us to disregard (Y_s) in the estimate of conditional expectation of (ξ_s). Random (A_s) requires the use Lemma 5.3.4 whose statement cannot be strengthened for $\mathbb{E}_t[Y_s^2]$ (this is easy to see) resulting in an estimate $O(\varepsilon)$. Using Lemma 5.3.4 in its present form would weaken the estimate of $\mathbb{E}_t[\xi_s]$ to $O(\varepsilon^{1/2})$ which is too little for the proof of Proposition 5.3.1.

Alternative proof of Lemma 5.3.4. The lemma can be proved directly. By taking conditional expectation of both sides of (5.9) and using the fact that then the stochastic integral disappears, we obtain

$$\mathbb{E}_t[\xi_s] = \int_t^s \Big(A_s\, \mathbb{E}_t[\xi_s] + \mathbb{E}_t[D_s v_s] \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\Big) ds.$$

Denoting by $\hat{\xi} = \mathbb{E}_t[\xi_s]$ and $b_s = \mathbb{E}_t[D_s v_s]$, the above is rewritten as an ordinary differential equation

$$d\hat{\xi}_s = \big(A_s \hat{\xi}_s + b_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\big) ds$$

with a solution

$$\hat{\xi}_s = \int_t^{(t+\varepsilon)\wedge s} b_u e^{\int_t^u A_r dr} du.$$

Consider a differential equation $d\theta_s = \big(|A_s|\theta_s + |b_s| \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\big) ds$ with solution

$$\theta_s = \int_t^{(t+\varepsilon)\wedge s} |b_u| e^{\int_t^u |A_r| dr} du.$$

Clearly, $\hat{\xi}_s \le \theta_s$. Moreover, if $L = \max\{\operatorname{ess\,sup}|v_s|, \operatorname{ess\,sup}|D_s|, \sup|A_s|\}$, then

$$\theta_s \le \int_t^{(t+\varepsilon)\wedge s} L^2 e^{\int_t^u L dr} du \le \int_t^{(t+\varepsilon)\wedge s} L^2 e^{L(u-t)} du \le L^2 e^{L(T-t)} \varepsilon,$$

which completes the proof. □

Lemma 5.3.5. *The following equalities hold*

$$\begin{aligned}\mathbb{E}_t[P_T\xi_T^2] = \mathbb{E}_t\left[\int_t^T \{P_s D_s^2 v_s^2 \mathbb{1}_{\{s\in[t,t+\varepsilon)\}} - \xi_s^2 Q_s\} ds\right] \\ + \mathbb{E}_t\left[\int_t^T 2\xi_s(P_sB_s + P_sC_sD_s + K_sD_s)v_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}} ds\right] \quad a.s.\end{aligned} \tag{5.12}$$

and

$$\mathbb{E}_t[p(T;t)\xi_T] = \mathbb{E}_t\int_t^T \left\{-\xi_s Q_s X_s^* + (p(s;t)B_s + k(s;t)D_s)v_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right\} ds \quad a.s.. \tag{5.13}$$

Proof. Apply the Itô's formula to the process $P_s\xi_s^2$, where P_s is the solution of BSDE (5.6):

$$\begin{aligned}d(P_s\xi_s^2) &= \xi_s^2 dP_s + P_s d\xi_s^2 + dP_s d\xi_s^2 \\ &= \big[2P_sB_sv_s\xi_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}} + P_sD_s^2v_s^2\mathbb{1}_{\{s\in[t,t+\varepsilon)\}} \\ &\quad + 2\xi_sP_sC_sD_sv_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}} - \xi_s^2Q_s + 2\xi_sK_sD_sv_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\big]ds \\ &\quad + \left[\xi_s^2K_s + 2P_s\xi_s^2C_s + 2P_s\xi_sD_sv_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right] dW_s.\end{aligned} \tag{5.14}$$

We will use the localisation method to prove (5.12). Let

$$\tau_n := \inf\{s \geq t; \int_t^s (\xi_s^4 K_s^2 + \xi_s^4 P_s^2) ds > n\}.$$

Random variables τ_n are (finite-valued) stopping times. Indeed,

$$\int_t^s (\xi_u^4K_u^2 + \xi_u^4P_u^2)du \leq \big(\sup_{u\in[t,T]} \xi_u^4\big)\int_t^s K_u^2 du + \big(\sup_{u\in[t,T]} \xi_u^4\big)\cdot\big(\sup_{u\in[t,T]} P_u^2\big).$$

Since all coefficients in (5.9) are essentially bounded, (Krylov, 2008, Corollary 6, Section 2.5) implies that every moment of ξ is bounded. Also, by Lemma 5.3.1, we have $(P(\cdot), K(\cdot)) \in L^2_{\mathcal{F}}(\Omega; C(t,T;\mathbb{R})) \times L^2_{\mathcal{F}}(t,T;\mathbb{R})$. Then

$$\int_t^s (\xi_s^4 K_s^2 + \xi_s^4 P_s^2) ds < \infty \quad \text{a.s.}$$

and

$$\lim_{n\to\infty} \tau_n = \infty.$$

In view of (5.14), we have

$$\begin{aligned}\mathbb{E}_t[P_{T\wedge\tau_n}\xi^2_{T\wedge\tau_n}] = \mathbb{E}_t\left[\int_t^{T\wedge\tau_n} \{P_s D_s^2 v_s^2 \mathbb{1}_{\{s\in[t,t+\varepsilon)\}} - \xi_s^2 Q_s\} ds\right] \\ + \mathbb{E}_t\left[\int_t^{T\wedge\tau_n} 2\xi_s(P_s B_s + P_s C_s D_s + K_s D_s) v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}} ds\right].\end{aligned} \tag{5.15}$$

For each n we have a.s.

$$\left|P_{T\wedge\tau_n}\xi^2_{T\wedge\tau_n}\right| \le \sup_{s\in[t,T]} |P_s| \cdot \sup_{s\in[t,T]} \left|\xi_s^2\right|,$$

By Cauchy-Schwartz inequality

$$\mathbb{E}\left[\sup_{s\in[t,T]} |P_s| \cdot \sup_{s\in[t,T]} \left|\xi_s^2\right|\right] \le \left(\mathbb{E}[\sup_{s\in[t,T]} |P_T|^2]\right)^{\frac{1}{2}} \left(\mathbb{E}[\sup_{s\in[t,T]} |\xi_T|^4]\right)^{\frac{1}{2}} < \infty,$$

hence, conditional dominated convergence theorem (Williams, 1991, Lemma 9.7) implies

$$\lim_{\tau_n\to\infty} \mathbb{E}_t\left[P_{T\wedge\tau_n}\xi^2_{T\wedge\tau_n}\right] = \mathbb{E}_t\left[P_T\xi_T^2\right].$$

Similarly, quantities under conditional expectation on the right-hand side of (5.15) are dominated by

$$\begin{aligned}&\left|\int_t^{T\wedge\tau_n} \left(P_s D_s^2 v_s^2 \mathbb{1}_{\{s\in[t,t+\varepsilon)\}} - \xi_s^2 Q_s\right) ds\right| + \left|\int_t^{T\wedge\tau_n} 2\xi_s(P_s B_s + P_s C_s D_s + K_s D_s) v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}} ds\right| \\ &\le \int_t^T \left(\left|P_s D_s^2 v_s^2 \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right| + \left|\xi_s^2 Q_s\right|\right) ds + \int_t^T \left|2\xi_s(P_s B_s + P_s C_s D_s + K_s D_s) v_s \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right| ds.\end{aligned}$$

The expectation of the above bound is finite since it is dominated by

$$C\left(\mathbb{E}\int_t^{t+\varepsilon}|P_s|\,ds+\mathbb{E}_t\int_t^T|\xi_s|^2\,ds\right)+C\left(\mathbb{E}\int_t^{t+\varepsilon}|\xi_sP_s|\,ds+\mathbb{E}\int_t^{t+\varepsilon}|\xi_sK_s|\,ds\right)$$
$$\le C\left(\mathbb{E}\int_t^{t+\varepsilon}|P_s|\,ds+\mathbb{E}\int_t^T|\xi_s|^2\,ds\right)$$
$$+C\varepsilon^{1/2}\Big(\sup_{s\in[t,T]}\mathbb{E}\,|\xi_s|^2\Big)^{\frac{1}{2}}\Big[\Big(\int_t^{t+\varepsilon}\mathbb{E}\,|P_s|^2\,ds\Big)^{\frac{1}{2}}+\Big(\int_t^{t+\varepsilon}\mathbb{E}\,|K_s|^2\,ds\Big)^{\frac{1}{2}}\Big]<\infty.$$

Therefore, we can again apply conditional dominated convergence theorem to infer (5.12) from (5.15). The proof of (5.13) is similar. □

Corollary 5.3.3. *The following estimate holds*

$$\mathbb{E}_t\left[\int_t^T 2\xi_s(P_sB_s+P_sC_sD_s+K_sD_s)v_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}ds\right]=o(\varepsilon)\quad a.s.\qquad(5.16)$$

Proof. Using conditional form of the estimates from the above proof we get

$$\mathbb{E}_t\left[\int_t^T 2\xi_s(P_sB_s+P_sC_sD_s+K_sD_s)v_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}ds\right]$$
$$\le C\varepsilon^{1/2}\Big(\sup_{s\in[t,T]}\mathbb{E}_t\,|\xi_s|^2\Big)^{\frac{1}{2}}\Big[\Big(\int_t^{t+\varepsilon}\mathbb{E}_t\,|P_s|^2\,ds\Big)^{\frac{1}{2}}+\Big(\int_t^{t+\varepsilon}\mathbb{E}_t\,|K_s|^2\,ds\Big)^{\frac{1}{2}}\Big].$$

By Lemma 5.3.3, $\sup_{s\in[t,T]}\mathbb{E}_t\,|\xi_s|^2=O(\varepsilon)$. It, therefore, suffices to prove that

$$\Big(\int_t^{t+\varepsilon}\mathbb{E}_t\,|P_s|^2\,ds\Big)^{\frac{1}{2}}+\Big(\int_t^{t+\varepsilon}\mathbb{E}_t\,|K_s|^2\,ds\Big)^{\frac{1}{2}}=o(1),$$

which is clear given that both processes are square integrable. □

We have now collected all preliminary results required in the proof of Proposition 5.3.1.

Proof of Proposition 5.3.1. From the definition of the functional, we have

$$
\begin{aligned}
&2(J(t,X_t^*;u^{t,\varepsilon,v})-J(t,X_t^*;u^*))\\
&=\mathbb{E}_t\int_t^T[Q_s(X_s^{t,\varepsilon,v})^2+R_s(u^{t,\varepsilon,v})^2]ds+G\,\mathbb{E}_t[(X_T^{t,\varepsilon,v})^2]\\
&\quad-h(\mathbb{E}_t[X_T^{t,\varepsilon,v}])^2-2(\mu_1X_t^*+\mu_2)\mathbb{E}_t[X_T^{t,\varepsilon,v}]\\
&\quad-\mathbb{E}_t\int_t^T[Q_s(X_s^*)^2+R_s(u^*)^2]ds-G\mathbb{E}_t[(X_T^*)^2]+h(\mathbb{E}_t[X_T^*])^2+2(\mu_1X_t^*+\mu_2)\mathbb{E}_t[X_T^*]\\
&=\mathbb{E}_t\int_t^T\big\{Q_s[(X_s^{t,\varepsilon,v})^2-(X_s^*)^2]+R_s[(u^{t,\varepsilon,v})^2-(u^*)^2]\big\}ds\\
&\quad+G\,\mathbb{E}_t[(X_T^{t,\varepsilon,v})^2-(X_T^*)^2]-h[(\mathbb{E}_t[X_T^{t,\varepsilon,v}])^2-(\mathbb{E}_t[X_T^*])^2]\\
&\quad-2(\mu_1X_t^*+\mu_2)\mathbb{E}_t[X_T^{t,\varepsilon,v}-X_T^*].
\end{aligned}
$$

Recalling that $X_s^{t,\varepsilon,v}=X_s^*+\xi_s$ and $(a+b)^2-b^2=a(a+2b)$, we rewrite the functional

$$
\begin{aligned}
&2(J(t,X_t^*;u^{t,\varepsilon,v})-J(t,X_t^*;u^*))\\
&=\mathbb{E}_t\int_t^T Q_s[2X_s^*+\xi_s]\xi_sds+\mathbb{E}_t\int_t^{t+\varepsilon}R_s(2u_s^*+v_s)v_sds\\
&\quad+G\,\mathbb{E}_t[\xi_T^2]+2G\mathbb{E}_t[\xi_TX_T^*]-h(\mathbb{E}_t[\xi_T])^2-2h\mathbb{E}_t[\xi_T]\mathbb{E}_t[X_T^*]-2(\mu_1X_t^*+\mu_2)\mathbb{E}_t[\xi_T]\\
&=\mathbb{E}_t\int_t^T Q_s[2X_s^*+\xi_s]\xi_sds+\mathbb{E}_t\int_t^{t+\varepsilon}R_s(2u_s^*+v_s)v_sds\\
&\quad+G\,\mathbb{E}_t[\xi_T^2]+2\mathbb{E}_t[(GX_T^*-h\mathbb{E}_t[X_T^*]-\mu_1X_t^*-\mu_2)\xi_T]-h(\mathbb{E}_t[\xi_T])^2.
\end{aligned}
$$

In view of (5.10), we have $h(\mathbb{E}_t[\xi_T])^2=o(\varepsilon)$. By (5.12) and (5.16)

$$
G\mathbb{E}_t[\xi_T^2]=\mathbb{E}_t\int_t^T[P_sD_s^2v_s^2\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}-\xi_s^2Q_s]ds+o(\varepsilon).
$$

In the final step of the proof we collect above results and also apply (5.13):

$$
\begin{aligned}
&2(J(t,X_t^*;u^{t,\varepsilon,v}) - J(t,X_t^*;u^*)) \\
&= 2\mathbb{E}_t\int_t^T Q_s X_s^*\xi_s ds + \mathbb{E}_t\int_t^{t+\varepsilon} R_s(2u_s^*+v_s)v_s ds + \mathbb{E}_t\int_t^T [P_s D_s^2 v_s^2 \mathbb{1}_{\{s\in[t,t+\varepsilon)\}}]ds \\
&\quad + 2\mathbb{E}_t\int_t^T \left[-\xi_s Q_s X_s^* + (p(s;t)B_s + k(s;t)D_s)v_s\mathbb{1}_{\{s\in[t,t+\varepsilon)\}}\right] ds + o(\varepsilon) \\
&= \mathbb{E}_t\int_t^{t+\varepsilon}\left\{2[R_s u_s^* + p(s;t)B_s + D_s k(s;t)]v_s + [P_s D_s^2 + R_s]v_s^2\right\} ds + o(\varepsilon) \\
&= \mathbb{E}_t\int_t^{t+\varepsilon}\left\{2\Lambda(s;t)v_s + H_s v_s^2\right\} ds + o(\varepsilon),
\end{aligned}
$$

where $\Lambda(s;t)$ and H_s are defined in the statement of the proposition. □

Lemma 5.3.6. *An admissible control u^* is an equilibrium control if and only if*

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[\Lambda(s;t)v_s]ds \geq 0 \quad a.s. \tag{5.17}$$

for any $t\in[0,T)$ and $v\in\mathcal{L}^\infty_{\mathcal{F}}(t,T;\mathbb{R})$.

Proof. If u^* is an equilibrium control, by Proposition 5.3.1, for any $A\in\mathcal{F}_t$, $v\in\mathcal{L}^\infty_{\mathcal{F}}(t,T;\mathbb{R})$ and $t\in[0,T)$

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\left(\Lambda(s;t)v_s + \frac{1}{2}H_s v_s^2\right)\mathbb{1}_A\right\} ds \geq 0. \tag{5.18}$$

Fix t and v. For any $N>0$, apply (5.18) to perturbation $\frac{1}{N}v$:

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\left(\frac{1}{N}\Lambda(s;t)v_s + \frac{1}{2N^2}H_s v_s^2\right)\mathbb{1}_A\right\} ds \geq 0.$$

Multiplying both sides by N and taking the limit as $N\to\infty$ yields

$$\lim_{N\to\infty}\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\left(\Lambda(s;t)v_s + \frac{1}{2N}H_s v_s^2\right)\mathbb{1}_A\right\} ds \geq 0.$$

Recall that $H_s = R_s + D_s^2 P_s \geq 0$ a.s. We can obtain an upper bound for

$\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}[\frac{1}{N}H_s v_s^2 \mathbb{1}_A]ds$ which does not depend on ε. Indeed,

$$\begin{aligned}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\frac{1}{N}H_s v_s^2\mathbb{1}_A\right\}ds &\leq C\frac{1}{\varepsilon}\frac{1}{N}\mathbb{E}\left\{\int_t^{t+\varepsilon}H_s\mathbb{1}_A ds\right\}\\ &\leq C\frac{1}{\varepsilon}\frac{1}{N}\mathbb{E}\left\{\varepsilon\sup_{s\in[t,t+\varepsilon)}|H_s|\,\mathbb{1}_A\right\}\leq\frac{C}{N},\end{aligned}$$

where, as before, the constant C may differ from line to line. Then,

$$\begin{aligned}0 &\leq \lim_{N\to\infty}\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\left(\Lambda(s;t)v_s+\frac{1}{2N}H_s v_s^2\right)\mathbb{1}_A\right\}ds\\ &\leq \lim_{N\to\infty}\liminf_{\varepsilon\downarrow 0}\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\Lambda(s;t)v_s\mathbb{1}_A\right\}ds+\frac{C}{N}\right)\\ &= \liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\Lambda(s;t)v_s\mathbb{1}_A\right\}ds.\end{aligned}$$

Conversely, assume (5.17) holds, since $H_s\geq 0$ almost surely for $s\in[t,T]$, then

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{\frac{1}{2}H_s v_s^2\right\}ds\geq 0\quad a.s.$$

which gives

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left\{(\Lambda(s;t)v_s+\frac{1}{2}H_s v_s^2)\right\}ds\geq 0\quad a.s.$$

□

Theorem 5.3.4. *u^* is an equilibrium control if and only if*

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds=0,\quad a.s. \tag{5.19}$$

for any $t\in[0,T)$ and $v\in\mathcal{L}^\infty_{\mathcal{F}}(t,T;\mathcal{R})$.

Proof. By Lemma 5.3.6, it suffices to show that

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[\Lambda(s;t)v_s]\,ds\geq 0\quad\Leftrightarrow\quad\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds=0.$$

If $\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[\Lambda(s;t)v_s]ds\geq 0$ holds almost surely, then setting $v_s = -\operatorname{sgn}(\Lambda(s;t))$, we have

$$-\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[\Lambda(s;t)\cdot\operatorname{sgn}(\Lambda(s;t))]ds\geq 0\quad a.s.$$

Hence,

$$\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds\leq 0,$$

this yields

$$0\geq\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds\geq\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds\geq 0.$$

Therefore,

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]ds=0.$$

Conversely, if $\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]ds = 0$ holds almost surely, this implies $\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[|\Lambda(s;t)|]\,ds=0$. Then, for any $v\in\mathcal{L}^\infty_{\mathcal{F}}(t,T;\mathcal{R})$, we have

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}_t\int_t^{t+\varepsilon}[\Lambda(s;t)v_s]ds &\geq -\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}_t[|\Lambda(s;t)|\,|v_s|]ds\\ &\geq -C\cdot\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}_t[|\Lambda(s;t)|]ds\\ &= 0\quad a.s.\end{aligned}$$

□

5.4 Conclusion

In this chapter, we study the open-loop equilibrium control for the linear-quadratic problem in continuous time. In contrast to the existing literature (see Hu *et al.* (2012)), where the equilibrium control is characterised via sufficient condition, we provide a necessary and sufficient condition for open-loop equilibrium control. The motivation is that the investor may have her own preference for the investment policy, and then our condition can be used to test her policy. Existing literature (see Djehiche & Huang (2016)), study necessary and sufficient condition for mean field type of equilibrium control under different assumption. The assumptions require the boundedness of the state dynamic. However, within the linear-quadratic framework, this condition is not satisfied.

Moreover, the equilibrium condition of a adapted process perturbation is identical to the condition of a random variable perturbation. Therefore, the type of equilibrium strategy does not depend on the perturbation type. In conclusion, the equilibrium approach can be viewed as a game for the investor. The equilibrium investor has to choose two investment plans: short-term investment and long-term investment plans. Although the investors might have a very good long-term investment plan with a good rewards, they still have to decides the short-term investment plan. As a result, the final investment strategies are modified and adjusted to the corresponding short-term plan. The game for the investor becomes choosing the best combination of the investment strategy.

Our results of equilibrium rationale indicate that there exists a family of equilibrium strategies by choosing different short-term investment plans. Therefore, the drawbacks we discussed in previous chapter are only responsible for the specific short-term plan. Regarding to improve the performance of the equilibrium strategy, it is crucial to decide a sensible and profitable short-term plan.

Chapter 6

Mean-Variance puzzle

6.1 Introduction

In this chapter, we are trying to address the feature of equilibrium control in discrete time setting (present both in open-loop and closed-loop formulations) that the investment in the risky asset increases as time approaches the investment horizon. This is what we call the mean-variance puzzle. The equilibrium concept has been understood through the game theoretic framework under finite time horizon. The basic idea is that the investor chooses a decision at any time t by agreeing to implement the policy she thinks that would also be optimal for future times. She then views the problem as a cooperative game only with her future self. Therefore, the investor will be able to find an equilibrium policy as long as the future time is finite. Provided there is a long enough investment horizon, the mean-variance puzzle manifests itself as the control goes to zero at the starting point. To ensure there is a long-enough investment horizon, we allow the starting point to move backwards in time.

6.2 Solving the mean-variance puzzle with Present-Biased preference

For the sake of simplicity, we assume that there is only one risky asset and one risk-free asset available in the market. The risky asset has the excess return factor $\mathbf{B}_t$, which mean μ_t and variance σ_t^2 for all $t = 0, \ldots, T-1$. In Chapter 3 and Chapter 4, we have shown that the open and closed-loop equilibrium strategies admit the following form:

$$u_t^{op} = O_t x_t,$$

and

$$u_t^{cl} = C_t x_t.$$

We can further simplify the closed-form solutions of open-loop (3.11) and closed-loop (4.7) equilibrium strategies, the following representations can be obtained:

$$u_t^{op} = O_t x_t = \frac{\mu_t}{\gamma \cdot \mathbb{E}_t\big[\prod_{i=t+1}^{T-1}(A_i + B_{i+1}O_i)\big]\sigma_t^2} x_t, \tag{6.1}$$

and

$$u_t^{cl} = C_t x_t = \frac{\mathbb{E}_t\big[\prod_{i=t+1}^{T-1}(A_i + B_{i+1}C_i)\big]\mu_t}{\gamma \cdot \mathbb{E}_t\big[\prod_{i=t+1}^{T-1}(A_i + B_{i+1}C_i)^2\big]\sigma_t^2} x_t. \tag{6.2}$$

We notice that, for B_t independently and identically distributed, the expression in the denominator of open-loop control (6.1) can be written as:

$$\begin{aligned} \mathbb{E}_t\Big[\prod_{i=t+1}^{T-1}(A_i + B_{i+1}O_i)\Big] &= \prod_{i=t+1}^{T-1} \mathbb{E}_t[A_i + B_{i+1}O_i] \\ &= \prod_{i=t+1}^{T-1} \big\{A_i + \mathbb{E}_t[B_{i+1}]\,\mathbb{E}_t[O_i]\big\}. \end{aligned}$$

The expression inside of the curly bracket is always positive for any $i \in \{t+1, \ldots, T-1\}$ and $\mathbb{E}_t[B_{i+1}] > 0$ (which is a reasonable assumption). Therefore, as t moves backwards, the control (6.1) will decay towards 0. Figure 6.1 shows plots of both controls (6.1) and (6.2) where the mean-variance puzzle is evident. This is in agreement with the numeric results in (Björk *et al.*, 2014) and (Hu *et al.*, 2012), where both open-loop and closed-loop equilibrium strategies would decrease rapidly as time goes backwards.

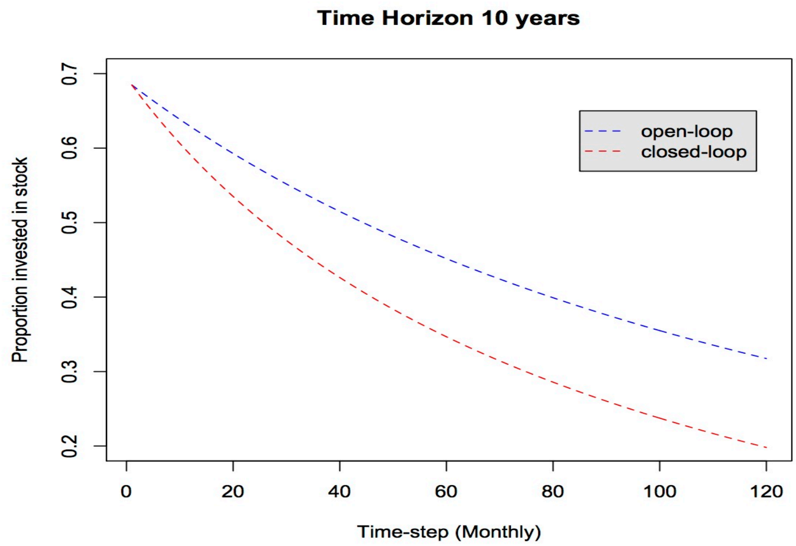


Figure 6.1: The comparison between open-loop and closed-loop equilibrium strategies as time moves away from terminal

Assuming that B_t is independent and identically distributed, then the result implies that the investor would invest much less at the beginning than towards maturity even if she faces exactly the same investment environment. This is because the variance term penalizes perturbation quadratically as well as the expectation term only increases linearly. If the investor deviates slightly from the equilibrium, as long as it's far away from the maturity, the variance would eventually penalise the perturbation more than any benefits added by the expectation term. As a consequence, the control should be sufficiently small in order to keep the variance low enough. In our view, this is unreasonable in practice and this feature is unrealistic. In the remaining of this section, we concentrate on addressing this unrealistic feature (the mean-variance puzzle).

Recall the independent and identical increment case, the open-loop control (3.12) can be written as

$$u_t = \frac{\mu}{\gamma\ \alpha_{t+1}\ \sigma^2} x_t,$$

where

$$\alpha_t = \alpha_{t+1} \cdot A_t + \frac{\mu^2}{\gamma\sigma^2} \quad \text{with} \quad \alpha_T = 1.$$

Such a strategy is in the form of one-period mean-variance strategy with the risk-aversion coefficient $\gamma\alpha_{t+1}$. Therefore, we can modify the risk-aversion depending on time, forcing the strategy to be riskier if the investor is far away from the maturity. We will call this type of risk aversion Present-Biased preference. The name follows from a long history on studying the Strotz-Pollak equilibrium problem (see for example (Phelps & Pollak, 1968) and (Peleg & Yaari, 1973)) with Present-Biased preference. The fundamental framework in (Phelps & Pollak, 1968) postulates that the preference for the investor at time t is represented by the sum of the utilities of future payoffs. Since the utilities are applied identically for the time t and future, then the future payoff at time $t+k$ should be discounted by a factor $\delta\beta^k$. The constant factor δ represents the degree of selfishness and β is the discounting factor. In what follows, we consider a Present-Biased preference in the same spirit to avoid th control decay to zero.

6.2.1 Case 1: $A_t = 1$ for all t (zero interest rate)

In this section, we will use the arithmetic excess return and set the interest rate equal to zero. Suppose the investor is sitting at time $t = T-k$, which is k periods away from the maturity. Then we have

$$u_t := u_{T-k} = \frac{\mu}{\gamma_{T-k}\ \alpha_{T-k+1}\ \sigma^2} \cdot X_{T-k}, \tag{6.3}$$

where

$$
\begin{aligned}
\alpha_{T-k+1} &= \alpha_{T-k+2} \cdot 1 + \frac{\mu^2}{\gamma_{T-k+1}\,\sigma^2} \\
&= \alpha_{T-k+3} + \frac{\mu^2}{\sigma^2}\left(\frac{1}{\gamma_{T-k+1}} + \frac{1}{\gamma_{T-k+2}}\right) \\
&= \alpha_T + \frac{\mu^2}{\sigma^2}\sum_{i=1}^{k-1}\frac{1}{\gamma_{T-i}} \\
&= 1 + \frac{\mu^2}{\sigma^2}\sum_{i=1}^{k-1}\frac{1}{\gamma_{T-i}}.
\end{aligned}
$$

Therefore, the denominator of control (6.3), we get

$$
\gamma_{T-k} \cdot \alpha_{T-k+1} = \gamma_{T-k} + \frac{\mu^2}{\sigma^2}\sum_{i=1}^{k-1}\frac{\gamma_{T-k}}{\gamma_{T-i}}. \tag{6.4}
$$

Suppose we choose the risk-aversion in the following form:

$$
\gamma_{T-k} = \gamma e^{-q(k-1)},
$$

for some positive constant q and γ. Under such risk attitude, if the investor is far away from the maturity (large k), then variance is less penalised and she invests more in the stocks. Therefore, the summation in Equation (6.4) can be written as:

$$
\sum_{i=1}^{k-1}\frac{\gamma_{T-k}}{\gamma_{T-i}} = e^{-qk}\sum_{i=1}^{k-1}e^{qi}.
$$

We are interested in the asymptotic behavior as time goes backwards, which means as k tends to infinity. We have

$$\begin{aligned}
\gamma_{T-k}\cdot\alpha_{T-k+1} &= \lim_{k\to\infty}\{\frac{1}{k^q}+\frac{\mu^2 e^{-qk}}{\sigma^2}\sum_{i=1}^{k-1}e^{qi}\} \\
&= \lim_{k\to\infty}\{\frac{1}{k^q}+\frac{\mu^2 e^{-qk}}{\sigma^2}\cdot\frac{e^q-e^{qk}}{1-e^q}\} \\
&= \lim_{k\to\infty}\{\frac{1}{k^q}+\frac{\mu^2}{\sigma^2}\cdot\frac{e^{-q(k-1)}-1}{1-e^q}\} \\
&= \frac{\mu^2}{\sigma^2}\cdot\frac{1}{e^q-1}.
\end{aligned}$$

In practice, the exponential risk-aversion eliminates the variance and the investor would immediately try to maximise the expectation. Therefore, we would like the exponential risk-aversion to decay as slow as possible. We suggest to consider the risk-aversion in the form:

$$\gamma_{T-k}=\gamma e^{-\frac{(k-1)}{T}}$$

In the case when k is small, such time-dependent risk-aversion coefficient constant risk-aversion. In the case when k is large, the investor becomes more risky. Consequently, as k tends to infinity, the equilibrium control tends to

$$u_{T-k}\to\frac{\mu}{\sigma^2}\cdot\frac{\sigma^2(e^{\frac{1}{T}}-1)}{\mu^2}\cdot X_{T-k}=\frac{e^{\frac{1}{T}}-1}{\mu}\cdot X_{T-k}$$

In the case $\gamma_{T-k}=\gamma e^{-\frac{(k-1)}{T}}$, the original functional can be reformulated as

$$J(t,x_t,s_t;\bar{\mathbf{u}}_t)=\mathbb{E}_t[e^{-\frac{(k-1)}{T}}X_T]-\frac{\gamma}{2x_t}\mathrm{Var}_t(e^{-\frac{(k-1)}{T}}X_T)$$

This is a Present-Biased preference mean-variance problem with the discounting fator $\beta=e^{-\frac{1}{T}}$.

Example: Black-Scholes model with present-biased risk aversion

Recall the mean-variance problem with present-biased preference and i.i.d B_t:

$$J(t, x_t; \bar{\mathbf{u}}_t) = \mathbb{E}_t\big[X_T\big] - \frac{3e^{-(T-t-1)}}{2x_t}\mathrm{Var}_t\big(X_T\big).$$

The figure below shows the open-loop strategies (the proportion of wealth invested in stock) for 30-years investment.

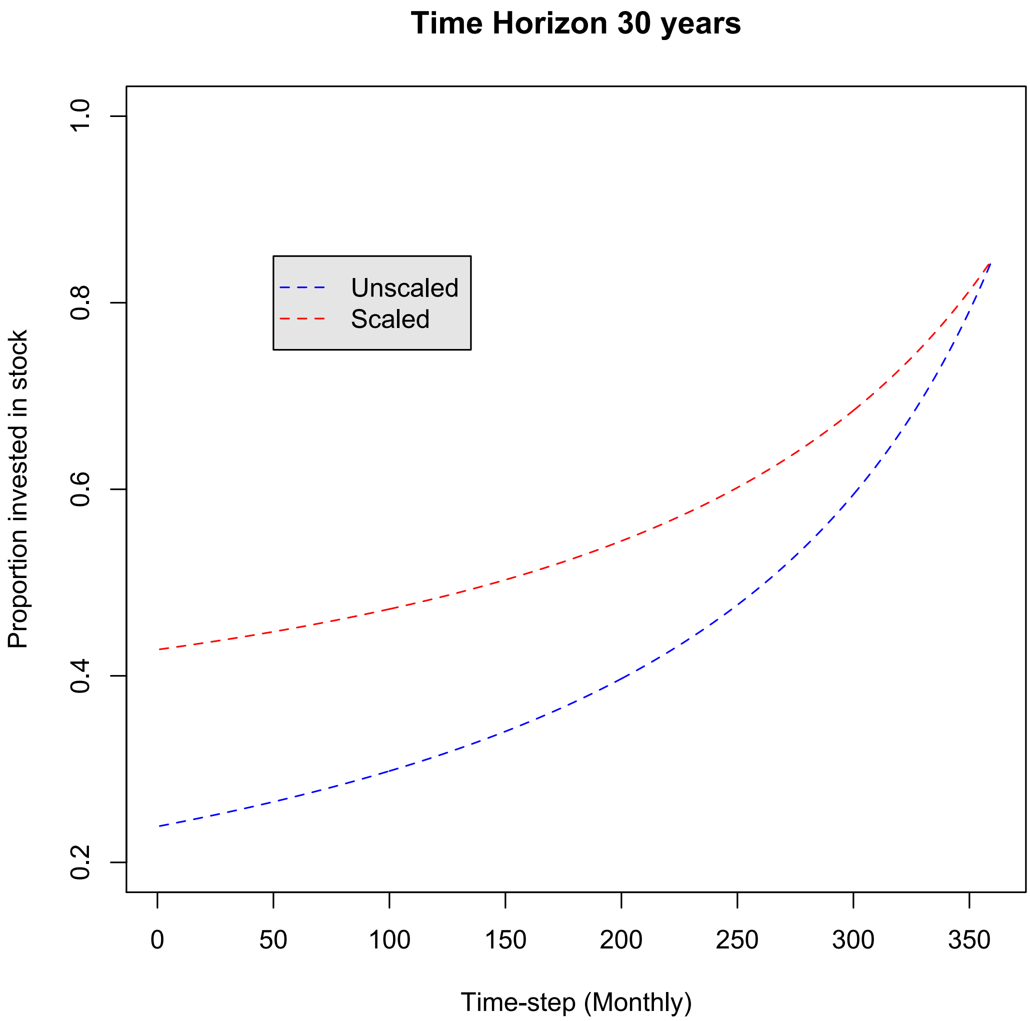


Figure 6.2: The comparison between constant risk-aversion and present-biased risk-aversion. The red dash line indicates the strategy using present-biased risk averision, whereas the blue dash line indicates the constant risk aversion

We see that, with Present-Biased risk aversion, the strategy does not go to zero as time goes backwards. However, it is still observed that the strategy becomes riskier towards the terminal time.

6.2.2 Case 2: $A_t > 1$ for all t. (positive interest rate)

Suppose the investor is sitting at time $t = T - k$, which is k periods away from the maturity. Then from the previous result, we have

$$u_t := u_{T-k} = \frac{\mu}{\gamma_{T-k}\ \alpha_{T-k+1}\ \sigma^2} \cdot X_{T-k},$$

where

$$\begin{aligned}
\alpha_{T-k+1} &= \alpha_{T-k+2} \cdot A_{T-k+1} + \frac{\mu^2}{\gamma_{T-k+1}\ \sigma^2} \\
&= \alpha_{T-k+2} \cdot A_{T-k+2} \cdot A_{T-k+1} + \frac{\mu^2}{\sigma^2}\left(\frac{1}{\gamma_{T-k+1}} + \frac{A_{T-k+1}}{\gamma_{T-k+2}}\right) \\
&= \alpha_T \cdot \prod_{i=1}^{k-1} A_{T-i} + \frac{\mu^2}{\sigma^2} \sum_{i=1}^{k-1} \frac{\prod_{j=1}^{k-1-i} A_{T-k+j}}{\gamma_{T-i}} \\
&= \prod_{i=1}^{k-1} A_{T-i} + \frac{\mu^2}{\sigma^2} \sum_{i=1}^{k-1} \frac{\prod_{j=1}^{k-1-i} A_{T-k+j}}{\gamma_{T-i}}
\end{aligned}$$

Recall that $A_i = e^r$ for all i, then

$$\alpha_{T-k+1} = e^{r\cdot(k-1)} + \frac{\mu^2}{\sigma^2} \sum_{i=1}^{k-1} \frac{e^{r\cdot(k-1-i)}}{\gamma_{T-i}}$$

Suppose we choose the risk-aversion in the following form:

$$\gamma_{T-k} = \gamma e^{-r\cdot q\cdot k},$$

for some positive constant $q > 1$ and γ. Therefore,

$$\begin{aligned}
\gamma_{T-k} \cdot \alpha_{T-k+1} &= \gamma e^{-r\cdot q\cdot k} e^{r\cdot(k-1)} + \gamma e^{-r\cdot q\cdot k} \cdot \frac{\mu^2}{\sigma^2} \cdot \sum_{i=1}^{k-1} \frac{e^{r\cdot(k-1-i)}}{\gamma e^{-r\cdot q\cdot i}} \\
&= \gamma e^{(1-q)rk-r} + \frac{\mu^2}{\sigma^2} \cdot \sum_{i=1}^{k-1} e^{rk-r-ri-rqk+rqi} \\
&= \gamma e^{(1-q)rk-r} + \frac{\mu^2}{\sigma^2} \cdot e^{r(k-1-qk)} \sum_{i=1}^{k-1} e^{(q-1)ri} \\
&= \gamma e^{(1-q)rk-r} + \frac{\mu^2}{\sigma^2} \cdot e^{r(k-1-qk)} \cdot \frac{e^{(q-1)r} - e^{(q-1)rk}}{1 - e^{(q-1)r}} \\
&= \gamma e^{(1-q)rk-r} + \frac{\mu^2}{\sigma^2} \cdot \frac{e^{(1-q)rk+r(q-2)} - e^{-r}}{1 - e^{(q-1)r}}
\end{aligned}$$

Since $q > 1$, then as time goes backwards, we have

$$\lim_{k\to\infty} \gamma_{T-k} \cdot \alpha_{T-k+1} = \frac{\mu^2}{\sigma^2} \cdot \frac{e^{-r}}{e^{(q-1)r} - 1}.$$

Hence, as k tends to infinity, the equilibrium control tends to

$$u_{T-k} \to \frac{e^r(e^q - 1)}{\mu} \cdot X_{T-k}.$$

Since our risk-aversion decays exponentially fast, as k tends to infinity, the penalty from the risk is negligible. Therefore, the asymptotic open-loop equilibrium control does not depend on the volatility of the stocks. On the other hand, when we use the parameter $q > 1$, the discount rate is larger than the interest rate. This suggests an alternative way to look at the model in the case where interest rate is negative.

6.2.3 Case 3: $A_t \in (0, 1)$ for all t. (negative interest rate)

There are two circumstances for A_t being in such range: Firstly, when the inflation rate is higher than the interest rate. The central bank tends to implement the policy based on the Friedman's k-percent rule. Milton Friedman, a Nobel-prize-winning economist, proposed that the central bank should increase the money supply by a constant percentage rate every year, irrespective of business cycles. In many developing countries, especially in China, the inflation rate is on average around 3% to 4% per year and the interest rate for 1-year zero coupon bond is 1.75%. If the investor deposits the money into the bank account, then the purchasing power of the money would actually depreciate. However, the economy in the developing countries is growing fast, which makes the real estate or stocks more attractive than the zero-coupon bond. For example, an investor decides to invest from 1999 to 2014 for 15 years. The average return rate for 1-year zero coupon is 2% in China, in which the equivalent return rate is 35% for 15 years. However, the Shanghai Composite Stock Market Index raised from 1123.70 to 3000, in which the return rate is 300%. Therefore, the investor should buy stocks rather than make savings to keep the value of money.

This situation is not uncommon in modern market. In 2015, Switzerland has become the first government to sell 10-year bonds with a yield of -0.055%. Sweden, Denmark, European and Japanese Central Bank have followed and cut their key interest rate below zero. This means investors buying risk-free assets and holding to maturity will not get the same money back. This situation encourages the investor to invest in stock in order to keep the value of her assets, which makes the investor more risk-seeking.

Hence, we consider the wealth model as following:

$$X_{t+1} = A_t X_t + B_{t+1} u_t$$

where

$$A_t = e^{-c},$$

$$B_{t+1} = \frac{S_{t+1}}{S_t} - e^{-c},$$

for some positive constant c. We can understand the model as following: in the case where the interest rate is negative, then $r = -c$. In the case where the inflation rate is higher than the interest rate, then (Interest rate$-$Inflation rate) $= -c$. Such a model indicates that if the investor decides not to invest the money into stock market, then the real value of money would actually depreciate. On the other hand, Hence,

$$\begin{aligned}\alpha_{T-k+1} &= \alpha_{T-k+2} \cdot A_{T-k+1} + \frac{\mu^2}{\gamma\,\sigma^2} \\ &= \alpha_{T-k+2} \cdot A_{T-k+2} \cdot A_{T-k+1} + \frac{\mu^2}{\sigma^2}\big(\frac{1}{\gamma} + \frac{A_{T-k+1}}{\gamma}\big) \\ &= \alpha_T \cdot \prod_{i=1}^{k-1} A_{T-i} + \frac{\mu^2}{\gamma\,\sigma^2} \sum_{i=1}^{k-1} \prod_{j=1}^{k-1-i} A_{T-k+j} \\ &= e^{-(k-1)\cdot c} + \frac{\mu^2}{\gamma\,\sigma^2} \sum_{i=1}^{k-1} e^{-(k-1-i)\cdot c}.\end{aligned}$$

Then,

$$\gamma \cdot \alpha_{T-k+1} = e^{-(k-1)\cdot c} \cdot \gamma + \frac{\mu^2}{\sigma^2} \sum_{i=1}^{k-1} e^{-(k-1-i)\cdot c}.$$

We can see the last term is the geometric series. We can obtain

$$\begin{aligned}\sum_{j=1}^{k-1} e^{-(k-1-i)\cdot c} &= e^{(1-k)\cdot c} \cdot \sum_{j=1}^{k-1} e^{c\cdot i} \\ &= e^{(1-k)\cdot c} \cdot \frac{e^c - e^{ck}}{1-e^c} \\ &= \frac{e^{(2-k)\cdot c} - e^c}{1-e^c}.\end{aligned}$$

Hence, as t goes backward, and as k tends to infinity, we have

$$\lim_{k\to\infty} \gamma \cdot \alpha_{T-k+1} = \frac{e^c}{e^c - 1} \cdot \frac{\mu^2}{\sigma^2} = \frac{1}{1-e^{-c}} \cdot \frac{\mu^2}{\sigma^2}.$$

and the equilibrium control tends to

$$\begin{aligned}u_{T-k} &\to \frac{\mu}{\sigma^2} \cdot X_{T-k} \cdot \frac{\sigma^2}{\mu^2} \cdot (1-e^{-c}) \\ &= \frac{1-e^{-c}}{\mu} \cdot X_{T-k}, \qquad \text{as } k \text{ tends to infinity.}\end{aligned}$$

Notice that this control does not go to zero as k tends to infinity, even though the Present-Biased preference is not used. That is because the effect of negative interest rate counteracts the penalty for the variance.

Remark. One can always direct the problem into the case of negative interest rate by considering an anticipated inflation rate that is higher than the risk-free interest rate.

6.3 Conclusion

This chapter presents the idea of using state-dependent present-biased preference as the risk-aversion parameter. Our analysis addresses the drawbacks of state-dependent risk-aversion which is proposed by Björk *et al.* (2014). The focus is given to analyse the investment behaviour for equilibrium strategy from the behavioural economic perspective. Our result points to the risk attitude of the investor corresponding to the investment time point.

The present-biased risk aversion has been defined e^{-qk}, where k is the number of periods away from the maturity and q is degree of bias for the risk attitude as time goes backward. With using a present-biased risk aversion parameter in i.i.d case, we have shown that the open-loop equilibrium strategy tends to a constant proportion invested in the stock with respect to differnt wealth.

Chapter 7

Numerical Simulation

Based on the equilibrium strategies presented in Chapter 3 and Chapter 4, this chapter assesses the performance of equilibrium strategies together with pre-commitment strategies and explores numerically the impact of equlibrium investors' behaviour. The closed-form solution of pre-commitment strategy in Li & Ng (2000) will be introduced in next section. A variety of investment situations such as an identical and independent market return, a booming market and a depression market will be discussed. We analyse the investment impact of investors' present-biased preference in performance evaluation by varying the value of biased level q. Finally, we compare the investment performance by allowing investors to re-evaluate regularly.

7.1 Pre-commitment v.s. Equilibrium: I.I.D model

In this section, we study the difference between the performance of the pre-commitment and equilibrium strategies. Suppose there are two assets available in the market: one risk-free asset with risk-free rate r and one risky asset with independent and identical investment return: mean $\mathbb{E}[B]$ and variance $\mathrm{Var}(B)$. We refer to the results for pre-commitment strategy from Li & Ng (2000)[Section

3]. The optimal pre-commitment strategy $u^{p,*}$ follows:

$$u_t^{p,*} = -\frac{e^r\,\mathbb{E}[B]}{\mathbb{E}[B^2]}X_t + \frac{1}{2}(bX_0 + \frac{\nu}{2\gamma a})(\frac{A^1}{A^2})^{T-t-1}\frac{\mathbb{E}[B]}{\mathbb{E}[B^2]} \quad \text{for } t = 0, \ldots, T-1, \tag{7.1}$$

where

$$\begin{aligned}
&K = \frac{(\mathbb{E}[B])^2}{\mathbb{E}[B^2]}, \quad A^1 = e^r - \frac{e^r(\mathbb{E}[B])^2}{\mathbb{E}[B^2]}, \quad A^2 = e^{2r} - \frac{(e^rE[B])^2}{\mathbb{E}[B^2]} \\
&K_t^1 = K\frac{(A^1)^{T-t-1}}{2(A^2)^{T-t-1}}, \qquad K_t^2 = K(\frac{(A^1)^{T-t-1}}{2(A^2)^{T-t-1}})^2 \\
&\mu = (A^1)^T, \qquad \tau = (A^2)^T, \qquad \nu = \sum_{t=0}^{T-1}(A^1)^{T-t-1}K_t^1 \\
&a = \frac{\nu}{2} - \nu^2, \qquad b = \frac{\mu\nu}{a}, \qquad c = \tau - \mu^2 - ab^2.
\end{aligned}$$

r	0.05	$\mathbb{E}[B]$	0.1
Var(B)	0.09	X_0	1
γ	0.5		

Table 7.1: Parameters setting for the market with independent and identical returns

Table 7.1 shows the values for different parameter. In this case, we have $K = 0.158$, $A^1 = 0.885$, $A^2 = 0.930$, $\mu = 0.295$, $\nu = 0.486$, $a = 0.0367$, $b = 3.297$, and $c = 0$. The expected terminal wealth and variance under the pre-commitment strategy $u^{p,*}$ with respect to a given risk-aversion parameter γ follow:

$$\mathbb{E}[X_T] = e^{rT}X_0 + \frac{1-(1-K)^T}{2\gamma(1-K)^T} \tag{7.2}$$

$$\text{Var}(X_T) = \frac{1-(1-K)^T}{4\gamma(1-K)^T}. \tag{7.3}$$

The analytical expression of the mean-variance efficient frontier Li & Ng

(2000)[Section 3, Equation (27)] is

$$\mathrm{Var}(X_T) = \frac{(1-K)^T}{1-(1-K)^T}(\mathbb{E}[X_T] - X_0 e^{rT}) \quad \text{for} \quad \mathbb{E}[X_T] \geq X_0(e^r)^T \tag{7.4}$$

We set initial wealth X_0 equal to 1 and simulate a number of different wealth trajectories by using the open-loop equilibrium (Corollary 3.3.6), closed-loop equilibrium (Corollary 4.3.3) and pre-commitment strategies (equation (7.1)).The parameters for simulation are given in 7.1. Next, we record the terminal value X_T with respect to each simulation. For n simulations, we obtain a sequence of terminal wealth $(X_T^1, X_T^2, \dots, X_T^n)$. Denote $\mathbb{E}_{t=0,X_0}[\cdot] = \mathbb{E}[\cdot|t=0, X_0=1]$ and $\mathrm{Var}_{t=0,X_0}[\cdot] = \mathrm{Var}[\cdot|t=0, X_0=1]$. We can calculate the mean $\mathbb{E}_{t=0,X_0}[X_T]$ and the standard deviation $Std_{t=0,X_0}[X_T]$. Then, we will evaluate the performance by using the Sharpe ratio of simulated terminal wealth $\frac{\mathbb{E}_{t=0,X_0}[X_T]}{Std_{t=0}[X_T]}$. Furthermore, we will compare the performance in following two situations:

1. The different investment periods: 10 years and 40 years.
2. Using the present-biased risk aversion.

Figure 7.1 shows the probability density function of terminal wealth for $X_0 = 1$. For 10 years investment period, the pair $(Std^{u^p}_{t=0,x_0}[X_T], \mathbb{E}^{u^p}_{t=0}[X_T]) = (1.401584, 3.511251)$, while the open-loop equilibrium strategy with constant risk-aversion gives $(Std^{u^o}_{t=0,x_0}[X_T], \mathbb{E}^{u^o}_{t=0,x_0}[X_T]) = (0.7824366, 2.4664584)$ and the closed-loop equilibrium strategy with constant risk-aversion gives $(Std^{u^c}_{t=0,x_0}[X_T], \mathbb{E}^{u^c}_{t=0}[X_T]) = (0.5628208, 2.1860717)$. The Sharpe ratios for terminal wealth for pre-commitment, open-loop and closed-loop strategies are 2.5, 3.15 and 3.88, respectively. Although the pre-commitment gives better expected return $\mathbb{E}[X_T]$, the pre-commitment is more risky compared to the equilibrium strategy due to the left tail of the density function.

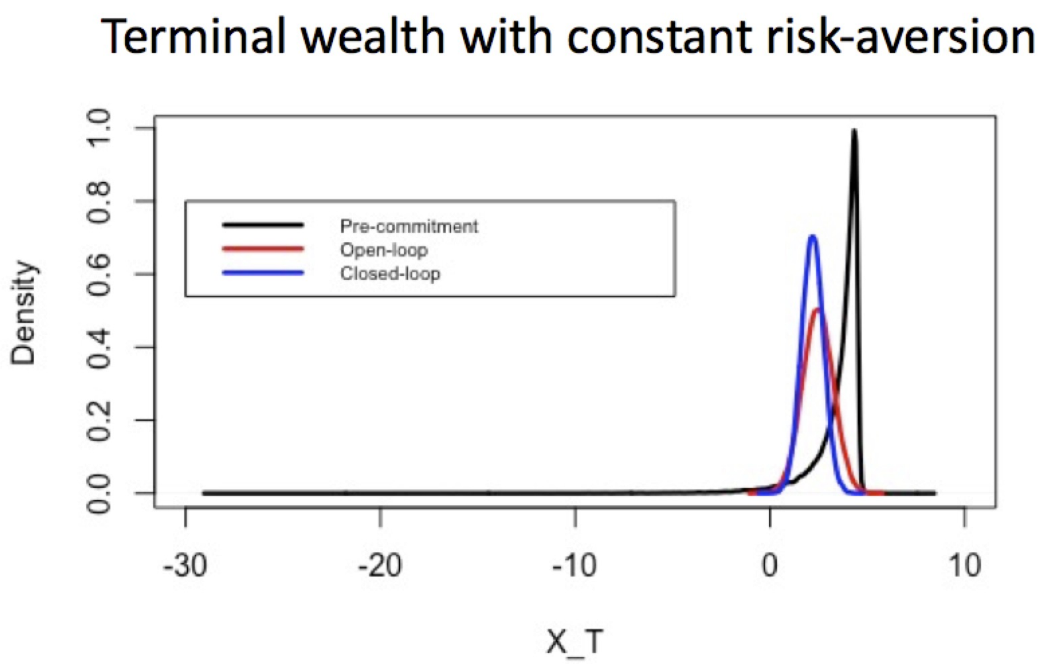


Figure 7.1: Probability density function for 10-year investment with 65,000 Monte Carlo simulations. The parameters are given in Table 7.1 with a constant risk-aversion parameter $\gamma = 0.5$.

The figures 7.2 shows the mean and standard deviation for the terminal wealth for various length of investment period. It can be seen that both open-loop and closed-loop equilibrium strategies have a low terminal return variance, whereas pre-commitment starts to out-perform the equilibrium strategies for longer investment period. For example, in table 7.2, although the Sharpe ratio of equilibrium strategy is higher than the pre-commitment strategy, the pre-commitment strategy achieves a high return with a reasonable level of variance. Compared with pre-commitment, the equilibrium strategy is a state-independent strategy in which the investors only invest a small amount of money during the whole investment period. As a result, the small amount of investment leads to a low standard deviation of terminal wealth.

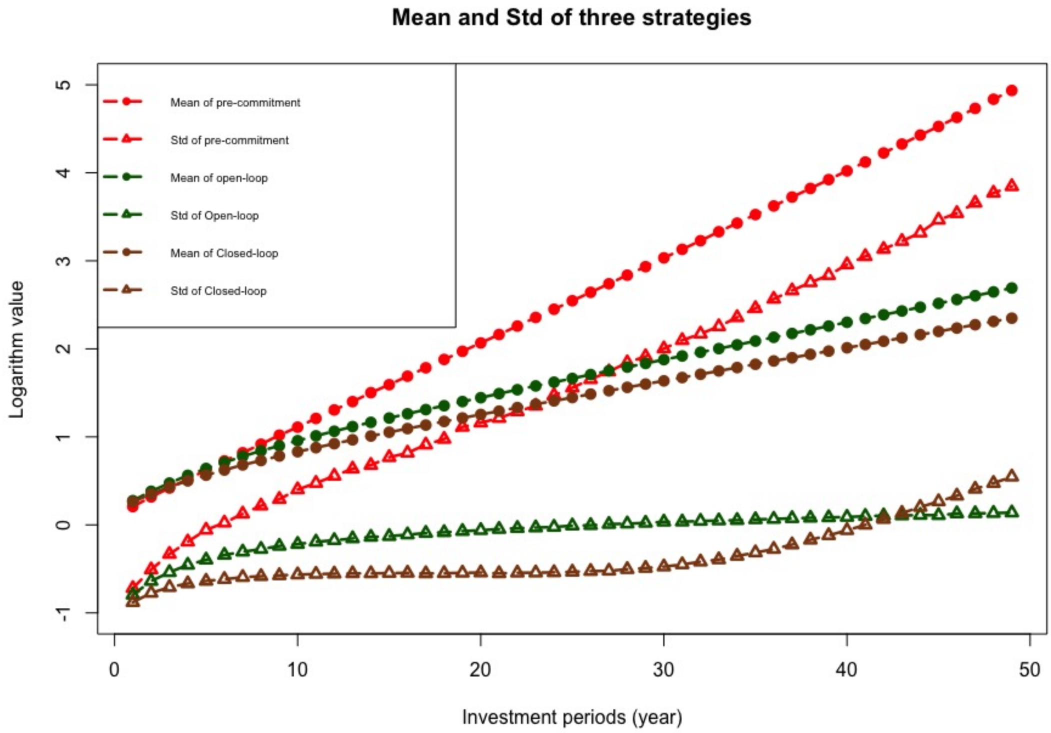


Figure 7.2: Mean and standard deviation for three different strategies. There are 65000 Monte Carlo simulations. The parameters are given in Table 7.1.

.

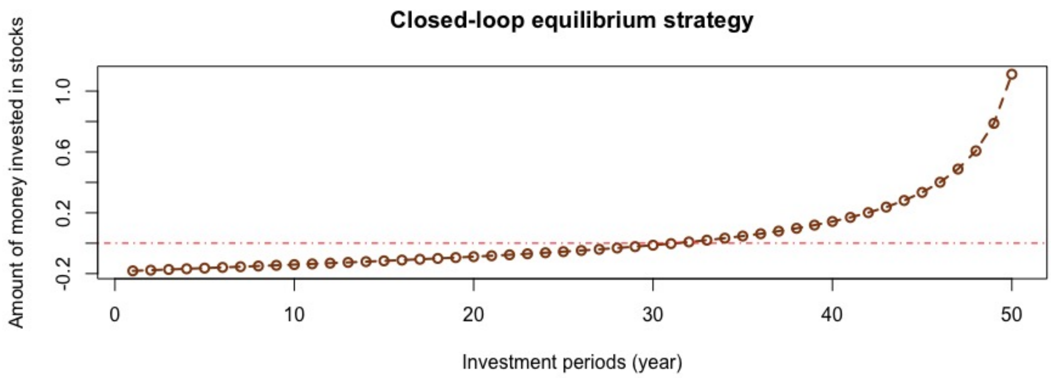


Figure 7.3: The closed-loop strategy with the parameters given in Table 7.1 for investment periods between 1 and 50 years.

.

Figure 7.2 also provides a comparison for two types of equilibrium strategies. First of all, we observe the standard deviation of the closed-loop equilibrium strategies increase much faster after 31 years. This phenomenon is caused by the negative initial investment of closed-loop equilibrium strategies. As shown in Figure 7.3, the closed-loop equilibrium strategies approach to negative value

40 yrs	Standard deviation	Mean
Pre-commitment	7.016938	74.04783
Open-loop	1.0879517	9.568016
Closed-loop	0.8855937	7.198412

Table 7.2: Means and standard deviations of terminal wealth for 40-years investment under different strategies

around year 31. Next, to compare the performance between two equilibrium strategy, we analyse the ratio of Sharpe ratios, which is defined by:

$$\text{ratio of sharpe ratios} = \frac{\text{Expected return}^{open}/\text{Expected return}^{closed}}{\text{Standard deviation}^{open}/\text{Standard deviation}^{closed}}.$$

If the ratio is above 1, then it describes that the open-loop equilibrium strategy achieves a higher return than closed-loop with respect to the same level of risk. For example, in table 7.2, the open-loop equilibrium policy yields an expected return of $\frac{9.568016}{1.0879517} = 8.794522$, while the closed-loop equilibrium policy only gives $\frac{10.467957}{1.725999} = 8.128346$. The ratio of Sharpe ratios is equal to 1.081957. Therefore, we conclude that the open-loop equilibrium strategy is preferable to closed-loop strategy for 40-years investment period.

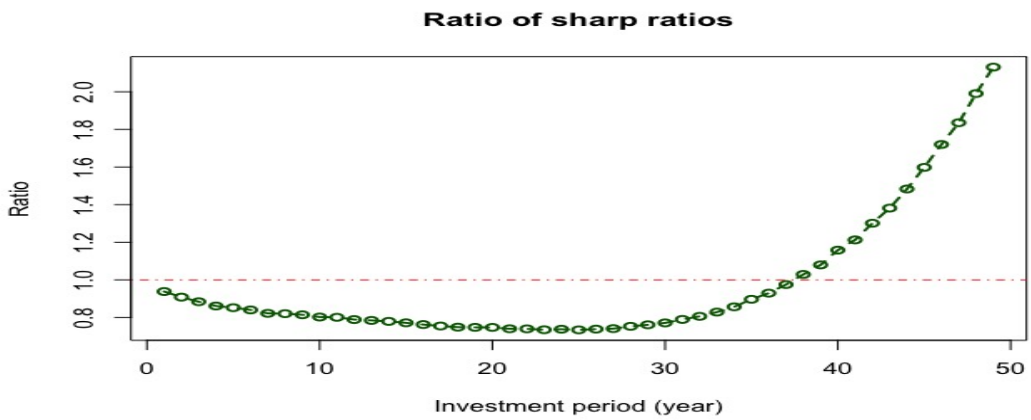


Figure 7.4: The ratio of terminal wealth Sharpe ratios for equilibrium strategies.

Figure 7.4 shows the ratio defined as above for various investment periods. We can see that, before 37-years investment time, the closed-loop equilibrium performs better than the open-loop equilibrium. After 37-years investment time, the open-loop equilibrium becomes better than the closed-loop equilibrium.

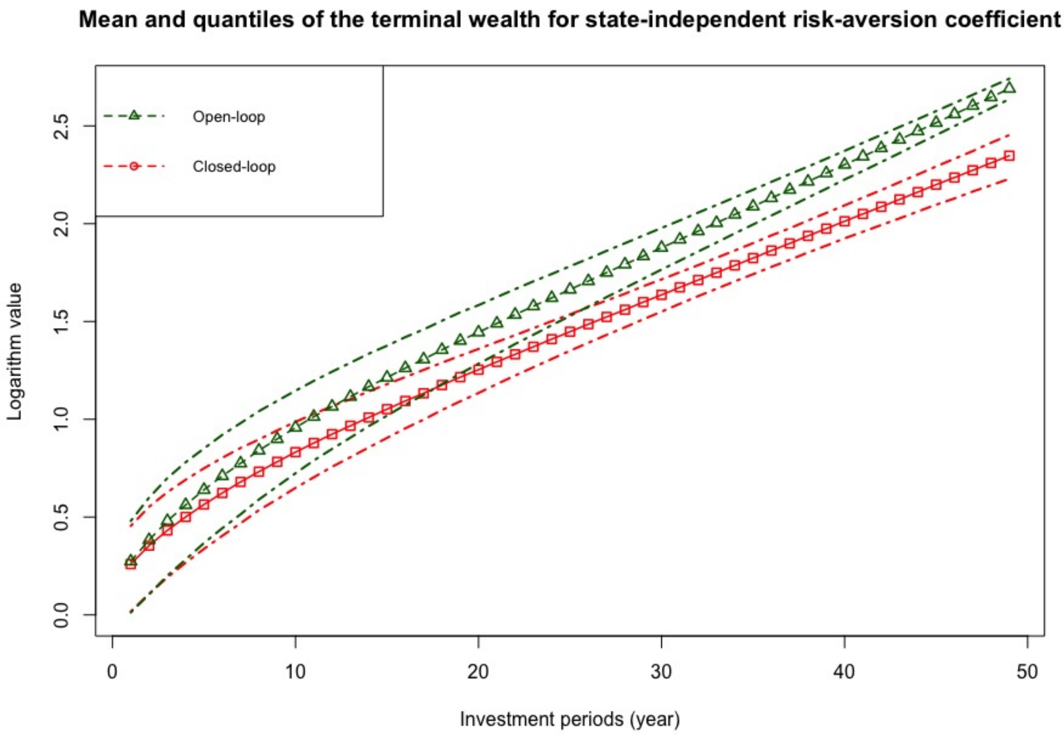


Figure 7.5: The ratio of terminal wealth Sharpe ratios for equilibrium strategies

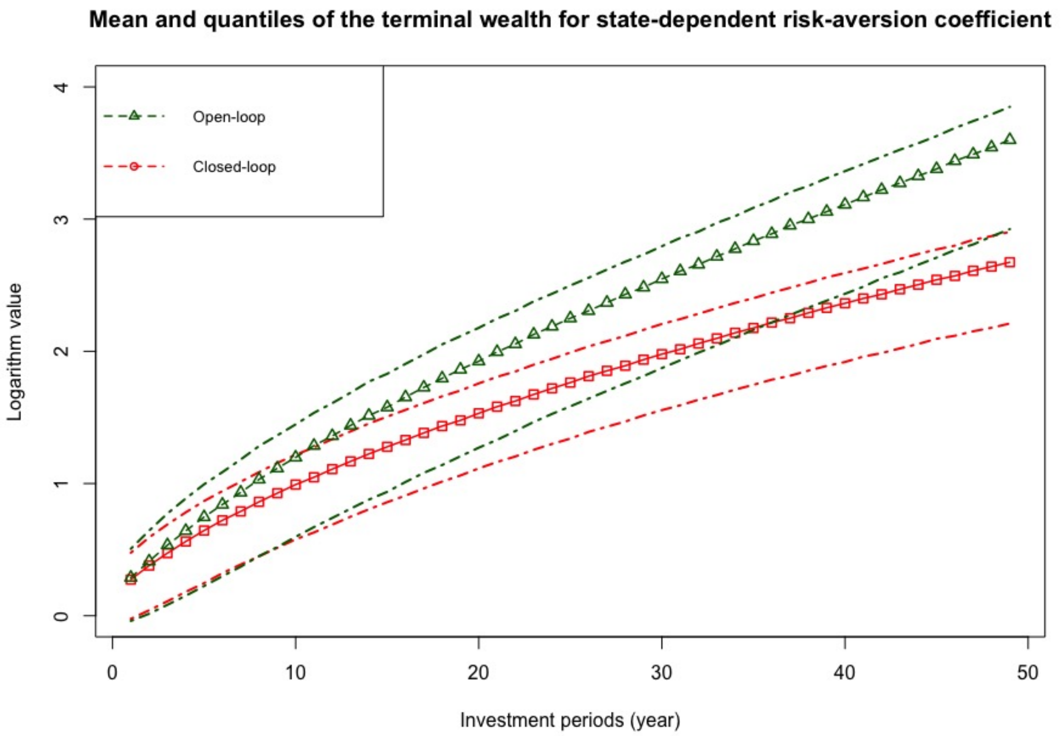
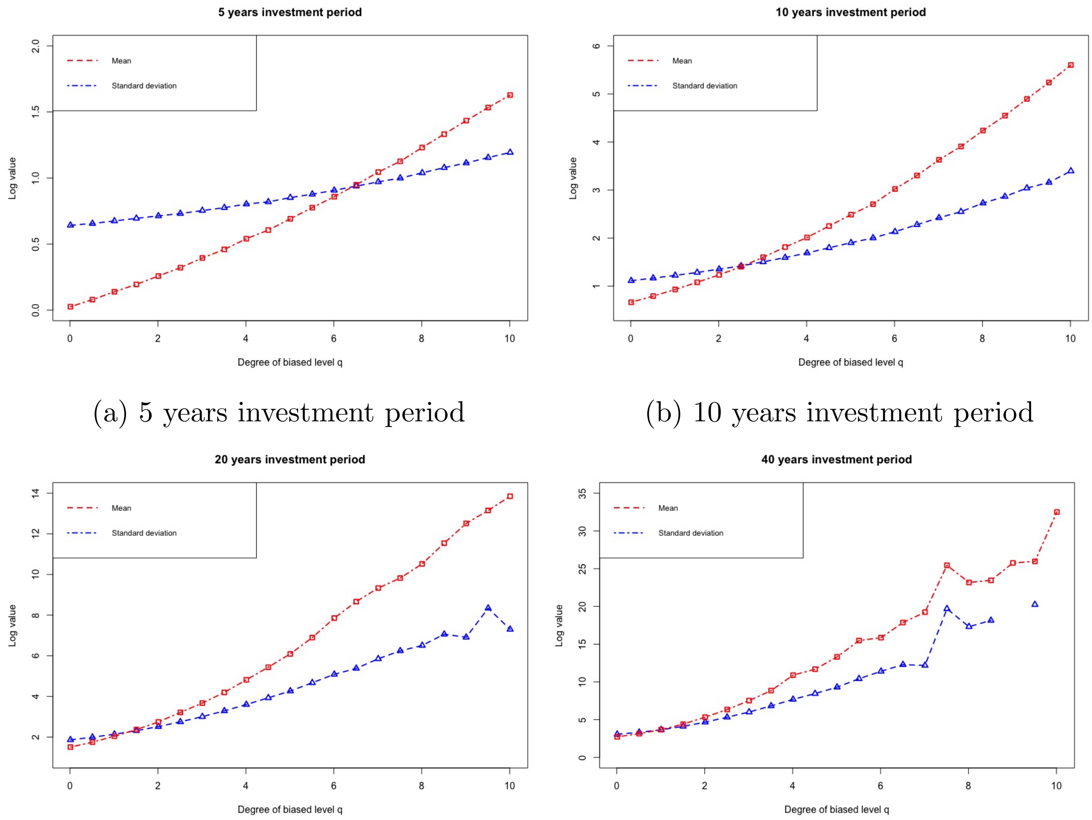


Figure 7.6: The ratio of terminal wealth's Sharpe ratios for equilibrium strategies

Figure 7.5 - Figure 7.6 show the mean path of $65,000$ Monte Carlo simulated trajectories for both open-loop equilibrium and closed-loop equilibrium strategies with state-dependent risk aversion $\frac{1}{2x_t}$ and state-independent risk aversion $\frac{1}{2}$, respectively. First, we notice that the state-dependent risk-aversion coefficient yields an increase in the mean of open-loop terminal wealth. In contrast, the mean of closed-loop equilibrium is almost identical to the state-independent case. This observation shows that the improvement of closed-loop equilibrium investment may depend on the initial wealth level. Secondly, the open-loop equilibrium wealth has a wide lower quantile and a narrow upper quantile in the state-dependent case. In addition, the distance between mean and quantiles does not change along with an increase in investment period, whereas the quantiles shrink towards the mean in the state-independent case. Finally, as shown in figure 7.6, the lower quantile of open-loop terminal wealth increases and crosses the upper quantile of closed-loop terminal wealth as investment period increases, which means three quarter of the open-loop terminal wealth is greater than three quarter of closed-loop equilibrium terminal wealth. Therefore, the open-loop equilibrium strategy is more likely to yield a higher terminal wealth than the closed-loop equilibrium strategy.

Figure 7.7 shows the mean and standard deviation of open-loop terminal wealth with present-biased risk-aversion coefficient for various investment periods. As described in Chapter 6.2, the present-biased risk-aversion has the form: $\frac{e^{-qr(T-t-1)}}{2x_t}$. The value of biased level q is tested between 0 and 10 with a step size of 0.5. It can be noticed that the present-biased risk-aversion leads to a deterioration in the performance of the wealth's Sharpe ratio. As biased level increases, the Sharpe ratios decreases. This is because an increase in biased level causes the investors to invest in the stock aggressively from the beginning of the investment period. The variance of terminal wealth grows quadratically corresponding to this adjusted investment, while the mean of terminal wealth grows linearly. Furthermore, we notice that the crossing point between red and blue lines shifts towards left along with the increase in investment period. Therefore-for long term investment, using the present-biased risk coefficient will result a drop in the Sharpe ratio.



(a) 5 years investment period

(b) 10 years investment period

(c) 20 years investment period

(d) 40 years investment period

Figure 7.7: Mean and standard deviation with present-biased risk-aversion parameter for different investment periods.

7.2 Pre-commitment v.s. Equilibrium: Case study

In the previous section, we study the performance of different types of strategies by using the Monte Carlo simulations. In this section, we carry out numerical tests on comparing the strategies for Apple Inc stock shares. We will test our comparison in the following ways:

- **Case Study A** Suppose the investor enters the market in Jan, 2000 with all the information from 2000 to 2017, i.e., the stock price between 2000 and 2017 is known. Then we compare the pre-commitment strategy and equilibrium strategy with monthly rebalancing for three different time lengths: 5 years, 10 years and 15 years.

- **Case Study B** Suppose the investor enter the market in Jan, 2000 with the parameters estimated from past 5 years $(1995 - 1999)$ and rebalances the strategies monthly. Then the investor aims to invest for 15 years and re-evaluate his strategies for every 1, 3 and 5 years.

The primary goal in this chapter is to compare the different investment behaviours and investment outcomes of equilibrium strategies. We classify the investment situations into two cases: first, we suppose that the investors evaluate their strategies at the beginning of the investment period and the investors do not intend to change the strategies afterwards. Such situation is analysed in the *Case Study A*. Since our study focuses on the comparison of the performance between different strategies, without considering the estimated parameter sensitivity for different strategies, we would like to assume the investors have all useful information about future stock market. Therefore, the investors are assumed to observe all relevant information of future share prices.

In contrast, *Case Study B* illustrates a different investment behaviour by allowing regular re-evaluation of the strategies. Since the pre-commitment has been well-studied and widely used under this circumstance, the *Case study B* explores the performance of equilibrium strategies compared to the pre-commitment. Furthermore, as the estimated parameters change along with re-evaluation time, we can analyse the parameter sensitivity of equilibrium strategies. For the sake of simplicity, we assume the independence for the future asset returns and the investors enter the market with 1000 dollars.

7.2.1 Case Study A

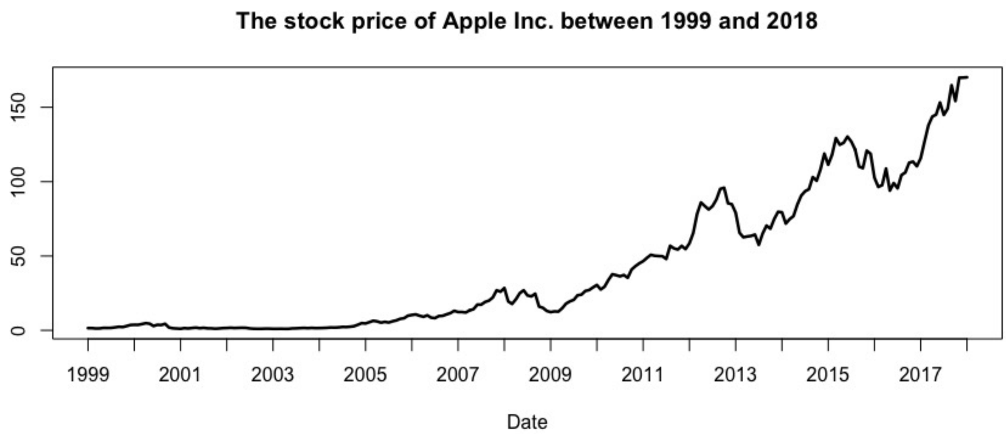


Figure 7.8: Apple Inc. share price from Jan,1999 to Jan,2018

The Figure 7.8 shows the stock price of Apple. We observe the uptrend and downtrend in the price movement. The 5 years investment parameters are estimated for 60 months from 2000 to 2005. The mean and standard deviation of monthly excess return rate estimated for 5 years investment period are 0.01673431 and 0.1688279, respectively. The 95% confidence interval for estimated 5-years excess return is $(-0.007146538, 0.04061516)$. The 10 years investment parameters are estimated for 120 months from 2000 to 2010. The mean and standard deviation of monthly excess return rate estimated for 10 years investment period are 0.02658799 and 0.1466011, respectively. The 95% confidence interval for estimated 10-years excess return is $(0.9787574, 1.019367)$. The 15 years investment parameters are estimated from 180 months from 2000 to 2015. The mean and standard deviation of monthly excess return rate estimated for 15 years investment period are 0.02568134 and 0.126899, respectively. The 95% confidence interval for estimated 15-years excess return is $(0.007731372, 0.0436313)$.

Based on the parameter listed above, we will show the resulting strategies for three different types of risk-aversion parameters: constant risk-aversion parameter $(\frac{1}{2})$, state-dependent risk-aversion parameter $(\frac{1}{2x_t})$, and present-biased risk-aversion parameter $(\frac{e^{-qr(T-t-1)}}{2x_t})$. The aims of studying these risk-aversion parameters are different. In state-independent risk-aversion parameter case, we focus on studying the difference in structure between open-loop and closed-loop equilibrium strategies with different investment periods. In the state-dependent risk-aversion parameter case, we focus on assessing different investment performance between pre-commitment, open-loop equilibrium and closed-loop equilibrium strategies. In the present-biased risk-aversion parameter case, we focus on improving the investment outcome for open-loop equilibrium strategy with different biased level q.

State-independent Risk-aversion Coefficient Case

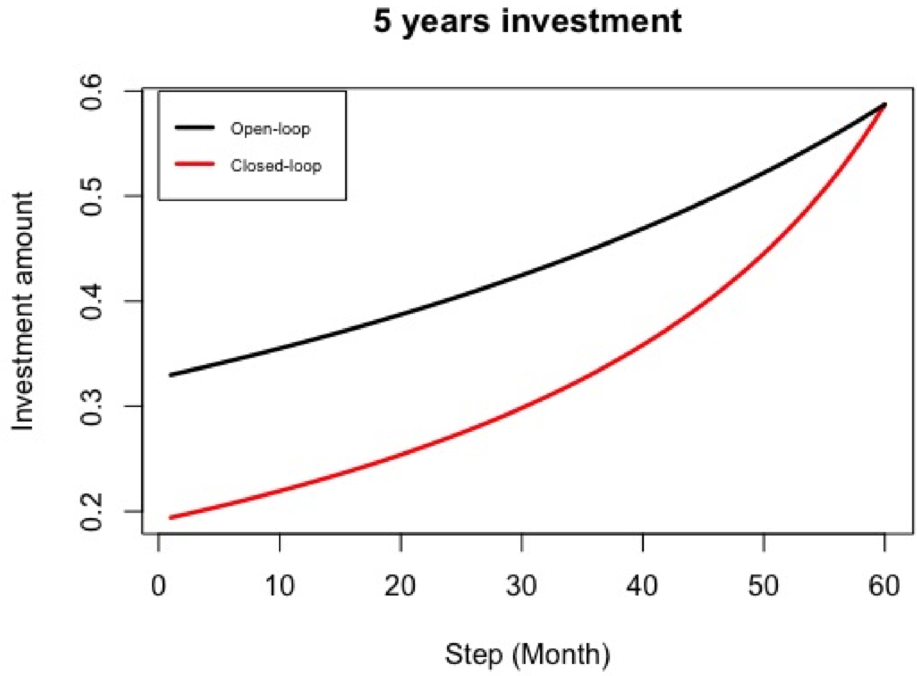


Figure 7.9: The equilibrium strategies for 5 years investment period. The parameters are estimated from 2000 to 2005.

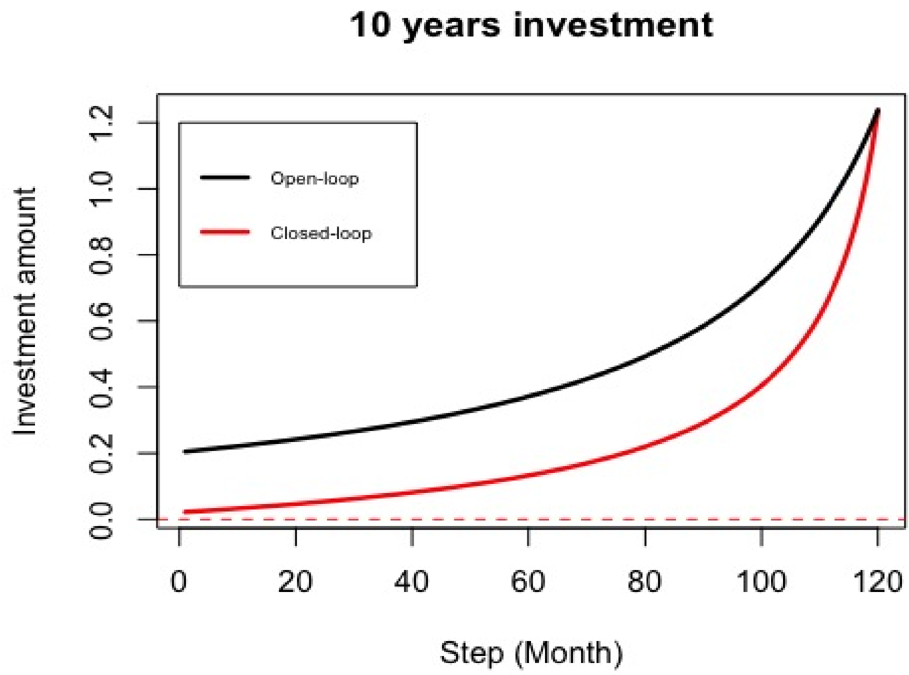


Figure 7.10: The equilibrium strategies for 15 years investment period. The parameters are estimated from 2000 to 2010.

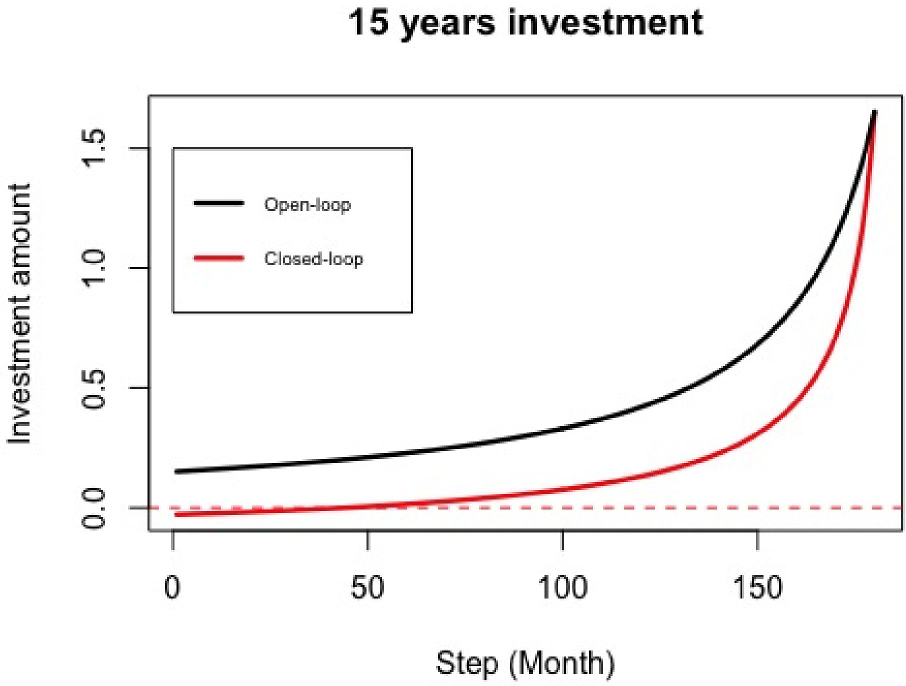


Figure 7.11: The equilibrium strategies for 5 years investment period. The parameters are estimated from 2000 to 2015.

Figures 7.9 - 7.11 show the amount of money invested in stock for different equilibrium strategies. First, we observe that the open-loop equilibrium strategies are always above the closed-loop equilibrium strategies. This is because the short-term plan of open-loop equilibrium strategies only invests in the stock market for one period. Compared with the closed-loop equilibrium investor, the open-loop equilibrium investors take less future risk into account in the short-term investment plan. Then the open-loop equilibrium strategies are more risky than the closed-loop equilibrium strategies. Second, the open-loop and closed-loop equilibrium strategies approach to the same value at the end of investment period. This value is the mean-variance strategy for single-period framework, which depends on the estimated values of excess return and variance. Finally, as we mentioned in Chapter 4, for long period (e.g. 15 years), the initial closed-loop equilibrium strategy becomes negative.

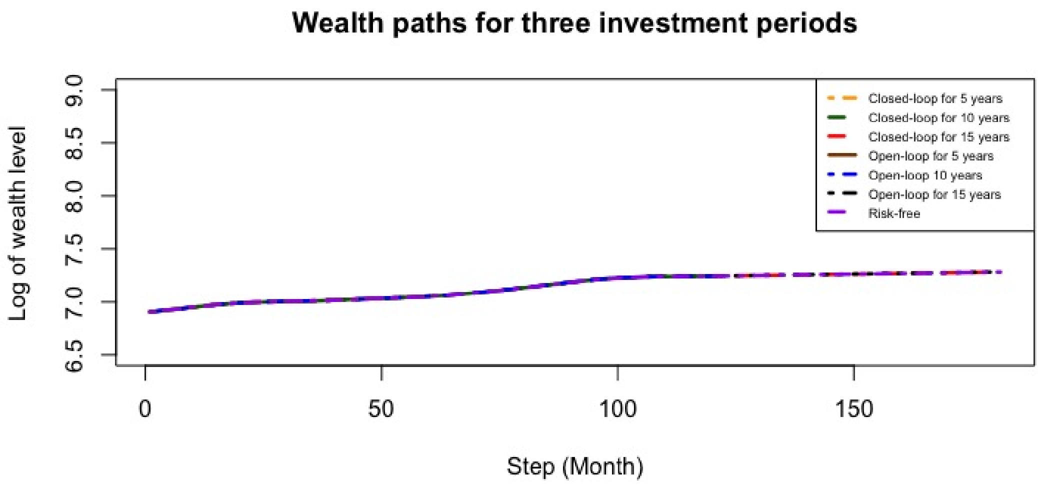


Figure 7.12: The logarithm of wealth paths under different investment strategies for various investment periods.

Figure 7.12 shows the logarithm of wealth paths for different investment periods. First of all, we notice that there is no differences in terminal wealth between open-loop and closed-loop equilibrium strategies for 5 years, 10 years and 15 years. Also, the equilibrium wealth paths are identical to the risk-free wealth path. This shows the equilibrium wealth is driven by the risk-free asset return rather than the stock investment return.

State-dependent Risk-aversion Coefficient Case

It can be criticised that the investment amount is fixed for any amount of initial wealth. As we described above, the state-independent wealth path is driven by the risk-free asset return. The state-dependent risk-aversion coefficient solves such problem. Considering the state-dependent risk-aversion coefficient, the resulting figures for equilibrium strategies are identical to the Figures 7.9 - 7.11. However, instead of the amount of money invested in the stock market, the y-axis indicates the proportion of the current wealth investing in the stocks.

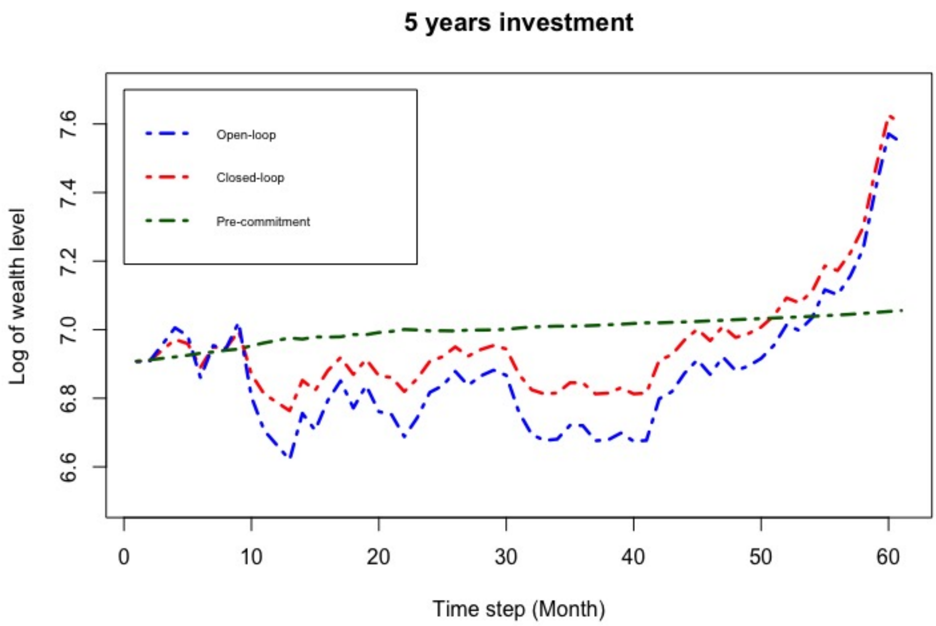


Figure 7.13: The logarithm of wealth paths by pre-commitment, open-loop and closed-loop equilibrium strategies for 5 years investment period.

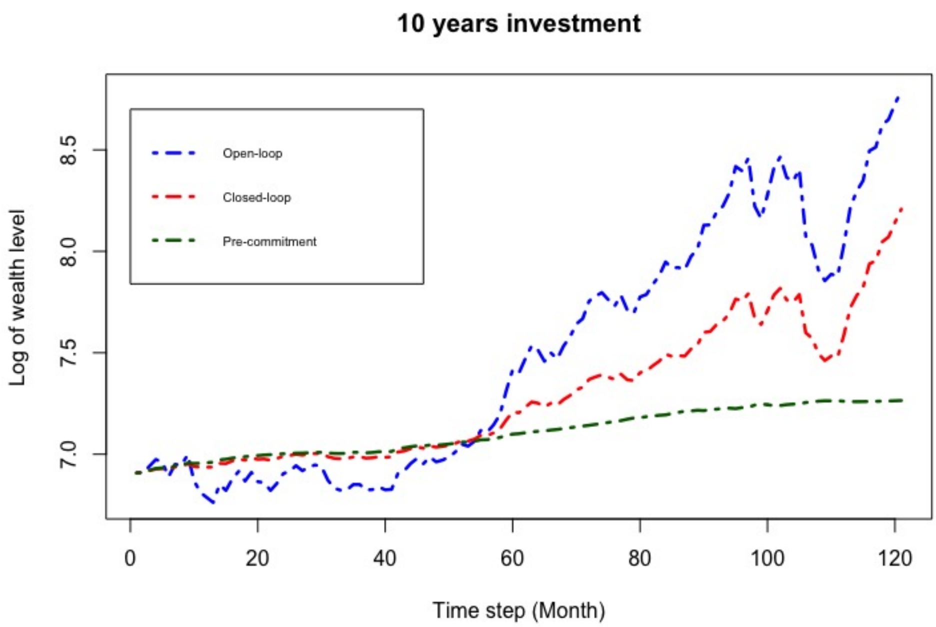


Figure 7.14: The logarithm of wealth paths by pre-commitment, open-loop and closed-loop equilibrium strategies for 10 years investment period.

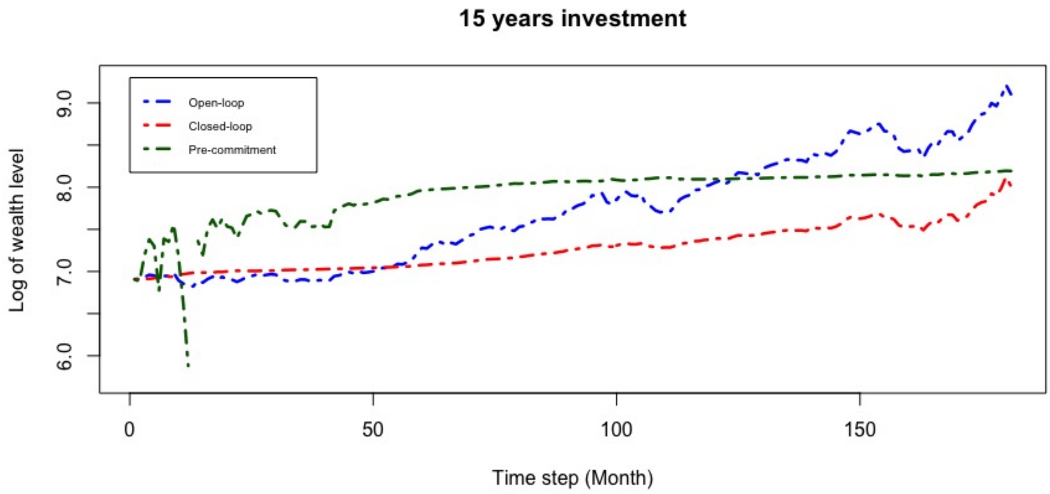


Figure 7.15: The logarithm of wealth paths by pre-commitment, open-loop and closed-loop equilibrium strategies for 15 years investment period.

Figures 7.13 - 7.15 show the logarithm of wealth paths for three types of strategies: pre-commitment, open-loop equilibrium and closed-loop equilibrium strategies. Fist of all, we notice that, in figure 7.15, the wealth under pre-commitment strategies becomes negative at 11th month. This is because the pre-commiment investors try to achieve the target investment and invest a dramatical amount of money into the stock market (see Figure 7.16). For illustration, the investment proportion for pre-commitment strategy has a range between -40 and 10. Once the target is achieved, the pre-commitment investors will stabilise their wealth around the target by reducing the amount of investment. Such behaviour can be observed in Figure 7.16. In contrast, the equilibrium investors always start with a small proportion and become risk-seeking at the maturity. Secondly, compared with closed-loop equilibrium strategies, open-loop equilibrium strategies yield a better terminal wealth in general. Although Figure 7.13 shows that the closed-loop equilibrium strategy gives higher terminal wealth, the open-loop achieves a better performance between month 12 and month 60. Finally, the equilibrium strategies yield better terminal wealth than the pre-commitment strategies. This is because the pre-commitment strategy optimises the objective function at the initial time. The pre-commitment investors set their target investment based on the initial wealth level, whereas there is no targeting threshold for equilibrium investors. As the wealth level changes dramatically over a long period, the targeting wealth evaluating from pre-commitment is undervalued at initial time.

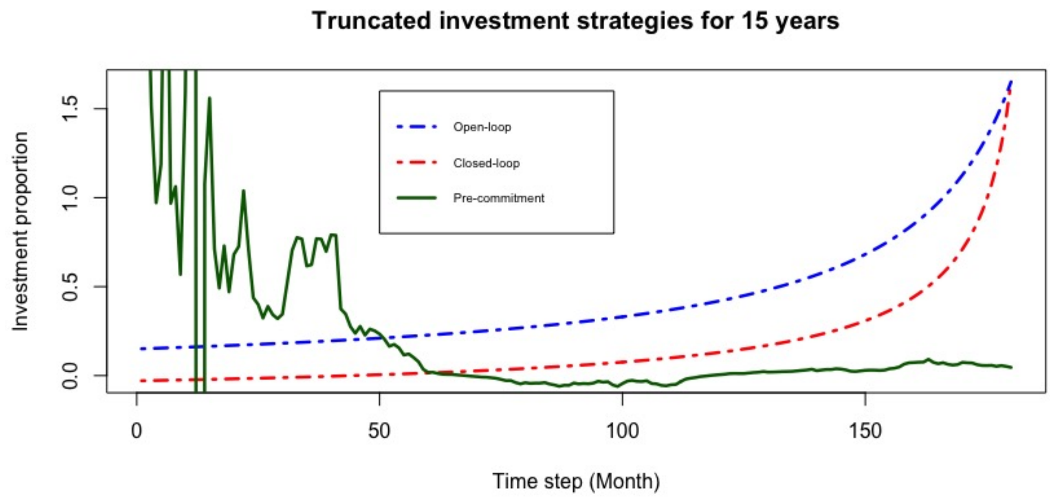


Figure 7.16: The truncated investment proportions of pre-commitment, closed-loop and open-loop equilibrium strategies.

Present-biased risk-aversion Coefficient Case

We study the present-biased risk-aversion coefficient with the range of biased level q between 0.01 and 10. As a result of applying present-biased risk aversion, the sizes of the open-loop equilibrium strategies (the absolute value of the investment proportion) increase at the beginning of investment period. Figures 7.17 shows the terminal wealth for three different investment periods. We notice the present-biased risk-aversion parameter increases the terminal investment wealth for 10 and 15 years investment. In contrast, the terminal wealth strictly decreases as the biased level increases for 5 years investment period. It is interesting to explore whether such observation holds for general. Therefore, we use the share prices of Sony Corporation to compare the result between state-dependent and present-biased risk-aversion coefficients.

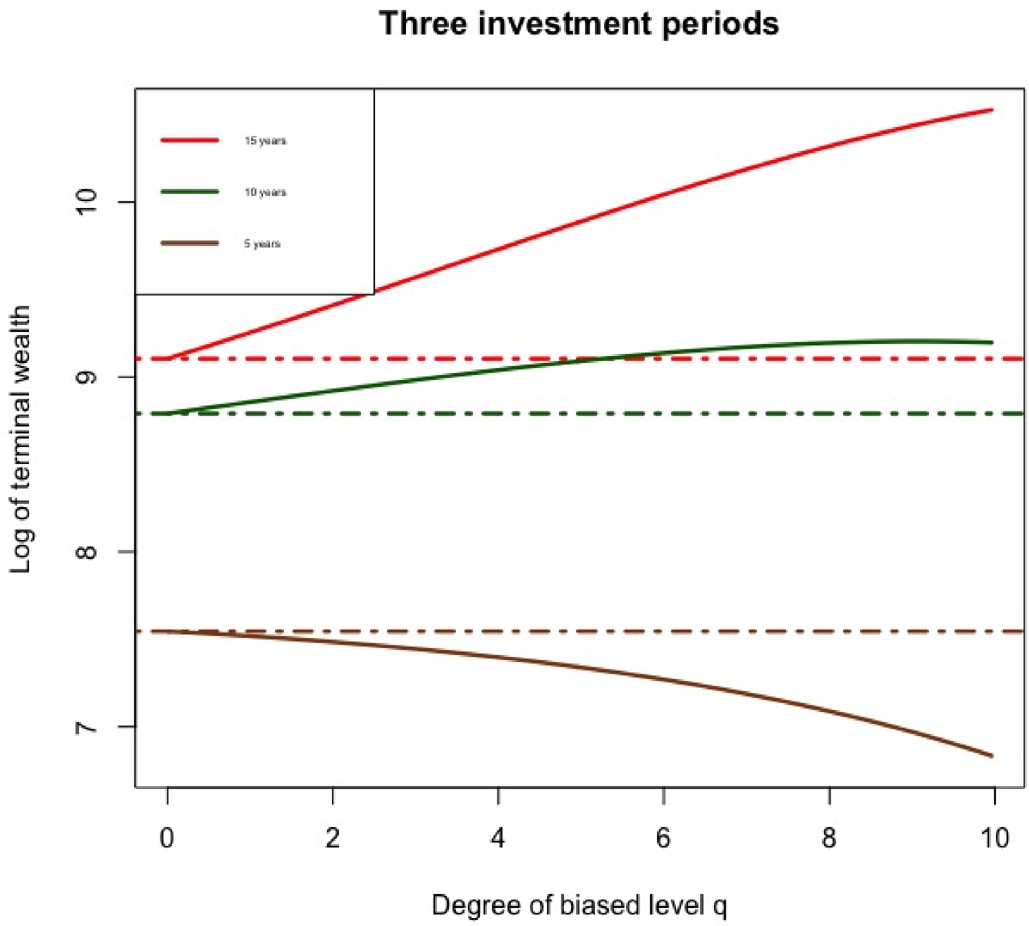


Figure 7.17: The logarithm terminal wealth of present-biased risk-aversion coefficients with different biased level q for Apple inc. stock. The investment periods are 5, 10 and 15 years. The horizontal dash lines are the terminal threshold with state-dependent risk-aversion coefficients.

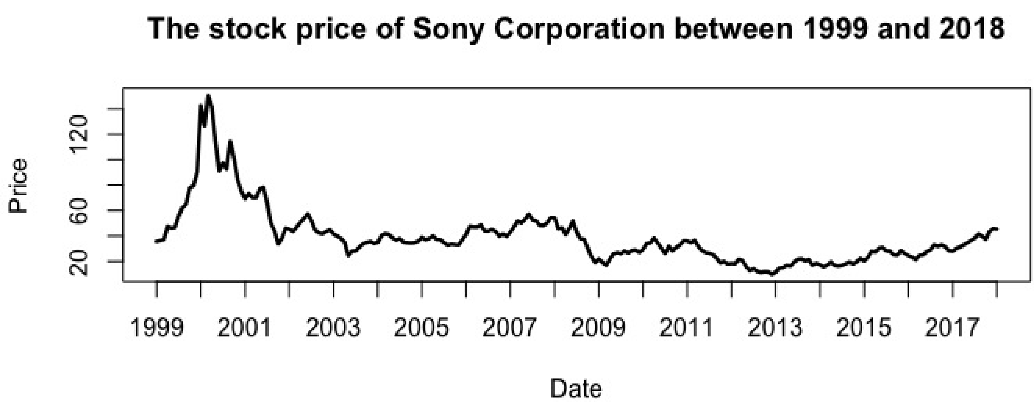


Figure 7.18: Sony Corporation share price from Jan, 1999 to Jan, 2018.

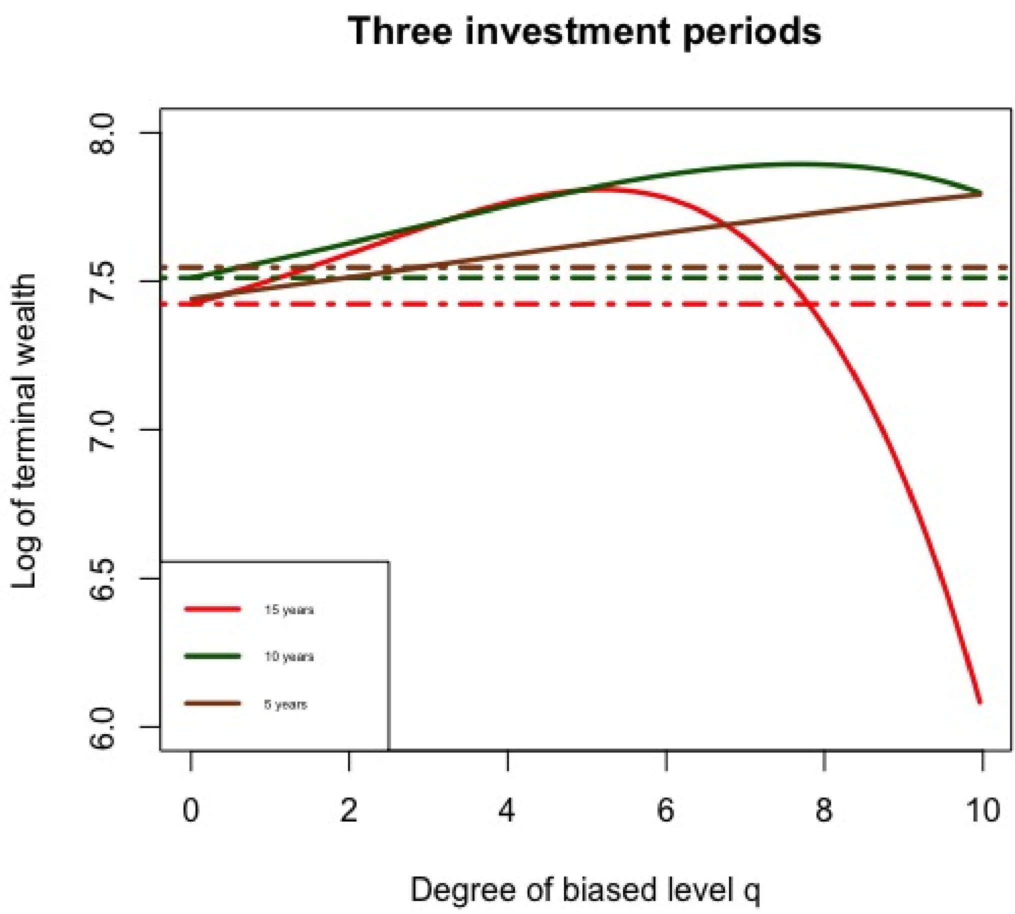


Figure 7.19: The logarithm terminal wealth of present-biased risk-aversion coefficients with different biased level q for Sony corporation stock. The investment periods are 5, 10 and 15 years. The horizontal dash lines are the terminal threshold with state-dependent risk-aversion coefficients.

Figure 7.18 shows the share prices of Sony corporation between 1999 and 2018. It can be noticed that, in the first two years, there is a recession in the stock market. As a result, the estimated means of excess return are all negative and open-loop strategies become all negative (short-selling). The higher biased level q yields a larger investment in the risk-free asset. Since the price of Sony shares is stabilised after first few years, it is not profitable to keep investing a large proportion of wealth in the risk-free asset for long term. For illustration, in figure 7.19, the red curve decays with a higher biased level. Compared with state-dependent risk-aversion parameter, the present-biased risk-aversion parameters is able to achieve an improvement in terminal wealth. However, it is crucial to determine the biased level q based on the anticipated future.

7.2.2 Case Study B

In previous subsection, we address the issue with "Case Study A": without adjusting the targeting threshold for pre-commitment strategies, the equilibrium strategies out-perform the pre-commitment in the long-term investment. Therefore, in "Case Study B", we would like to further explore the situation in which the adjustment is allowed. While the adjustment is allowed, the input parameters are estimated from past 5 years.

State-dependent Risk-aversion Coefficient Case

To illustrate the impact of frequent evaluation of the equilibrium strategies, Figure 7.20 -7.21 indicate that the re-evaluation does not help the equilibrium investors to improve the performance of the investment. The intuition is straightforward. Since the equilibrium strategies only becomes risky towards the maturity, apart from those strategies near the maturity, the changes of remaining equilibrium strategies are relatively small. Then the open-loop equilibrium strategies with three different updating frequency are relatively close to each other. Also, as long as the updating estimated parameters (mean of excess return, variance of excess return and the interest rate) do not change significantly, the resulting terminal strategies are almost identical near the maturity. For example, as shown in Figure 7.22, the terminal size of equilibrium strategy for the brown line estimated from 2010 is almost identical the green dash line estimated from 2012.

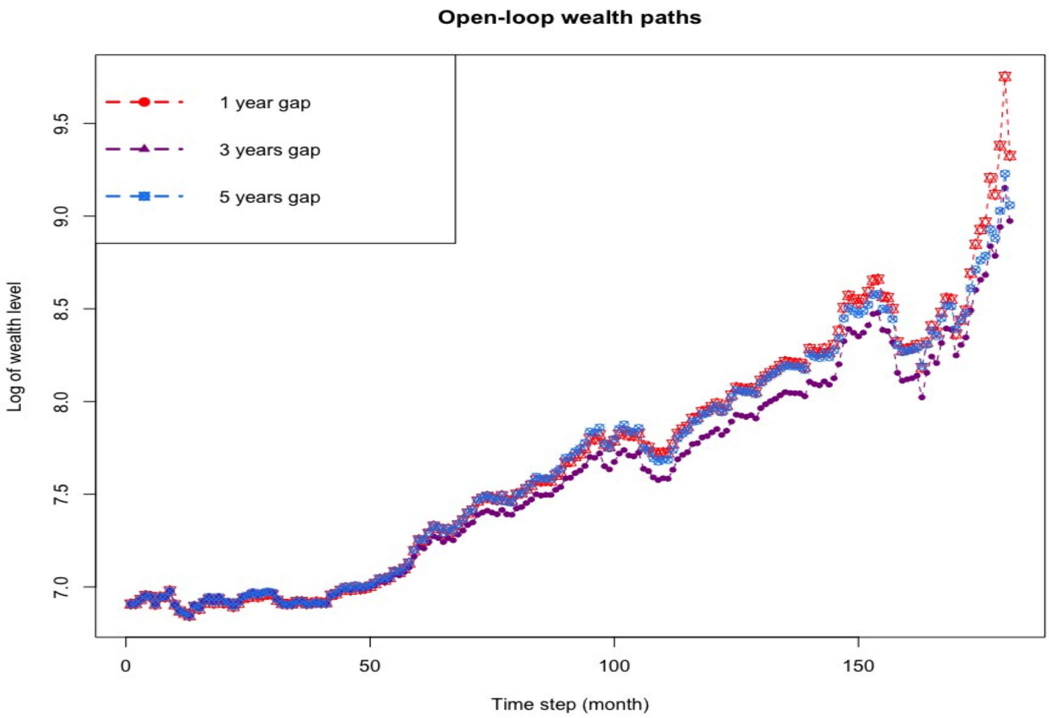


Figure 7.20: The logarithm of wealth paths by open-loop equilibrium strategies for 15 years investment period. The equilibrium strategies have been re-evaluated for every 1 year, 3 years and 5 years.

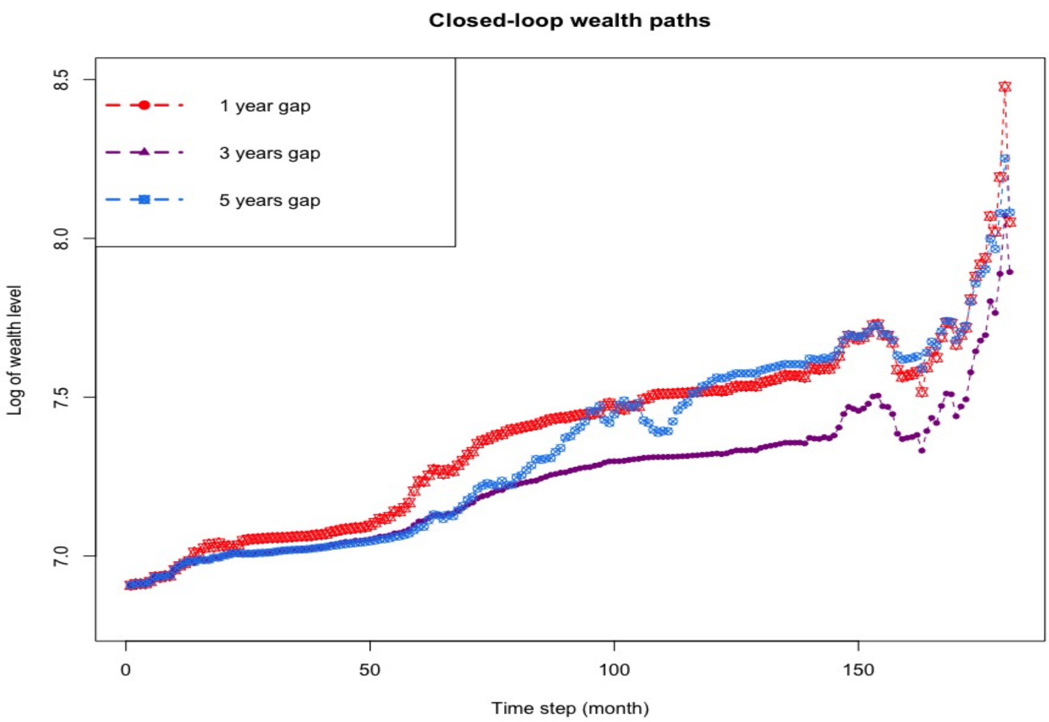


Figure 7.21: The logarithm of wealth paths by closed-loop equilibrium strategies for 15 years investment period. The equilibrium strategies have been re-evaluated for every 1 year, 3 years and 5 years.

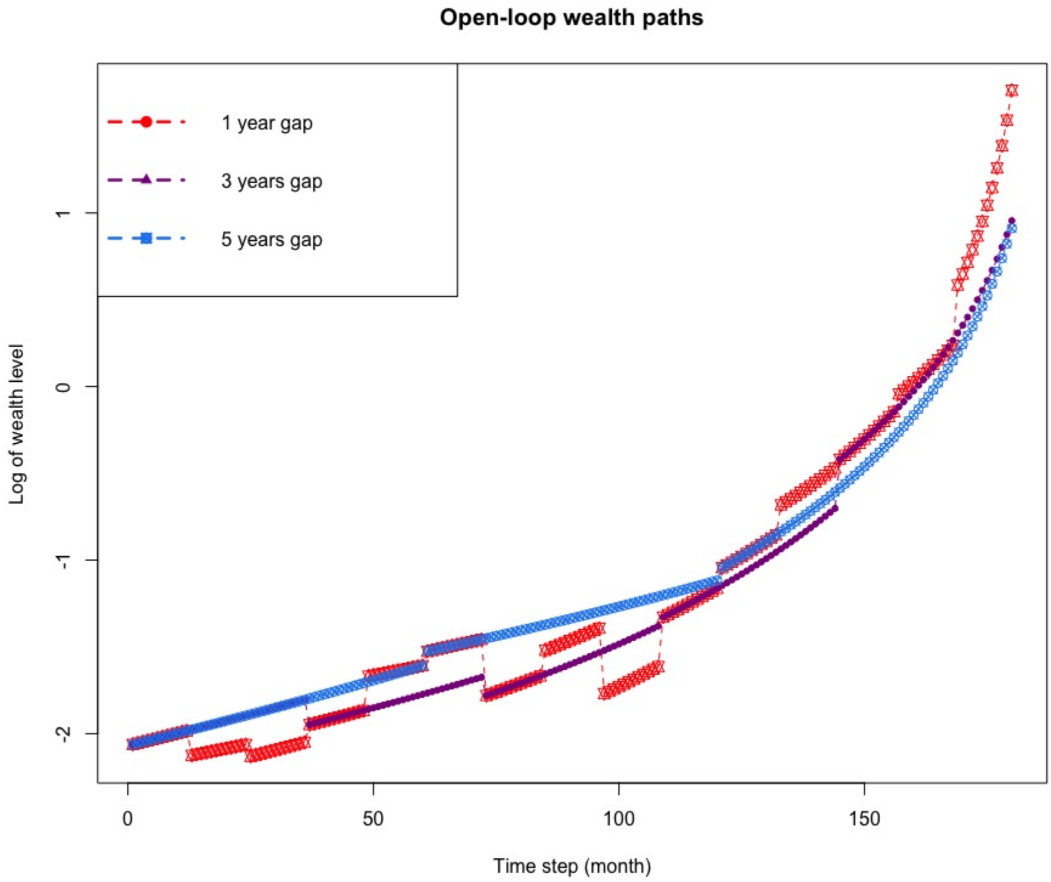


Figure 7.22: The open-loop equilibrium strategies for 15 years investment period by re-evaluating in every 1 year, 3 years and 5 years.

The figures 7.23-7.24 further illustrate the wealth paths and size of pre-commitment strategies with different re-evaluating frequency. Compare Figure 7.23 with 7.15, we notice that a high frequency of re-evaluation is able to improve the performance of investment under the pre-commitment strategies. This is because the targeting threshold is amended based on the current wealth and estimated parameters. In Figure 7.24, we observe a few peaks for green dash line around month 73. These peaks are caused by the updated mean and variance of the excess return in 2006. As the mean of excess return increases by more than twice, the pre-commitment investors immediately invest a huge amount of money into the risky asset. As a result, we notice logarithm wealth (the green dash line in Figure 7.23) jumps from 8 to 11. Compared with equilibrium strategies, the pre-commitment strategies are more sensitive to the estimated parameters.

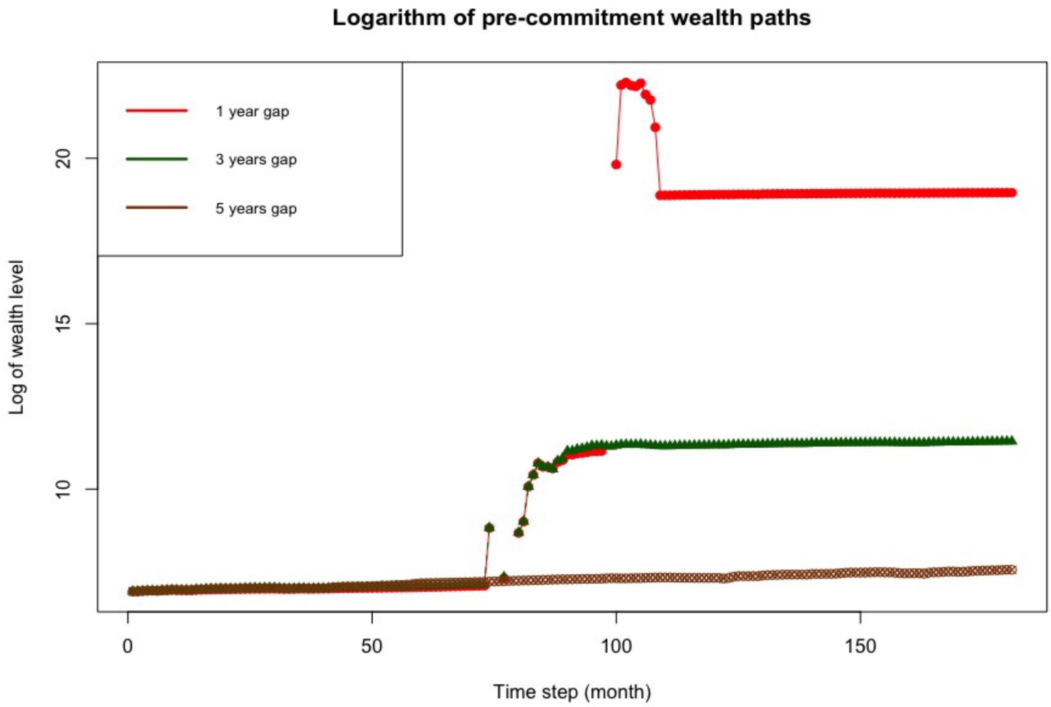


Figure 7.23: The logarithm of wealth paths by pre-commitment equilibrium strategies for 15 years investment period. The pre-commitment strategies have been re-evaluated for every 1 year, 3 years and 5 years.

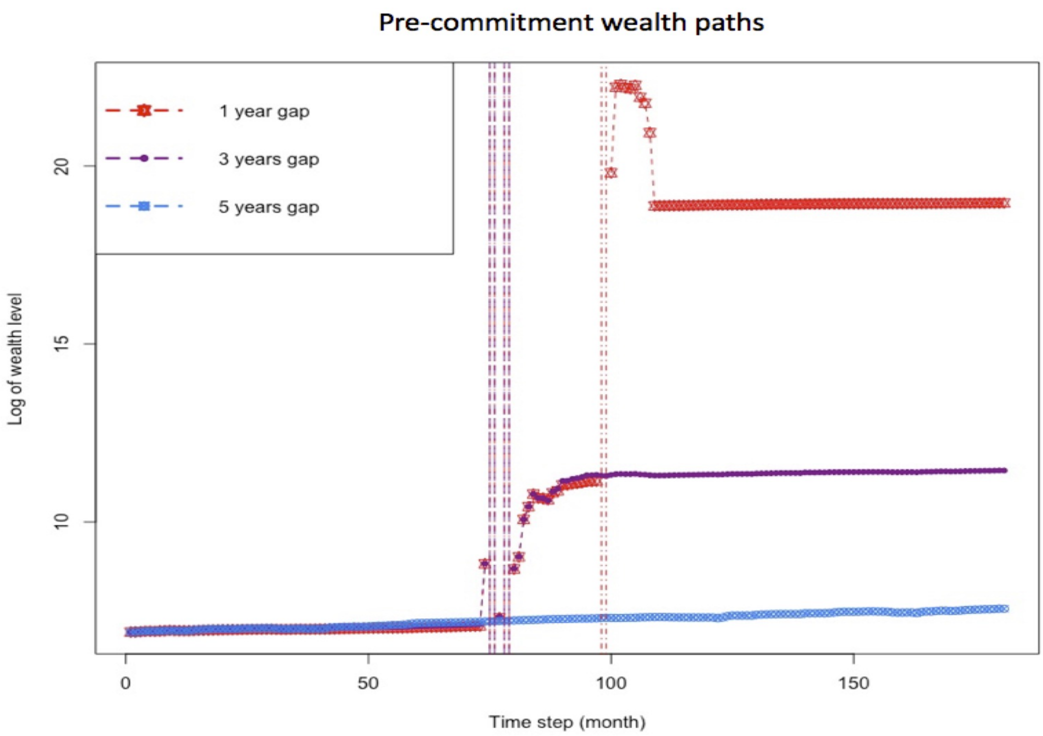


Figure 7.24: The pre-commitment strategies for 15 years investment period by re-evaluating in every 1 year, 3 years and 5 years. The vertical dash lines indicate the negative strategies (where logarithm is undefined).

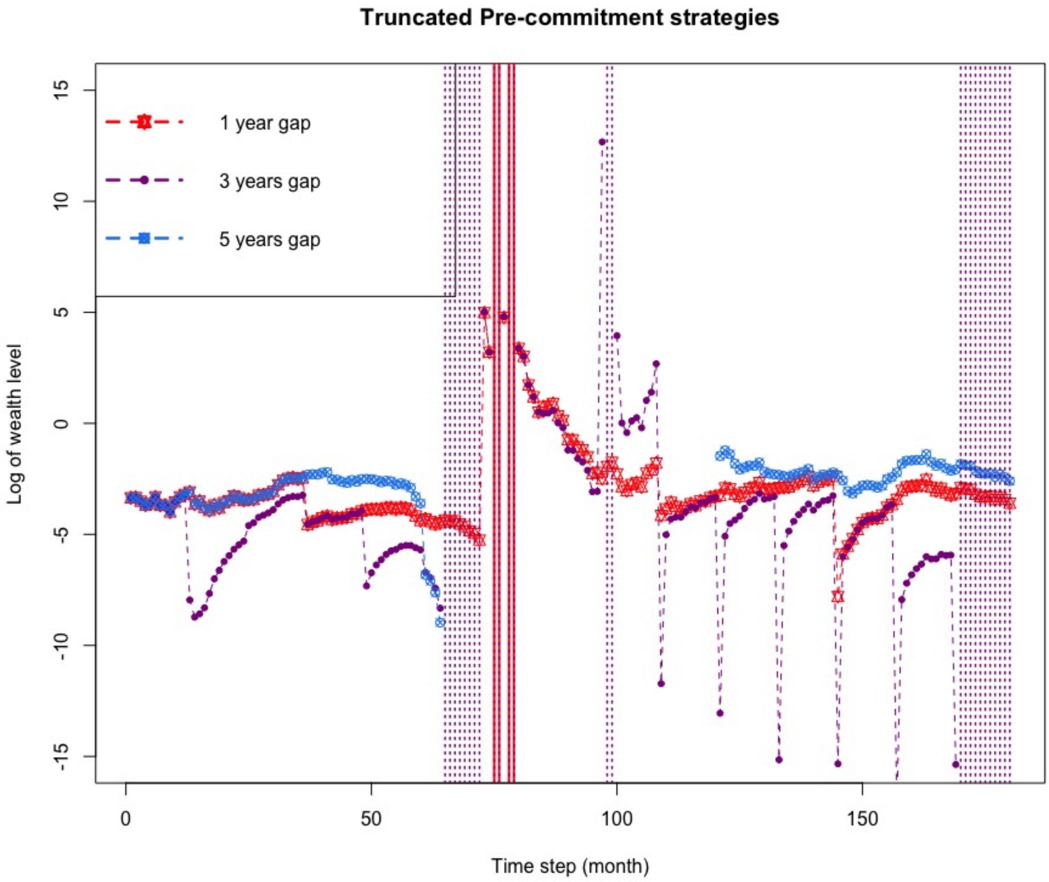


Figure 7.25: The truncated logarithm of pre-commitment strategies for 15 years investment period by re-evaluating in every 1 year, 3 years and 5 years. The vertical dash lines indicate the negative strategies. The strategies in the gap between the blue lines are all negative.

7.3 Conclusion

This chapter uses numerical tools to assess the performance of three types of strategies: pre-commitment, open-loop equilibrium and closed-loop equilibrium. One of the main focuses of the research has been on the Sharpe ratio and probability density of the terminal wealth under different strategies. For the market with independent and identical return, it has been shown that the equilibrium strategies with constant risk aversion achieve better Sharpe ratio of terminal wealth, whereas pre-commitment strategies are the profit-driven strategy which yields better expected return with higher risk.

The second focus of the research is on a case study to compare the performance of different strategies. The case study has been classified into two situations: re-evaluation prohibited and re-evaluation allowed. In the re-evaluation prohibited case, the equilibrium strategies out-performs the pre-commitment strategies. For pre-commitment strategies, once the target threshold is reached, the investors will not take anymore risk. Our results indicate that the pre-commitment investors may undervalue the target investment threshold due to the length of investment period and initial wealth level. In contrast, there is no target threshold for equilibrium investors. Comparing different equilibrium strategies, we find the open-loop equilibrium strategies out-performs the closed-loop equilibrium strategies as investment period increases, which is consistent with one observed in the Monte Carlo simulation case.

When the re-evaluation is allowed, the terminal pay-off of pre-commitment strategies increases along with the increase in the frequency of re-evaluation. The updated asset parameters and wealth level for pre-commitment strategies lead to a huge change in the trading volume. However, we show that there is no guarantee on the terminal pay-off for equilibrium investors as the re-evaluation frequency increases. This is because the updated asset parameters does not affect the equilibrium strategies dramatically.

Finally, we evaluate performance of the investment by using the present-biased risk-aversion coefficient. Similar to varying the risk attitude by wealth level, the present-biased risk aversion allows the investors to adjust their risk attitude based on the anticipated future investment environment. We show that the improvement can be achieved in both booming market and recession market. However, our results indicate that the investor's biased level is able to cause a more diverse wealth distribution.

Chapter 8

Conclusion

8.1 Summary

This research studies the different types of equilibrium strategies for mean-variance problem in discrete and continuous time. The focus is on analysing the role of perturbation in different equilibrium methodologies in affecting the dynamic of the wealth process. The goal is to explore the difference between open-loop and closed-loop equilibrium strategies existing in the current literature and identify the issues for implementing the equilibrium strategies in practice.

In Chapter 1, we describe the background and the motivation of the research. Also, we propose a number of research questions and an overview of the thesis.

Chapter 2 presents a literature review on the main three strategies existing in the literature namely: pre-commitment, open-loop equilibrium and closed-loop equilibrium. The purpose is to identify the outcomes, issues and gaps in this new implementation of equilibrium methodology for mean-variance problem. By reviewing the equilibrium strategies for mean-variance problem, we notice the different results for equilibrium strategies which may impact investors' financial decisions. Understanding the different rationale behind the equilibrium strategies helps the investors to choose the best strategy corresponding to their own preference.

Chapter 3 initiates the main research by studying the equilibrium strategies in discrete time. Although the continuous framework already provides an explicit solution for equilibrium strategy, there is still a gap between understanding the behaviour of equilibrium strategy and implementing the solution of the problem. In a discrete time setting, the behaviour of the strategy and the dynamic of the wealth are trackable by constructing the equilibrium strategy recursively. By incorporating concepts from open-loop controls in engineering, we take the perturbation affecting the wealth dynamic in a specific way and derive a necessary and sufficient condition for open-loop equilibrium strategy. A analytic study of existence and uniqueness of the open-loop equilibrium strategy has been carried out.

Chapter 4 shows a perturbation acts on the wealth in a different way in order to obtain the closed-loop equilibrium strategies. Together with results from Chapter 3, a key behavioural feature of equilibrium investors is that the investors divide their investment plan into two parts: short-term and long-term plans. Once the short-term plan is chosen, the investors combine the short-term plan with any long-term investment plan and test different amount of investment in short-term investment plan. Therefore, the equilibrium strategy is the strategy that yields the best outcome of the objective function among these combinations. This interpretation allows us to draw out the key different behaviour between open-loop and closed-loop investors. The open-loop short-term strategy (perturbation) only invests in the stock market for a single-period and deposits the investment into bank account until the maturity, whereas the closed-loop short-term strategy invests in the stock market and leaves the investment until the maturity.

Chapter 5 develops open-loop equilibrium strategies in continuous time by using a different type of perturbation. Compared with the random variable of perturbation used in Hu *et al.* (2012), we relax the perturbation as a random process. We have shown that the random process perturbation yields the same condition as the random variable perturbation. Therefore, the type of equilibrium strategies depends on the way that the perturbation affects the wealth dynamics. Compared with the continuous time framework, the discrete time framework provides more flexibility for the equilibrium investor in choosing the short-term investment plan (perturbation).

Chapter 6.2 addresses issues for equilibrium strategies for practical implementation. As time moves away from maturity, the investment amount for both closed-loop and open-loop strategies decays. The open-loop equilibrium decays to zero, whereas the closed-loop equilibrium decays to a negative limit. A proper relation between the value of the limit and the model parameters has been identified by the Sharpe ratio of the risky asset in Chapter 4. By incorporating concepts from behavioural economics, we focus on modifying the risk aversion parameter which represents the preference of the incarnations of investor at different time point.

Chapter 7 uses numerical tools to assess the performance of pre-commitment, open-loop and closed-loop equilibrium strategies in various situations. Our analysis shows that, the equilibrium strategies are preferable to pre-commitment strategies when the re-evaluation is prohibited. Comparing between the closed-loop and open-loop equilibrium strategies, we find that the performance of the closed-loop equilibrium deteriorates as the investment period increases. When the re-evaluation is allowed, the equilibrium strategies are not as sensitive as the pre-commitment strategies to the updated parameters of the assets. As a result, the pre-commitment strategies out-performs the equilibrium strategies as the re-evaluation frequency increases. Finally, we show that the improvement of the investment can be achieved by implementing the present-biased risk-aversion coefficient.

References

AIVALIOTIS, G. & PALCZEWSKI, J. (2014). Investment strategies and compensation of a mean–variance optimizing fund manager. *European journal of operational research*, **234**, 561–570. 16

ALIA, I., CHIGHOUB, F. & SOHAIL, A. (2016). A characterization of equilibrium strategies in continuous-time mean–variance problems for insurers. *Insurance: Mathematics and Economics*, **68**, 212–223. 13, 15

ANKIRCHNER, S., HOBSON, D., STRACK, P. *et al.* (2015). Finite, integrable and bounded time embeddings for diffusions. *Bernoulli*, **21**, 1067–1088. 10

ARROW, K.J. (1971). The theory of risk aversion. *Essays in the theory of risk-bearing*, 90–120. 8, 12

ARTZNER, P., DELBAEN, F., EBER, J.M. & HEATH, D. (1999). Coherent measures of risk. *Mathematical finance*, **9**, 203–228. 15

BASAK, S. & CHABAKAURI, G. (2010). Dynamic mean-variance asset allocation. *Review of Financial Studies*, **23**, 2970–3016. 3, 11, 15

BAWA, V.S. & LINDENBERG, E.B. (1977). Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics*, **5**, 189–200. 16

BENSOUSSAN, A., WONG, K., YAM, S.C.P. & YUNG, S.P. (2014). Time-consistent portfolio selection under short-selling prohibition: From discrete to continuous setting. *SIAM Journal on Financial Mathematics*, **5**, 153–190. 13

BIELECKI, T.R., JIN, H., PLISKA, S.R. & ZHOU, X.Y. (2005). Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Mathematical Finance*, **15**, 213–244. 10

Bjork, T. & Murgoci, A. (2010). A general theory of markovian time inconsistent stochastic control problems. *Available at SSRN 1694759*. 3, 5, 11, 12, 13, 64

Bjork, T. & Murgoci, A. (2014). A theory of markovian time inconsistent stochastic control in discrete time. *Finance and Stochastics*, **18**, 545–592. 15, 40, 48, 62

Björk, T., Murgoci, A. & Zhou, X.Y. (2014). Mean–variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance*, **24**, 1–24. 12, 13, 19, 86, 95

Björk, T., Khapko, M. & Murgoci, A. (2017). On time-inconsistent stochastic control in continuous time. *Finance and Stochastics*, **21**, 331–360. 12

Chen, H.F. (2006). *Stochastic approximation and its applications*, vol. 64. Springer Science & Business Media. 71

Cui, X., Li, D., Wang, S. & Zhu, S. (2012). Better than dynamic mean-variance: Time inconsistency and free cash flow stream. *Mathematical Finance*, **22**, 346–378. 16

Czichowsky, C. (2013). Time-consistent mean-variance portfolio selection in discrete and continuous time. *Finance and Stochastics*, **17**, 227–271. 12

Djehiche, B. & Huang, M. (2016). A characterization of sub-game perfect equilibria for sdes of mean-field type. *Dynamic Games and Applications*, **6**, 55–81. 83

Ekeland, I. & Lazrak, A. (2006). Being serious about non-commitment: subgame perfect equilibrium in continuous time. *arXiv preprint math/0604264*. 12

El Karoui, N., Hamadène, S. & Matoussi, A. (2008). Backward stochastic differential equations and applications. *Indifference pricing: theory and applications*, 267–320. 67

Fleming, W.H. & Rishel, R.W. (2012). *Deterministic and stochastic optimal control*, vol. 1. Springer Science & Business Media. 3

Harlow, W.V. & Rao, R.K. (1989). Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence. *Journal of Financial and Quantitative Analysis*, **24**, 285–311. 16

Henderson, V. & Hobson, D. (2013). Risk aversion, indivisible timing options, and gambling. *Operations Research*, **61**, 126–137. 12

Hu, Y., Jin, H. & Zhou, X.Y. (2012). Time-inconsistent stochastic linear–quadratic control. *SIAM journal on Control and Optimization*, **50**, 1548–1572. 4, 5, 13, 15, 17, 38, 39, 67, 83, 86, 122

Hu, Y., Jin, H. & Zhou, X.Y. (2015). The uniqueness of equilibrium for time-inconsistent stochastic linear-quadratic control. *Available at SSRN 2589518*. 13, 39

Krokhmal, P., Palmquist, J. & Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of risk*, **4**, 43–68. 16

Krylov, N.V. (2008). *Controlled diffusion processes*, vol. 14. Springer Science & Business Media. 65, 70, 76

Lam, C.K., Xu, Y. & Yin, G. (2016). Dynamic portfolio selection without risk-free assets. *arXiv preprint arXiv:1602.04975*. 13

Li, D. & Ng, W.L. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, **10**, 387–406. 2, 5, 10, 96, 97

Li, X., Zhou, X.Y. & Lim, A.E. (2002). Dynamic mean-variance portfolio selection with no-shorting constraints. *SIAM Journal on Control and Optimization*, **40**, 1540–1555. 10

Li, Z., Zeng, Y. & Lai, Y. (2012). Optimal time-consistent investment and reinsurance strategies for insurers under hestons sv model. *Insurance: Mathematics and Economics*, **51**, 191–203. 13

Liang, Z. & Song, M. (2015). Time-consistent reinsurance and investment strategies for mean–variance insurer under partial information. *Insurance: Mathematics and Economics*, **65**, 66–76. 15

Lim, A.E. (2004). Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Mathematics of Operations Research*, **29**, 132–161. 10

Lim, A.E. & Zhou, X.Y. (2002). Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, **27**, 101–120. 10

MARKOWITZ, H. (1952). Portfolio selection. *The journal of finance*, **7**, 77–91. 1, 7

MARKOWITZ, H., TODD, P., XU, G. & YAMANE, Y. (1993). Computation of mean-semivariance efficient sets by the critical line algorithm. *Annals of Operations Research*, **45**, 307–317. 16

MARKOWITZ, H.M. (1968). *Portfolio selection: efficient diversification of investments*, vol. 16. Yale university press. 8, 16

MERTON, R.C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of economic theory*, **3**, 373–413. 1

MOSSIN, J. (1968). Optimal multiperiod portfolio policies. *The Journal of Business*, **41**, 215–229. 8

PEDERSEN, J. & PESKIR, G. (2015). Constrained dynamic optimality and binomial terminal wealth. *Research Report*, **10siem**. 14

PEDERSEN, J.L. & PESKIR, G. (2017). Optimal mean-variance portfolio selection. *Mathematics and Financial Economics*, **11**, 137–160. 14

PELEG, B. & YAARI, M.E. (1973). On the existence of a consistent course of action when tastes are changing. *The Review of Economic Studies*, **40**, 391–401. 87

PHELPS, E.S. & POLLAK, R.A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, **35**, 185–199. 87

PLISKA, S.R. (1982). A discrete time stochastic decision model. In *Advances in Filtering and Optimal Stochastic Control*, 290–304, Springer. 10

ROCKAFELLAR, R.T. & URYASEV, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, **2**, 21–42. 16

RUBINSTEIN, M.E. (1973). A comparative statics analysis of risk premiums. *The Journal of Business*, **46**, 605–615. 8

RUDLOFF, B., STREET, A. & VALLADÃO, D.M. (2014). Time consistency and risk averse dynamic decision models: Definition, interpretation and practical consequences. *European Journal of Operational Research*, **234**, 743–750. 16

SHARPE, W.F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, **19**, 425–442. 1

Strotz, R.H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, **23**, 165–180. 2, 11, 12

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The review of economic studies*, **25**, 65–86. 8

Vigna, E. (2016). On time consistency for mean-variance portfolio selection. 14, 15

Wang, J. & Forsyth, P.A. (2011). Continuous time mean variance asset allocation: A time-consistent strategy. *European Journal of Operational Research*, **209**, 184–201. 12

Wei, J., Wong, K., Yam, S. & Yung, S. (2013). Markowitzs mean–variance asset–liability management with regime switching: A time-consistent approach. *Insurance: Mathematics and Economics*, **53**, 281–291. 13

Williams, D. (1991). *Probability with martingales*. Cambridge university press. 70, 77

Yao, H., Lai, Y. & Li, Y. (2013). Continuous-time mean–variance asset–liability management with endogenous liabilities. *Insurance: Mathematics and Economics*, **52**, 6–17. 13

Yong, J. (2012). Time-inconsistent optimal control problems and the equilibrium hjb equation. *arXiv preprint arXiv:1204.0568*. 12

Yong, J. & Zhou, X.Y. (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43. Springer Science & Business Media. 67

Zeng, Y. & Li, Z. (2011). Optimal time-consistent investment and reinsurance policies for mean-variance insurers. *Insurance: Mathematics and Economics*, **49**, 145–154. 13, 15

Zeng, Y., Li, Z. & Lai, Y. (2013). Time-consistent investment and reinsurance strategies for mean–variance insurers with jumps. *Insurance: Mathematics and Economics*, **52**, 498–507. 13

Zhou, X.Y. & Li, D. (2000). Continuous-time mean-variance portfolio selection: A stochastic lq framework. *Applied Mathematics & Optimization*, **42**, 19–33. 10, 17, 72

Zhou, X.Y. & Yin, G. (2003). Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization*, **42**, 1466–1482. 10