Log Mean-Variance Portfolio Theory and Time Inconsistency

Trinity 2018

Geshan Rugjee

This thesis is dedicated to my parents and grandparents.

Abstract

In this paper, we will survey a particular class of continuous-time stochastic control problems which are time inconsistent. With time inconsistency, we will see that the Bellman optimality principle – which is the standard idea used to deal with stochastic control problems – will no longer be valid. Therefore, to solve a time-inconsistent stochastic control problem, we have to consider a different approach. One such approach that we will study is based on some of the concepts of game theory. More precisely, we will be looking for Nash subgame perfect equilibrium points for our problem. To determine such an equilibrium point (or strategy), we will make use of an extended version of the standard Hamilton-Jacobi-Bellman equation. In particular, we will use it to find an equilibrium strategy for a variant of the traditional problem presented in the Modern Portfolio Theory – which we will formulate in a continuous-time and dynamic setting. Finally, we will show how to derive an equivalent time-consistent formulation to our time-inconsistent problem.

Acknowledgement

I would like to thank my supervisor for the considerable support, illuminating ideas and aspiring guidance he has given me over the course of this project.

Contents

1 Introduction **5**

2 An overview of the Modern Portfolio Theory **7**

3 Log return Mean-Variance Portfolio Theory **12**

4 Stochastic Control and Time Inconsistency **16**
4.1 Equilibrium control and the extended HJB equation 19

5 Applications of the extended HJB equation **25**
5.1 Constant risk aversion index 31
5.2 Modified Basak-Chabakauri problem 36

6 Equivalent time consistent formulation **41**

7 Conclusion **43**

1 Introduction

One of the most important and enduring problems in finance relates to the construction of a portfolio with a certain risk-return profile. This is a difficult task not only because of the sheer number of securities that can be included in a portfolio but also because the returns on those securities at any given time are random variables. This, in turn, presents an additional difficulty as we will be required to build models for the returns on those securities. However, even if the returns were to be modelled accurately, the investor still faces the task of choosing the weights of the securities in his/her portfolio to achieve the expected return he/she seeks given the risk level that he/she is willing to undertake over a given investment period.

Of course, the complexity of the problem has prompted several economists, finance analysts and mathematicians to investigate the "ideal" way on how to construct a portfolio. Indeed, the literature on the subject is very rich and broad with different approaches being suggested. One such approach is the Modern Portfolio Theory (MPT) which was proposed by the economist Harry Markowitz [1] in 1952 and for which he was awarded the Nobel Prize. The MPT is definitely one of the most famous theories in finance which, to a large extent, is due to its simplicity and intuitive approach to building a portfolio.

In his work, Harry Markowitz assumed, quite reasonably, that investors are risk-averse – which loosely speaking means that investors shy away from taking on risk. Thus, given the same expected return, investors would prefer an asset which is less risky over a more risky asset. Therefore, Markowitz proposes a methodology whereby to build a portfolio, an investor should set an expected return that he/she is aiming for and find the corresponding weights of the different securities in his/her portfolio such that the risk undertaken is minimised.

Mathematically speaking, this is a constrained optimisation problem whereby we are trying to find the weights of the different securities which result in the least risk given a prescribed expected return level. It should also be clear that we have one additional constraint which is that the weights of the securities should add up to one. Equivalently, the investor could set the risk level he/she is prepared to accept and then maximise the expected return on his/her portfolio.

While Markowitz's Modern Portfolio Theory has revolutionised the way that we think about portfolio construction, its limitations must, however, not be overlooked. First and foremost, for the constrained optimisation problem to make sense, we need to find a metric to measure how risky our portfolio is. In Markowitz view – and as is the general view in finance – risk is measured by the variance of the returns of our portfolio. This, in itself, is a major assumption and its accurate estimation using historical data can be fiercely debated; and thus raising questions as to whether this methodology is reasonable. This is why different metrics, most notably the Value-at-Risk, have been considered (Brachinger [1999] [2], Sentana [2003] [3] and Byrne & Stephen Lee [2004][4] to mention a few). However, since the choice of metric is not the focus of this paper, we will therefore assume that the variance of the returns depicts accurately the risks in a portfolio.

One of the most important drawback, however, is that the constrained optimisation problem stated above is a static one. What this means is that our optimisation is over a single time period $[0, T]$ and that the optimal weights found at time $t = 0$ will be used only once throughout the interval $[0, T]$. It is clear that this is a major drawback and is highly unrealistic in practice as investment managers/investors actively manage their portfolio and do not simply choose the weights of the securities at time $t = 0$. Indeed, as time goes by, it is more realistic and advantageous for investment managers to react to the additional information that they have gathered rather than not doing anything at all.

One of the ways to address the above problem is to change the static Markowitz formulation into a dynamic one. This was famously done in the discrete setting by Hakansson [1971] [5] and Samuelson [1969] [6]. However, given that the present technology allows for High-Frequency Trading, it seems more appropriate and realistic to formulate our problem in continuous time rather than in discrete time (see Merton [1969] [7] and Bajeux-Besnainou & Portait [1998] [8]). What this means is that instead of solving only one constrained optimisation problem, we will solve the problem "continuously" at every time $t \in [0, T]$ conditional on the additional information that we have gathered up to that point in time. We will then use the "optimal" weights that we have found to rebalance our portfolio at every time $t \in [0, T]$. This, of course, is very similar to the approach proposed by Bajeux-Besnainou & Portait [1998] where continuous rebalancing was also allowed.

Naturally, when we formulate our problem in the continuous and dynamic setting, the degree of risk-aversion of the investor will not be required to be static (as in the MPT) and can change depending on factors such as the current wealth and the remaining time in his/her investment period. In effect, this means that when we formulate our problem, we will allow the investor's preference/risk-appetite to change over time.

Our aim will then be to find an optimal control law which, in our case, represents the proportion of our wealth that we should invest in the risky asset at any given time and wealth level. However, it turns out that our optimisation problem suffers from time inconsistency which means that the traditional concept of optimality will no longer apply. Moreover, this will also mean that we will not be able to use the standard Hamilton-Jacobi-Bellman (HJB) equation to solve our problem. So, we will need an alternative angle to tackle our problem.

The alternative that we will consider is the game theoretic approach which has been discussed in detail in the works of Ekeland & Lazrak [2006] [9] and Ekeland & Privu [2007] [10]. In particular, to solve our problem, we will make heavy use of the concept of equilibrium control and the extended HJB equation which was introduced and defined rigorously in those papers. In short, our problem will be viewed as a series of games whereby at every point in time, the player is allowed to choose a control. The key to solving the problem will then be to, loosely speaking, try to find the control law which will result in a subgame perfect Nash equilibrium. The control found is what we will call an equilibrium control – which is what we will attempt to solve for in this paper.

To achieve that, however, we will first need to study the Markowitz formulation, discuss its limitation and develop an intuition on how to formulate a similar problem in continuous time and in a Black-Scholes economy. This will be done in section 2 and 3. We will then move on to section 4 where we will discuss the issues that our formulation faces – the main one being time inconsistency. In section 5, the equilibrium controls for our problem will be derived. Finally, in section 6, we will construct a time-consistent formulation for our time-inconsistent problem.

2 An overview of the Modern Portfolio Theory

Before presenting the ideas of the MPT, let us first state some of the assumptions involved in the model. So, we will assume that the investor is risk-averse and has an investment period of $[0, T]$, where $T > 0$. Moreover, suppose that we are in a frictionless economy where there are $n \in \mathbb{N}_{>0}$ risky assets from which we can build a portfolio. Let us denote the value of those risky assets at time $t \in [0, T]$ by $S_1(t), \cdots, S_n(t)$ respectively, which of course are non-negative random variables. Then, we can define the single-period return on asset $i \in \{1, \cdots, n\}$ over $[0, T]$ as $R_i(T) := \dfrac{S_i(T) - S_i(0)}{S_i(0)}$.

To build a portfolio, we have to choose which assets to include and in what proportion to include them. In other words, we have to choose the weights of each of the n risky assets in the portfolio. However, the caveat here is that we can choose the weights in the portfolio **only** at time $t = 0$. So, suppose that we buy/short $\Delta_i \in \mathbb{R}$ of asset $i \in \{1, \cdots, n\}$ – assuming that the assets are perfectly divisible – at time $t = 0$. Therefore, the value of our portfolio at any time $t \in [0, T]$, denoted by $X(t)$, is given by:

$$X(t) = \sum_{i=1}^{n} \Delta_i S_i(t).$$

Thus, the one-period return over $[0, T]$ on the portfolio is given by:

$$\begin{aligned} R(T) &:= \frac{\sum_{i=1}^{n} \Delta_i S_i(T) - \sum_{i=1}^{n} \Delta_i S_i(0)}{\sum_{i=1}^{n} \Delta_i S_i(0)} \\ &= \frac{\sum_{i=1}^{n} \Delta_i [S_i(T) - S_i(0)]}{\sum_{i=1}^{n} \Delta_i S_i(0)} \\ &= \sum_{i=1}^{n} \left[\frac{\Delta_i S_i(0)}{\sum_{j=1}^{n} \Delta_j S_j(0)} \right] R_i(T) \\ &= \sum_{i=1}^{n} w_i R_i(T). \end{aligned}$$

Therefore, the one-period return on the portfolio over $[0, T]$ is actually the weighted average of the single-period returns on each risky asset over $[0, T]$ with the weight of asset $i \in \{1, \cdots, n\}$ being given by $w_i := \dfrac{\Delta_i S_i(0)}{\sum_{j=1}^{n} \Delta_j S_j(0)}$. Moreover, notice that we have $\sum_{i=1}^{n} w_i = 1$.

Now, suppose that $\mathbb{E}[R_i(T)] = \mu_i \in \mathbb{R}_{>0}$ and that $\mathrm{Cov}[R_i(T), R_j(T)] = \sigma_{ij} \in \mathbb{R}$, where $i, j \in \{1, \cdots, n\}$. Therefore, if we define $\boldsymbol{w}^T := (w_1, \cdots, w_n)$, $\boldsymbol{\mu}^T := (\mu_1, \cdots, \mu_n)$, $\boldsymbol{\Sigma}_{i,j} := \sigma_{ij}$ and $\boldsymbol{R}^T := (R_1(T), \cdots, R_n(T))$, then we can calculate the expected value [11], denoted by $\mu_{\mathbf{w}}$, of the portfolio return at time T as follows:

$$\begin{aligned} \mu_{\mathbf{w}} &= \mathbb{E}[R(T)] \\ &= \sum_{i=1}^{n} w_i \mathbb{E}[R_i(T)] \\ &= \mathbf{w}^T \boldsymbol{\mu}. \end{aligned}$$

The variance [12], denoted by $\sigma^2_{\mathbf{w}}$, of the portfolio return at time T is given by:

$$\begin{aligned}\sigma^2_{\mathbf{w}} &= \text{Cov}[R(T), R(T)] \\ &= \text{Cov}[\mathbf{w}^T\boldsymbol{R}, \mathbf{w}^T\boldsymbol{R}] \\ &= \mathbf{w}^T\text{Cov}[\boldsymbol{R}, \boldsymbol{R}]\mathbf{w} \\ &= \mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}.\end{aligned}$$

Having shown the preliminary results on how to find the expected value and variance of the portfolio return at time T for a certain weight vector $\mathbf{w}$, the question we face now is how to choose the weights in order to meet a given objective.

Before formulating the objective as proposed by Markowitz, it is necessary to understand how to rank two portfolios with different weights in the Markowitz mean-variance sense [13]. So, suppose we have two portfolios A and B with the weights of assets at time $t = 0$ being $\mathbf{w}$ and $\mathbf{w'}$ respectively. Then, we say that we prefer portfolio A to B in the mean-variance sense if $\mu_{\mathbf{w}} \geq \mu_{\mathbf{w'}}$ and $\sigma^2_{\mathbf{w}} \leq \sigma^2_{\mathbf{w'}}$ with at least one strict inequality.

As was mentioned above, the variance of the returns is a measure of the risk associated with a portfolio. Therefore, what the above definition suggests is that a risk-averse investor will always prefer a portfolio A over a portfolio B [14] if:

1. Portfolio A has an expected return at time T at least as high as portfolio B given that it is also strictly less risky than portfolio B.

2. Portfolio A has a strictly higher expected return at time T than portfolio B given that its risk is at most at the same level as portfolio B.

This order preference of portfolios above brings us to the definition of efficient portfolios – which is what Markowitz aims to find. So, a portfolio A is said to be efficient if there is no other portfolio B such that we prefer B to A in the mean-variance sense as described above [15]. Using the definition of the order preference of portfolios given above and that of efficient portfolios, it should be clear that the objective of the MPT is then to find those portfolios that give the highest expected return for a given risk level. Alternatively, from the definition, this is also equivalent to finding those portfolios which are the least risky for a given level of expected return.

Having found the efficient portfolios, we can then plot the highest expected return attainable for each level of risk associated with a portfolio. This is more commonly known as the efficient frontier.

Of course, the above problem has been extended to include a risk-free asset in the economy, but we will not go into further detail about these extensions in this paper.

So, mathematically, to find the weights of an efficient portfolio associated with an expected return of $r^* \in \mathbb{R}_{>0}$ at T, we need to solve the following problem [16]:

$$\min_{\mathbf{w}\in\mathbb{R}^n} \text{Var}[R(T)],$$
$$\text{s.t.} \;\; \mathbb{E}[R(T)] = r^* \text{ and } \sum_{i=1}^{n} w_i = 1.$$

Alternatively, to find the weights of an efficient portfolio associated with a prescribed risk level $(\sigma^*)^2 \in \mathbb{R}_{>0}$ at time T, we need to solve the following problem [17]:

$$\max_{\mathbf{w}\in\mathbb{R}^n} \; \mathbb{E}[R(T)],$$
$$\text{s.t.} \text{Var}[R(T)] = (\sigma^*)^2 \text{ and } \sum_{i=1}^{n} w_i = 1.$$

While the above formulations both yield efficient portfolios, we will focus more on another equivalent formulation [18] which will prove to be more useful when we present our formulation of the continuous (and modified) version of the Markowitz problem. This is given as follows:

$$\max_{\mathbf{w}\in\mathbb{R}^n} \; \mathbb{E}[R(T)] - \lambda \text{Var}[R(T)], \tag{‡}$$
$$\text{s.t.} \sum_{i=1}^{n} w_i = 1.$$

In the above formulation, λ is a positive real number (as investors are assumed to be risk-averse) and is commonly referred to as the Arrow-Pratt risk aversion index [19]. This is used as a measure of risk-aversion of an investor and intuitively reflects the trade-off between risk and reward that an investor is willing to accept. Indeed, the higher the Arrow-Pratt risk aversion index is, the more risk-averse the investor is and vice-versa. This can be easily seen from (‡) above, as the higher that λ is, the more will our objective function be penalised - as variance is always non-negative.

Before moving on and using a variant of (‡) to formulate the dynamic and continuous version of our problem, we must first ensure that solving (‡) is indeed equivalent to the initial Markowitz constrained optimisation problem. This is proved below.

Proof. Consider a fixed $\lambda > 0$ and suppose that by solving the constrained optimisation problem (‡), we obtain the following optimal weights $\boldsymbol{w}^*_{\boldsymbol{\lambda}} = (w^*_{1,\lambda}, \cdots, w^*_{n,\lambda})^T$.

So, we have to show that there is no other portfolio with different weights such that we prefer it to the one with weight vector $\boldsymbol{w}^*_{\boldsymbol{\lambda}}$.

First of all, notice that $\boldsymbol{w}^*_{\boldsymbol{\lambda}}$ is admissible as the sum of the individual weight components of $\boldsymbol{w}^*_{\boldsymbol{\lambda}}$ is equal to one.

Now, suppose that $\boldsymbol{w} = (w_1, \cdots, w_n)^T$, where $w_1, \cdots, w_n \in \mathbb{R}$, such that $\sum_{i=1}^n w_i = 1$. Then, by the principle of optimality, we have:

$$\begin{aligned}
& \mathbb{E}[R_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}(T)] - \lambda \mathrm{Var}[R_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}(T)] \geq \mathbb{E}[R_{\boldsymbol{w}}(T)] - \lambda \mathrm{Var}[R_{\boldsymbol{w}}(T)] \\
\Longrightarrow \quad & \mathbb{E}[R_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}(T)] - \mathbb{E}[R_{\boldsymbol{w}}(T)] \geq \lambda \{\mathrm{Var}[R_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}(T)] - \mathrm{Var}[R_{\boldsymbol{w}}(T)]\} \\
\Longrightarrow \quad & \mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \mu_{\boldsymbol{w}} \geq \lambda \{\sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \sigma^2_{\boldsymbol{w}}\} \qquad (\dagger)
\end{aligned}$$

We will now prove our claim by contradiction. So suppose that we prefer the portfolio with weight vector $\boldsymbol{w}$ over one with weights given by $\boldsymbol{w}^*_{\boldsymbol{\lambda}}$ in the mean-variance sense. Therefore, we must have $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} \leq \mu_{\boldsymbol{w}}$ and $\sigma^2_{\boldsymbol{w}} \leq \sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}$ with at least one strict inequality.

Suppose in the first instance that $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} \leq \mu_{\boldsymbol{w}}$ and $\sigma^2_{\boldsymbol{w}} < \sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}$. So, from (†) above, $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \mu_{\boldsymbol{w}} \geq \lambda\{\sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \sigma^2_{\boldsymbol{w}}\} > 0$. So, $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \mu_{\boldsymbol{w}} > 0 \iff \mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} > \mu_{\boldsymbol{w}}$ – which is a contradiction.

On the other hand, suppose that $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} < \mu_{\boldsymbol{w}}$ and $\sigma^2_{\boldsymbol{w}} \leq \sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}$. Then, using (†) again, we have $\lambda\{\sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \sigma^2_{\boldsymbol{w}}\} \leq \mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} - \mu_{\boldsymbol{w}} < 0$. Since, $\lambda > 0$, we have $\sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} < \sigma^2_{\boldsymbol{w}}$. This of course is a contradiction.

In the case where $\mu_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}} < \mu_{\boldsymbol{w}}$ and $\sigma^2_{\boldsymbol{w}} < \sigma^2_{\boldsymbol{w}^*_{\boldsymbol{\lambda}}}$ holds, the contradiction is obvious.

Therefore, we must have that the portfolio with weights $\boldsymbol{w}^*_{\boldsymbol{\lambda}}$ at time $t = 0$ must indeed form an efficient portfolio.

□

So, by solving the formulation (‡), we know that we will get an efficient portfolio provided that it exists for a given $\lambda > 0$. Notice that different values of λ will correspond to expected returns and variances of different efficient portfolios [20].

It must be noted that the formulation (‡) can be thought to be a slight simplification of two formulations initially presented as we have one less constraint to deal with – the preset expected returns at time T in the first case and the preset variance at time T in the second case. This type of formulation, although equivalent to the first two formulations, is therefore easier to construct and to deal with. In fact, when moving to the continuous and dynamic setting, we will also benefit from this advantage; which is why we will make use of a formulation akin to (‡) when presenting our problem.

3 Log return Mean-Variance Portfolio Theory

As we have seen above, the traditional MPT problem is a single-period one. As such, we are solving a single optimisation problem at time $t = 0$ and use the corresponding optimal weights to build our portfolio at time $t = 0$. Then, from $t = 0$ onwards till $t = T$, the portfolio will be left undisturbed. Clearly, this is highly unrealistic in practice, as portfolio managers constantly rebalance their portfolios to try to meet their specific objectives.

So, it is clear that we need to move in a dynamic setting to better represent reality. Moreover, we will assume that we are in a continuous time setting. Of course, this is not exactly accurate; but then again, it is not entirely unrealistic given the popularity of High-Frequency Trading in the financial industry. So, operating in continuous time and in a dynamic setting gives us the possibility to adjust the weights of the assets in our portfolio at any moment in time. Obviously, our aim will then be to adjust those weights optimally given what we know now and have already seen in the past. This is the motivation for our problem. However, before delving further into it, we will first state our assumptions.

We will assume a frictionless and complete market similar to the Black-Scholes setting (including the filtration) whereby there are only two assets: a stock which is risky and a bond which is risk-free. Furthermore, we assume that the risk-free rate is given by $r \in \mathbb{R}_{>0}$ and that the dynamics of assets are given by:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad \text{and} \quad dB_t = rB_t dt.$$

In the above, S_t and B_t represent the prices of the stock and of the bond at time t respectively; while W_t represents a Wiener process. Also, we have $\mu, \sigma \in \mathbb{R}_{>0}$. Since the stock is a risky asset, we expect μ to be greater than r. We will therefore assume this to be true in this paper.

We will now move on to construct a self-financing portfolio out of those two assets. So, suppose that at any time t over a period $[0, T]$, our wealth/portfolio is valued at $Y(t) \in \mathbb{R}$ and that we hold $\Delta(t) \in \mathbb{R}$ stocks. Then,

$$\begin{aligned} Y(t) &= \Delta(t)S_t + \left(\frac{Y(t) - \Delta(t)S_t}{B_t}\right) B_t \\ \implies dY(t) &= \Delta(t)dS_t + \left(\frac{Y(t) - \Delta(t)S_t}{B_t}\right) dB_t \\ \implies dY(t) &= \Delta(t)[\mu S_t dt + \sigma S_t dW_t] + r[Y(t) - \Delta(t)S_t]dt \\ \implies dY(t) &= [rY(t) + (\mu - r)\Delta(t)S_t]dt + [\sigma\Delta(t)S_t]dW_t. \end{aligned} \tag{1}$$

Now, in the traditional Markowitz formulation, the return used is a discrete-period one. Moreover, the return is not annualised which means that it does not take into account the length of the investment horizon. To remediate this issue and given that we are operating in continuous time, it makes sense and is actually advantageous for us to instead use the annualised log-returns when defining our problem.

Thus, let us define by $R(t)$ the annualised log-return on our portfolio over the interval $[t, T]$. Thus,

$$R(t) = \frac{1}{T-t} \log\left[\frac{Y(T)}{Y(t)}\right] = \frac{1}{T-t}\left[X(T) - X(t)\right].$$

Notice that given our definition of the annualised log-return $R(t)$, the latter will only make sense if and only if our wealth process is always positive. Thus, if we are to use $R(t)$ in our model, we are implicitly implying that our wealth can never be negative. However, since $Y(t)$ is governed by the SDE given by (1) above, there is no reason why it should always be positive for any $\Delta(t)$. Therefore, to explicitly ensure that our wealth process is always positive, we will assume that at every point in time, the value of stocks we hold in our portfolio is proportional to our current wealth. To that end, we will denote by u_t the proportion of wealth that we hold in stocks at time t.

On a more intuitive level, the proportionality assumption is explained by the fact that should the wealth of a **risk-averse** investor ever reach zero, then to avoid the possibility of it going negative, it makes sense for him/her not to hold any risky asset. Therefore, by holding stock value proportional to our wealth, if the latter reaches zero, then we will not have any position in the risky asset. In fact, as we will show below, this assumption will actually ensure that our wealth always remains positive provided that the investor starts with a positive wealth.

So, let us assume that at any time $t \in [0, T]$ that the value of stocks that we hold is proportional to our wealth i.e. $\Delta(t)S_t = u_t Y(t)$, where $(u_t)_{0 \leq t \leq T}$ is a real-valued process. Then from (1) above, we get:

$$\begin{aligned}
& dY(t) = [rY(t) + (\mu - r)\Delta(t)S_t]dt + [\sigma\Delta(t)S_t]dW_t \\
\implies \quad & dY(t) = [rY(t) + (\mu - r)u_t Y(t)]dt + [\sigma u_t Y(t)]dW_t \\
\implies \quad & dY(t) = Y(t)\left\{[r + (\mu - r)u_t]dt + [\sigma u_t]dW_t\right\}.
\end{aligned}$$

The solution to the above SDE is given by the Doléans-Dade exponential [21]. Therefore, our wealth at time T starting at time $t \in [0, T]$ is given by:

$$Y(T) = Y(t)\exp\left\{\int_t^T r + (\mu - r)u_s - \frac{\sigma^2 u_s^2}{2}\, ds + \int_t^T \sigma u_s\, dW_s\right\}.$$

Notice therefore that if we assume that $\Delta(t)S_t = u_t Y(t)$, then if we start off with a positive initial wealth i.e. $Y(0) > 0$, then our wealth process will always be positive – which is what we wanted. This implies that $X(t) := \log[Y(t)]$ is a well defined process and can be calculated to be:

$$\begin{aligned}
& \log[Y(T)] = \log[Y(t)] + \int_t^T r + (\mu - r)u_s - \frac{\sigma^2 u_s^2}{2}\, ds + \int_t^T \sigma u_s\, dW_s \\
& \implies X(T) = X(t) + \int_t^T r + (\mu - r)u_s - \frac{\sigma^2 u_s^2}{2}\, ds + \int_t^T \sigma u_s\, dW_s \\
& \implies dX(t) = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t}{2}\right] dt + [\sigma u_t]\, dW_t.
\end{aligned}$$

Notice that we have assumed that $(u_t)_{0 \leq t \leq T}$ is a regular enough process for the above derivations to hold. As we shall see in the next section, we will assume that $u_t = \boldsymbol{u}(t, X(t))$, where $\boldsymbol{u} : [0, T] \times \mathbb{R} \to \mathbb{R}$ is a deterministic function. For the time being, however, it is only necessary to note that u_t will serve as our control at time t and more generally, we will assume that the control process lives in a suitable space $\mathcal{U}$ and takes values in the restricted subset $U \subseteq \mathbb{R}$.

So, from the above, we have that $X(t)$ is given by:

$$X(T) = X(t) + \int_t^T r + (\mu - r)u_s - \frac{\sigma^2 u_s^2}{2} \, ds + \int_t^T \sigma u_s \, dW_s \; .$$

Therefore, if we suppose that u_t is a deterministic function of time only i.e. $u_t = \boldsymbol{u}(t)$, then the distribution of $X(T)$ given $X(t) = x \in \mathbb{R}$ is given by:

$$X(T)\,|_{X(t)=x} \quad \sim \; \mathcal{N}\left(x + \int_t^T r + (\mu - r)\boldsymbol{u}(s) - \frac{\sigma^2 \boldsymbol{u}^2(s)}{2} \, ds, \int_t^T \sigma^2 \boldsymbol{u}^2(s) \, ds \right)$$

Having ensured that $R(t)$ is now well defined over $[0, T]$, let us now return back to our objective. In a similar fashion to the discrete case, we wish to minimise the variance of the annualised log-return on our portfolio given a prescribed level of expected return. This is of course equivalent to maximising the expected return on our portfolio given a certain risk level because of the ranking methodology of portfolios that we discussed before.

Notice that since we are now operating in continuous time, our risk-preference can also change as time goes by. Therefore, instead of using only a constant and positive risk aversion index λ as in the Markowitz formulation, we can now use a positive function $\lambda(t, x)$ to reflect our changing risk-preference. Note that while the risk aversion index depends on a number of factors in reality, we will assume in this paper that it depends on where we are in our investment horizon $[0, T]$ and our current log-wealth.

Our problem is thus to find the control law which solves the following optimisation problem:

$$\max_{u \in \mathcal{U}} \mathbb{E}_{t,x}[R_T] - \lambda(t, x)\text{Var}_{t,x}[R(T)],$$

s.t. $dX(t) = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2} \right] dt + [\sigma u_t] \, dW_t.$

We will now simplify our objective function above and represent it in standard formulation – which will be useful when we present the extended HJB equation in the next section.

So, $\mathbb{E}_{t,x}[R(T)] = \frac{1}{T-t}\{\mathbb{E}_{t,x}[X(T)] - \mathbb{E}_{t,x}[X(t)]\} = \frac{1}{T-t}\{\mathbb{E}_{t,x}[X(T)] - x\}$, by simple linearity.

Also, $\text{Var}_{t,x}[R(T)] = \text{Var}_{t,x}\left\{\frac{1}{T-t}\left[X(T) - X(t)\right]\right\} = \frac{1}{(T-t)^2}\left\{\mathbb{E}_{t,x}[X^2(T)] - (\mathbb{E}_{t,x}[X(T)])^2\right\}.$

Therefore, our objective function which we will henceforth denote by $J(t, x, \mathbf{u})$, where (t, x) is our starting point and $\boldsymbol{u}$ is a control law, is given by:

$$J(t, x, \mathbf{u}) = \frac{1}{T-t}\{\mathbb{E}_{t,x}[X(T)] - x\} - \frac{\lambda(t,x)}{(T-t)^2}\left\{\mathbb{E}_{t,x}[X^2(T)] - (\mathbb{E}_{t,x}[X(T)])^2\right\}.$$

Our aim is therefore to find $\max_u J(t, x, \mathbf{u})$ for every $(t, x) \in [0, T] \times \mathbb{R}$ given our SDE of $X(t)$ above. This, of course, is mathematically equivalent to solving the problem given below:

$$\max_{u \in \mathcal{U}} \; \mathbb{E}_{t,x}[(T-t)X(T) - \lambda(t,x)X^2(T)] + \lambda(t,x)\left(\mathbb{E}_{t,x}[X(T)]\right)^2, \tag{$\dagger$}$$

s.t. $dX(t) = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right] dt + [\sigma u_t]\, dW_t.$

Presented in the above form, it is clear the above closely resembles a standard stochastic control problem. However, unlike the standard stochastic control problem, in this case, we cannot make use of the standard HJB equation to derive an optimal control law. This is because the problem outlined above turns out to be time inconsistent. This, in turn, makes the concept of optimality very problematic and thus requires a different approach to solve as we will see in the next section.

4 Stochastic Control and Time Inconsistency

In the standard stochastic optimal control problem [22], we are faced with problems of the following form:

$$\max_{u \in \mathcal{U}} \; \mathbb{E}_{t,x}\left[\int_t^T C(s, X_s^u, u_s)ds + F(X_T^u)\right], \tag{$\star$}$$

where u_t is the control process and X_t^u is a controlled Markov process taking real values which is governed by the following SDE:

$$dX_t^u = \mu(t, X_t^u, u_t)dt + \sigma(t, X_t^u, u_t)dW_t; \quad X_t = x. \tag{$\star\star$}$$

In the above, $\mu(t, X_t^u, u_t)$ and $\sigma(t, X_t^u, u_t)$ are both real-valued functions and W_t is a Wiener process. Moreover, from ($\star$), the function $C(t, X_t^u, u_t)$ is often called the running reward/penalty and in the standard formulation of a stochastic control problem, this is allowed to depend on the time t, the controlled process at time t i.e. X_t^u and the control itself at time t i.e. u_t. Notice, however, that the bequest/terminal reward represented by the function $F(.)$ is only allowed to depend on the terminal value of the controlled process i.e. X_T^u.

As far as the controls are concerned, we will be using only admissible feedback control laws [23]. By feedback control (or Markovian control), we mean that the control at time t, where $X_t^u = x$, is given by $\boldsymbol{u}(t,x)$, with $\boldsymbol{u}(.,.)$ being a deterministic real-valued function. By admissible, we mean that $\boldsymbol{u}(t,x) \in U$ for every $(t,x) \in [0,T] \times \mathbb{R}$ and that the SDE $(\star\star)$ corresponding to the control law $\boldsymbol{u}(t,X_t)$ admits a unique strong solution for every starting point $(s,y) \in [0,T] \times \mathbb{R}$. The use of feedback controls is an important assumption not only in the standard formulation of the stochastic control problem but also in a more general setting, which we will present later in this section, as it helps our model to maintain a Markovian structure [24]. Without this crucial assumption, all the results presented in this paper would be invalidated. In fact, getting any useful results at all without this assumption becomes incredibly hard (see Björk [2010] [25] for more detail). Therefore, in this paper, we will consider only admissible feedback control laws.

Now, in the standard stochastic problem problem, our aim is to solve $(\star)$ for every starting point t and every initial value x of the controlled process X_t^u. Intuitively, it would seem reasonable for us to expect that the optimal control law would be different for every starting point and initial wealth (t,x). However, it turns out that problems of this form have the time consistency property. What this means is that no matter where we start and what our initial wealth is, the optimal control law stays the same. So, the optimal law on $[0,T]$ will still remain optimal when restricted to any other interval $[t,T]$, where $t < T$ [26]. Mathematically, if we denote the optimal control at (s,X_s) when starting at (t_1,x_1) by $\boldsymbol{u}^*_{t_1,x_1}(s,X_s)$ and the optimal control at (s,X_s) when starting at (t_2,x_2) by $\boldsymbol{u}^*_{t_2,x_2}(s,X_s)$, where $0 \leq t_1 \leq t_2 \leq s \leq T$, then we say that the standard problem has the time consistency property[27] if:

$$\boldsymbol{u}^*_{t_1,x_1}(s,X_s) = \boldsymbol{u}^*_{t_2,x_2}(s,X_s),\ t_2 \leq s \leq T.$$

Notice that this means that we only need to solve the problem at time $t = 0$ to get the optimal control law over any other sub-interval $[t,T], t \leq T$.

The time consistency property as described above is a direct consequence of the Bellman optimality equation [28], which in this case can be summarised as follows:

$$H(t,x) = \sup_{u \in \mathcal{U}} \ \mathbb{E}_{t,x}\left[\int_t^\tau C(s,X_s^u,u_s)ds + H(\tau,X_\tau^u)\right],$$

where the above holds for all $(t,x) \in [0,T] \times \mathbb{R}$ and all stopping times $\tau \leq T$, and where we have $H^u(t,x) = \mathbb{E}_{t,x}\left[\int_t^T C(s,X_s^u,u_s)ds + F(X_T^u)\right]$ and $H(t,x) = \sup_u H^u(t,x)$.

Thus, the optimal control law starting at (t, x) coincides with the same optimal control law starting from (τ, X_τ) for every stopping time τ – in other words, we have time consistency.

To be able to derive the Bellman optimality equation above and have time consistency, we need our problem to be in the standard form. The main reasons as to why the standard problem has the time consistency [29] property are as follows:

1. The running reward/penalty $C(s, X_s^{\boldsymbol{u}}, u_s)$ is not allowed to depend on our starting time t and our initial wealth x.
2. The bequest/terminal reward is not allowed to depend on (t, x).
3. Most importantly, however, the terminal reward is of the form $\mathbb{E}_{t,x}[F(X_T^{\boldsymbol{u}})]$. Then, the Bellman optimality equation can then be derived from the tower property. In particular, we are not allowed to have a term in the form $G\left(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]\right)$, where $G(.)$ is a non-linear function. This is because we cannot, in general, apply the tower property for any non-linear function $G(.)$ and any random variable $X_T^{\boldsymbol{u}}$. More precisely, $\mathbb{E}_{\tau, X_\tau^{\boldsymbol{u}}}[G\left(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]\right)] \neq G\left(\mathbb{E}_{\tau, X_\tau^{\boldsymbol{u}}}[X_T^{\boldsymbol{u}}]\right)$ in general.

Notice that for our problem (†) presented in section 3, we can define $H^{\boldsymbol{u}}(t, x) := \mathbb{E}_{t,x}[(T-t)X(T) - \lambda(t,x)X^2(T)] + \lambda(t,x)\left(\mathbb{E}_{t,x}[X(T)]\right)^2$. We can see that our terminal reward/penalty is a function of (t, x). Moreover, we have a non-linear function $G(t, x, y) = \lambda(t, x)y^2$ acting on $\mathbb{E}_{t,x}[X(T)]$. Therefore, given the reasons presented above, our formulation will not exhibit time consistency – we call such a problem a time inconsistent one.

Clearly, time inconsistency is a serious problem because even if we were to solve the problem starting at (t, x) and obtain some optimal control law, then when we move along in time – to say $t + \Delta t$ and with new wealth $x + \Delta x$ – then the "optimal" control law that we have previously obtained will no longer be optimal. In other words, once we move from (t, x) to $(t + \Delta t, x + \Delta x)$, then it is possible for us to improve our functional given the information that we have gathered on $[t, t + \Delta t]$ if we choose a control law different from the "optimal" one that we calculated at (t, x). Therefore, it is clear that very concept of optimality becomes quite problematic and therefore, to solve the problem, we need to try a different approach.

One such approach, known as pre-commitment [30], is where we solve the optimisation problem at (t, x) and use the control law we obtain on $[t, T]$ despite knowing that the control law obtained will no longer be optimal in the future.

Another interesting methodology, and one that we will consider in this paper, is known as the game-theoretic approach – which we will introduce in the next subsection.

4.1 Equilibrium control and the extended HJB equation

To get an idea of how to think about the game-theoretic approach and the equilibrium control, we will first discuss it in the discrete setting [31]. Thereafter, we will use the same ideas to generalise the approach in continuous time. Before we move on, however, we will henceforth denote the control laws in bold (like $\mathbf{u}$) whereas the possible values they take will not be denoted in bold (e.g: $u_n = \boldsymbol{u_n}(t_n, x)$).

So, suppose we are facing a time inconsistent optimisation problem where we want to maximise a given functional $J_{t_0}(x, \boldsymbol{u})$ over a discrete time period $0 = t_0 < \cdots < t_n = T$, where $n \geq 1$. Moreover, we will assume that our functional is of the form:

$$J_i(x, \boldsymbol{u}) = \mathbb{E}_{t_i,x}\left[\sum_{k=i}^{n-1} C(x, X_k^{\boldsymbol{u}}, \boldsymbol{u}_k(X_k^{\boldsymbol{u}})) + F(x, X_T^{\boldsymbol{u}})\right] + G\left(x, \mathbb{E}_{n,x}[X_T^{\boldsymbol{u}}]\right).$$

Unlike in a time consistent problem where we can simply find the optimal control at time t_0 to solve the entire problem; in a time inconsistent problem, if a control law $\hat{\boldsymbol{u}}$ is optimal starting at (t_k, x), then it will no longer be optimal from (t_j, X_j) onwards, for $j > k$. An informal way to think of it is that our preferences are changing over time [32].

So, instead we will make use of the game-theoretic approach. First, remember that at every point in time $t_0, \cdots, t_{n-1}$, we can **only** choose **one** control u_{t_k} which will be of the form $\boldsymbol{u}(t_k, X_{t_k}), k = 0, \cdots, n-1$. The idea then behind the game-theoretic approach is to view the above problem as consisting of n games whereby game k, where $0 \leq k \leq n-1$, is played at time t_k and where we can choose **only** our control which will be in the form of u_{t_k} as described above [33].

Since we know that our preferences are changing in the future, our choice of the control u_{t_k} at time t_k should reflect that fact. In other words, we will choose our control at time t_k by taking into consideration our future changing preferences when we play the games at times $t_{k+1}, \ldots, t_{n-1}$. We then proceed backwards in time, starting at time t_{n-1}, and solve the standard (and static) optimisation problem of maximising our objective function $J_{n-1}(x, \boldsymbol{u})$ for every starting point $x \in \mathbb{R}$. We will denote the control law we obtain as $\hat{\boldsymbol{u}}_{\boldsymbol{t_{n-1}}}$. This represents the first equilibrium control.

Then at time t_{n-2}, we will solve the optimisation problem $\sup_{u_{t_{n-2}}} J_{n-2}(x, \boldsymbol{u})$ – where we can only choose $u_{t_{n-2}}$ – for every starting point $x \in \mathbb{R}$ given that once we reach t_{n-1}, we will use the control law $\boldsymbol{\hat{u}_{t_{n-1}}}$. The control law we obtain as a result, denoted by $\boldsymbol{\hat{u}_{t_{n-2}}}$, represents the equilibrium control law at time t_{n-2}.

Therefore, proceeding inductively backwards in time [34], we will get n equilibrium control laws namely $\boldsymbol{\hat{u}_{t_0}}, \cdots, \boldsymbol{\hat{u}_{t_{n-1}}}$ which we will use at time $t_0, \cdots, t_{n-1}$ respectively. Glueing those individual control laws together, we thus get our equilibrium strategy.

We will now present the above ideas more formally with the following definition from Björk & Murgoci [2010] [35]:

Definition 4.1. *Let us consider a fixed admissible control law* $\boldsymbol{\hat{u}}$ *and do the following construction:*

1. *Fix an arbitrary point* (t_k, x) *where* $k < n$*, and choose an arbitrary control value* $u \in U$*.*
2. *Now define a control law* $\boldsymbol{\bar{u}}$ *on the time set* $t_k, \cdots, t_{n-1}$ *by setting for any* $y \in \mathbb{R}$*,* $\boldsymbol{\bar{u}_{t_i}}(y) = \boldsymbol{\hat{u}_{t_i}}(y)$ *for* $i = k+1, \cdots, n-1$ *and* $\boldsymbol{\bar{u}_{t_k}}(y) = u$*.*

We say that $\boldsymbol{\hat{u}}$ *is a subgame perfect Nash equilibrium strategy if, for every fixed* (t_k, x)*, the following condition holds:*

$$\sup_u J_n(x, \boldsymbol{\bar{u}}) = J_n(x, \boldsymbol{\hat{u}}).$$

If an equilibrium control $\boldsymbol{\hat{u}}$ *exists, we then define the equilibrium value function as:*

$$V_n(x) = J_n(x, \boldsymbol{\hat{u}}).$$

Remark:

1. Ideally, a player starting at (t_k, x) would have liked to maximise $J_k(x, \boldsymbol{u})$ over all possible feedback control laws $\boldsymbol{u}$ [36]. However, in our setting the player at time t_k cannot do so as he is only allowed to choose his control at time t_k i.e. he is only allowed to choose $\boldsymbol{u_k}$ [37]. This is a very important assumption because without it, the problem becomes considerably harder to solve [38].
2. Moreover, for the game-theoretic approach as presented above to make sense, it is important that at t_k, the player playing game k should only be able to choose his action at that point! Had he/she been able to choose the control at, say, t_{k+1}, then we would have two players, namely player k and $k+1$, who would be playing the game at t_{k+1}.

3. The Markovian structure of our model above is guaranteed by the use of feedback control laws. Had this assumption not been made, the problem would have been considerably harder to solve and the results presented in this paper would have been invalidated [51].

We will now move on to the continuous time setting where again we will have different subgames being played. The functional that we wish to maximise this time will be of the form:

$$J(t,x,\boldsymbol{u}) = \mathbb{E}_{t,x}[F(t,x,X_T^{\boldsymbol{u}}] + G(t,x,\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]).$$

In continuous time, we will be playing infinitely many games on $[0,T]$. So, at time $t \in [0,T]$, player t will play a game where he/she can only choose a control u_t. Similar to the discrete time case, we will still look for a subgame perfect Nash equilibrium point [38]. Intuitively and informally, one can think that $\hat{\boldsymbol{u}}$ is the equilibrium control law if it is optimal for player t to use $\hat{\boldsymbol{u}}$ given that all players on $(t,T]$ will use $\hat{\boldsymbol{u}}$ [39].

Notice that in the continuous case, it is slightly more difficult to formalise this intuitive notion of equilibrium control as compared to the discrete setting. This is because player t can only choose the control u_t at time t and is therefore acting on a time set of Lebesgue measure zero [40]. This means that his/her actions will not have any bearing on the dynamics of the controlled process. Thus, we will need another definition for an equilibrium control law, which was provided given by Ekeland, Lazrak and Privu [41] [42], which is as follows:

Definition 4.2. *Consider an admissible control law $\hat{\boldsymbol{u}}$. Then, choose an arbitrary admissible control law $\boldsymbol{u}$ and a fixed real number $h > 0$. Also, we fix an arbitrarily chosen initial point (t,x). Then, for $y \in \mathbb{R}$, we define a control law $\boldsymbol{u_h}$ by $\boldsymbol{u_h}(s,y) := \boldsymbol{u}(s,y)$ for $t \leq s < t+h$ and $\boldsymbol{u_h}(s,y) := \hat{\boldsymbol{u}}(s,y)$ for $t+h \leq s \leq T$.*

Then, if we have:

$$\liminf_{h \to 0} \frac{J(t,x,\hat{\boldsymbol{u}}) - J(t,x,\boldsymbol{u_h})}{h} \geq 0,$$

for all admissible control law $\boldsymbol{u}$, we say that $\hat{\boldsymbol{u}}$ is an equilibrium control law. Moreover, the equilibrium value function is defined by:

$$V(t,x) = J(t,x,\hat{\boldsymbol{u}}).$$

Having defined and explained the concept of equilibrium control, we now face the problem of finding it. To do so, we will make use of the extended HJB equation

[52] – which we will introduce shortly. First, recall that we want to maximise a functional of the form:

$$J(t,x,\boldsymbol{u}) = \mathbb{E}_{t,x}[F(t,x,X_T^{\boldsymbol{u}}] + G(t,x,\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]),$$

where to ensure that we have a non-degenerate problem [52] , we will assume that $\mathbb{E}_{t,x}[|F(t,x,X_T^{\boldsymbol{u}}|] < \infty$ and $\mathbb{E}_{t,x}[|X_T^{\boldsymbol{u}}|] < \infty$ for each $(t,x) \in [0,T] \times \mathbb{R}$ and admissible control law $\boldsymbol{u}$.

Clearly, a functional of the above form is a simplification in the sense that we do not have a running reward/penalty. However, as we have seen previously in section 3, our problem formulation does not have a running reward/penalty. Therefore, the simplification presented above is enough to solve our problem. With this, let us define the extended HJB equation [44] where in the definition below, $\mathcal{A}^u h$ denotes the infinitesimal generator of the function h. Note that unlike the general definition of the infinitesimal generator, ours will include the term $\frac{\partial h}{\partial t}(t,x)$.

Definition 4.3. *[45] The extended HJB system of equations for V, f and g is defined as follows:*

1. *For $0 \le t \le T$, the function V is determined by:*

   $$\sup_{u \in \mathcal{U}} \big\{(\mathcal{A}^u V)(t,x) - (\mathcal{A}^u f)(t,x,t,x) + (\mathcal{A}^u f^{tx})(t,x) - \mathcal{A}^u(G \circ g)(t,x) + (\mathcal{H}^u g)(t,x)\big\} = 0.$$

   with boundary condition being given by:

   $$V(T,x) = F(T,x,x) + G(T,x,x).$$

2. *For every fixed $s \in [0,T]$ and $y \in \mathbb{R}$, the function $(t,x) \to f^{sy}(t,x)$ is defined by:*

   $$\begin{aligned} \mathcal{A}^{\hat{\boldsymbol{u}}} f^{sy}(t,x) &= 0,\ \ 0 \le t \le T. \\ f^{sy}(t,x) &= F(s,y,x). \end{aligned} \tag{4.1}$$

3. *The function $g(t,x)$ is defined by:*

   $$\begin{aligned} \mathcal{A}^{\hat{\boldsymbol{u}}} g(t,x) =&\ \ 0,\ \ 0 \le t \le T. \\ g(T,x) =&\ \ x. \end{aligned} \tag{4.2}$$

In the above definition, we have used the following notation:

$$\begin{aligned} f(t,x,s,y) &= f^{sy}(t,x). \\ (G \circ g)(t,x) &= G(t,x,g(t,x)). \\ \mathcal{H}^u g(t,x) &= G_y(t,x,g(t,x)) \times \mathcal{A}^u g(t,x). \\ G_y(x,y) &= \frac{\partial G}{\partial y}(t,x,y). \end{aligned}$$

Moreover, $\hat{\boldsymbol{u}}$ represents the control law which achieves the supremum in the first equation of the extended HJB system. Also, we have a probabilistic interpretation [53] for the functions f and g in the sense that they are both expectations given by:

$$f^{sy}(t,x) = \mathbb{E}_{t,x}[F(s,y,X_T^{\hat{\boldsymbol{u}}})], \quad 0 \le t \le T. \tag{4.3}$$

$$g(t,x) = \mathbb{E}_{t,x}[X_T^{\hat{\boldsymbol{u}}}], \quad 0 \le t \le T. \tag{4.4}$$

This probabilistic interpretation will be very useful when actually using the extended HJB system. This is because it will allow us to find an appropriate ansatz to solve the equations from the extended HJB. We will see this in more detail in the next section.

Notice that while the reader might reasonably expect $\hat{\boldsymbol{u}}$ to be an equilibrium control law, we actually need the verification theorem to confirm it. First, however, we will need to define a new function space. This is done below.

Definition 4.4. *[54] Consider an arbitrary admissible control law $\boldsymbol{u}$. We say that a function $h : \mathbb{R}_+ \times \mathbb{R}^n \to R$ belongs to the space $L_T^2(X^{\boldsymbol{u}})$ if it satisfies the condition:*

$$\mathbb{E}_{t,x}\left[\int_t^T ||h_x(s,X_s^{\boldsymbol{u}})\sigma^{\boldsymbol{u}}(s,X_s^{\boldsymbol{u}})||^2\, ds\right] < \infty,$$

for every (t,x). In this expression, h_x denotes the gradient of h in the x-variable and $\sigma^{\boldsymbol{u}}$ is the diffusion part of the SDE governing the controlled process $X_t^{\boldsymbol{u}}$.

We can now state the Verification theorem as follows:

Theorem 4.1 (Verification theorem)**.** *[55] Assume that (for all s and y) the functions $V(t,x), f^{sy}(t,x), g(t,x)$ and $\hat{\mathbf{u}}(t,x)$ have the following properties:*

1. *V, f^{sy} and g solve the extended HJB system in Definition 4.3.*
2. *$V(t,x)$ and $g(t,x)$ are smooth in the sense that they are in $C^{1,2}$, and $f(t,x,s,y)$ is in $C^{1,2,1,2}$.*

3. *The function* $\hat{\mathbf{u}}$ *realises the supremum in the V-equation, and* $\hat{\mathbf{u}}$ *is an admissible control law.*

4. V, f^{sy}, g *and* $G \circ g$ *as well as the function* $(t, x) \to f(t, x, t, x)$ *all belong to the space* $L^2_T(X^{\hat{\mathbf{u}}})$.

Then $\hat{\mathbf{u}}$ *is an equilibrium law, and* V *is the corresponding equilibrium value function. Furthermore,* f *and* g *can be interpreted according to equations* (4.3) *and* (4.4).

Therefore, given the regularity constraints given above, we know that if we were to find a solution to the extended HJB system, then we have effectively found an equilibrium control law for a time inconsistent problem with a functional $J(t, x, \boldsymbol{u})$ as defined previously.

While the extended HJB is undeniably helpful in finding the equilibrium control for a time inconsistent problem such as the one we posed in section 3, it is important to note the following points:

1. It is not always possible to solve the extended HJB system for any given function F and G as given in our functional $J(t, x, \boldsymbol{u})$.

2. The equilibrium control law that we get if we manage to solve the extended HJB system need not be unique [46].

3. In the cases where it turns out that the equilibrium control is not unique, the extended HJB does not tell us which equilibrium control is best for our problem.

4. The Verification theorem tells us that if we manage to solve the extended HJB, then we get an equilibrium control law $\hat{\boldsymbol{u}}$ (that is given by the maximising of the V-equation) together with its the corresponding value function V. Conversely, if there exists an equilibrium law $\hat{\boldsymbol{u}}$ with its the corresponding value function V, then we can only **conjecture** that V satisfies the extended HJB system and that $\hat{\boldsymbol{u}}$ is the supremum in the V-equation (given certain regularity conditions) [56]. As of yet, this conjecture has not been proven.

5. Notice that if our functional is in standard form i.e. we have a time consistent problem, then the extended HJB is simplified to the standard HJB equation.

So, we have seen that by solving the extended HJB system, we can find an equilibrium control law for our time inconsistent problem. So, in the next section, we will use this to derive some results on the form of the equilibrium control law.

5 Applications of the extended HJB equation

Having presented the extended HJB system, we will now focus on how to use it on our time-inconsistent problem which has the following form:

$$\max_{u\in\mathcal{U}} \mathbb{E}_{t,x}[(T-t)X(T) - \lambda(t,x)X^2(T)] + \lambda(t,x)\left(\mathbb{E}_{t,x}[X(T)]\right)^2,$$

given that:

$$\begin{aligned} dX(t) &= \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right] dt + [\sigma u_t]\, dW_t \\ &= \tilde{\mu} dt + \tilde{\sigma} dW_t, \end{aligned}$$

where $\tilde{\mu} := \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right]$ and $\tilde{\sigma} := [\sigma u_t]$.

<u>**Note:**</u> Clearly, $\tilde{\mu}$ and $\tilde{\sigma}$ are functions and not constants. However, as a matter of convenience, we will write them without explicitly showing their dependence on u_t (also keep in mind that $u_t = \boldsymbol{u}(t, X_t)$ for some function $\boldsymbol{u}$). Moreover, as a matter of notation, we will denote $X(t)$ by $X_t^{\boldsymbol{u}}$ to show its dependence on the control.

It is clear from our problem above and the extended HJB that the functions F and G are given by:

$$\begin{aligned} & F(t,x,X_T^{\boldsymbol{u}}) = (T-t)X_T^{\boldsymbol{u}} - \lambda(t,x)(X_T^{\boldsymbol{u}})^2 \\ \Longrightarrow \quad & F(t,x,y) = (T-t)y - \lambda(t,x)y^2. \end{aligned}$$

Moreover,

$$\begin{aligned} & G(t,x,\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]) = \lambda(t,x)\left(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]\right)^2 \\ \Longrightarrow \quad & G(t,x,y) = \lambda(t,x)y^2. \end{aligned}$$

Notice that this means that our terminal condition becomes:

$$V(T,x) = F(T,x,x) + G(T,x,x) = (T-T)x - \lambda(t,x)x^2 + \lambda(t,x)x^2 = 0.$$

To solve the problem, we will first start by calculating each of the individual terms inside $\sup_{u\in\mathcal{U}}\{.\}$ in the extended HJB system. In the derivations that will follow, we will use the following shorthand notations:

1. $V_t := \frac{\partial V}{\partial t}(t,x)$, $V_x := \frac{\partial V}{\partial x}(t,x)$, $V_{xx} := \frac{\partial^2 V}{\partial x^2}(t,x)$.
2. $f_t := \frac{\partial f}{\partial t}(t,x,t,x)$, $f_x := \frac{\partial f}{\partial x}(t,x,t,x)$, $f_s := \frac{\partial f}{\partial s}(t,x,t,x)$, $f_y := \frac{\partial f}{\partial y}(t,x,t,x)$, $f_{xx} := \frac{\partial^2 f}{\partial x^2}(t,x,t,x)$, $f_{xy} := \frac{\partial^2 f}{\partial y \partial x}(t,x,t,x)$ and $f_{yy} := \frac{\partial^2 f}{\partial^2 y}(t,x,t,x)$.

3. $g := g(t,x)$, $g_t := \frac{\partial g}{\partial t}(t,x)$, $g_x := \frac{\partial g}{\partial x}(t,x)$, $g_{xx} := \frac{\partial^2 g}{\partial x^2}(t,x)$
4. $G_t = \frac{\partial G}{\partial t}(t,x,g(t,x))$, $G_x = \frac{\partial G}{\partial x}(t,x,g(t,x))$, $G_y = \frac{\partial G}{\partial y}(t,x,g(t,x))$, $G_{xx} = \frac{\partial^2 G}{\partial x^2}(t,x,g(t,x))$, $G_{xy} = \frac{\partial^2 G}{\partial x \partial y}(t,x,g(t,x))$ and $G_{yy} = \frac{\partial^2 G}{\partial y^2}(t,x,g(t,x))$.
5. $\lambda := \lambda(t,x)$, $\lambda_t := \frac{\partial \lambda}{\partial t}(t,x)$, $\lambda_x := \frac{\partial \lambda}{\partial x}(t,x)$ and $\lambda_{xx} := \frac{\partial^2 \lambda}{\partial x^2}(t,x)$.

Moreover, we will also denote by $\mathcal{A}^u h$ the infinitesimal generator of a general function h corresponding to the control u (remember that our definition of the infinitesimal generator will include the term $\frac{\partial h}{\partial t}$).

So, we have:

$$(\mathcal{A}^u V)(t,x) = V_t + \tilde{\mu} V_x + \frac{\tilde{\sigma}^2}{2} V_{xx}.$$

Next we will calculate $(\mathcal{A}^u f)(t,x,t,x)$. But first, we will calculate $df(t,x,t,x)$ using Itô's lemma as follows:

$$\begin{aligned} df(t,x,t,x) &= f_t dt + f_x dX_t^u + f_s dt + f_y dX_t^u + \frac{1}{2}\left[f_{xx}(dX_t^u)^2 + 2f_{xy}(dX_t^u)^2 + f_{yy}(dX_t^u)^2\right] \\ &= \left[f_t + f_s + \tilde{\mu}(f_x + f_y) + \frac{\tilde{\sigma}^2}{2}(f_{xx} + 2f_{xy} + f_{yy})\right] dt + [\tilde{\sigma}(f_x + f_y)] dW_t. \end{aligned}$$

Therefore, we have:

$$(\mathcal{A}^u f)(t,x,t,x) = f_t + f_s + \tilde{\mu}(f_x + f_y) + \frac{\tilde{\sigma}^2}{2}(f_{xx} + 2f_{xy} + f_{yy}).$$

The third term we have to calculate is $(\mathcal{A}^u f^{tx})(t,x)$. Remember that the operator will not act on the upper case index variables as they are viewed as fixed. Then, it can easily calculated and is given by:

$$(\mathcal{A}^u f^{tx})(t,x) = f_t + \tilde{\mu} f_x + \frac{\tilde{\sigma}^2}{2} f_{xx}.$$

Next, we will calculate $\mathcal{A}^u(G \circ g)(t,x) = \mathcal{A}^u G(t,x,g(t,x))$. We will do so by first calculating $dG(t,x,g(t,x))$ using Itô's lemma. So,

$$\begin{aligned} dG(t,x,g(t,x)) = &\frac{1}{2}\left[G_{xx}(dX_t^u)^2 + 2G_{xy}(dX_t^u)(dg(t,X_t^u)) + G_{yy}(dg(t,X_t^u))^2\right] + G_t dt \\ &+ G_x dX_t^u + G_y dg(t,X_t^u) \end{aligned}$$

$$= G_t dt + G_x[\tilde{\mu} dt + \tilde{\sigma} dW_t] + G_y \left\{ g_t dt + g_x[\tilde{\mu} dt + \tilde{\sigma} dW_t] + \frac{\tilde{\sigma}^2}{2} g_{xx} dt \right\}$$

$$+ \frac{\tilde{\sigma}^2}{2} \left[G_{xx} + 2G_{xy} g_x + G_{yy} g_x^2 \right] dt.$$

Therefore,

$$\mathcal{A}^u (G \circ g)(t,x) = G_t + \tilde{\mu} G_x + G_y \left[g_t + \tilde{\mu} g_x + \frac{\tilde{\sigma}^2}{2} g_{xx} \right] + \frac{\tilde{\sigma}^2}{2} \left[G_{xx} + 2g_x G_{xy} + g_x^2 G_{yy} \right].$$

We are now left to calculate $(\mathcal{H}^u g)(t,x)$ which can be easily calculated and is given by:

$$(\mathcal{H}^u g)(t,x) = G_y(t,x,g(t,x)) \times \mathcal{A}^u g(t,x) = G_y \times \left[g_t + \tilde{\mu} g_x + \frac{\tilde{\sigma}^2}{2} g_{xx} \right].$$

Therefore, the function $J^*(t,x,u)$ inside $\sup_{u \in \mathcal{U}}\{.\}$ in the extended HJB equation can be simplified as follows:

$$J^*(t,x,u) = V_t + \tilde{\mu} V_x + \frac{\tilde{\sigma}^2}{2} V_{xx} - f_s - \tilde{\mu} f_y - \frac{\tilde{\sigma}^2}{2}(2f_{xy} + f_{yy}) - G_t - \tilde{\mu} G_x$$
$$- \frac{\tilde{\sigma}^2}{2}[G_{xx} + 2g_x G_{xy} + g_x^2 G_{yy}].$$

Notice that only $\tilde{\mu}$ and $\tilde{\sigma}$ depend on the control u. Therefore, the extended HJB equation can be simplified as follows:

$$V_t - f_s - G_t + \sup_{u \in \mathcal{U}} \left\{ [V_x - f_y - G_x]\tilde{\mu} + \frac{\tilde{\sigma}^2}{2}[V_{xx} - 2f_{xy} - f_{yy} - G_{xx} - 2g_x G_{xy} - g_x^2 G_{yy}] \right\} = 0.$$

Of course, the terminal condition is given by $V(T,x) = 0$.

Moreover, from our problem formulation, we know that the equilibrium value function will be given by:

$$V(t,x) = f(t,x,t,x) + \lambda(t,x) \left[g(t,x) \right]^2.$$

This will therefore allow us to simplify the extended HJB system further and find a system of PDEs which will help us in finding the equilibrium control – if indeed it does exist. So, we will now find the partial derivatives of $V(t,x)$ as follows:

$$V_t = f_t + f_s + [\lambda(t,x)][2gg_t] + g^2 \lambda_t$$
$$= f_t + f_s + 2gg_t\lambda + g^2\lambda_t.$$

Similarly,

$$
\begin{aligned}
V_x =& f_x + f_y + [\lambda(t,x)][2gg_x] + g^2\lambda_x \\
=& f_x + f_y + 2gg_x\lambda + g^2\lambda_x .
\end{aligned}
$$

Finally,

$$
\begin{aligned}
V_{xx} =& [f_{xx} + f_{xy}] + [f_{xy} + f_{yy}] + 2\left\{\lambda[gg_{xx} + g_x^2] + (gg_x)\lambda_x\right\} + g^2\lambda_{xx} + \lambda_x[2gg_x] \\
=& f_{xx} + 2f_{xy} + f_{yy} + 2\lambda[gg_{xx} + g_x^2] + g^2\lambda_{xx} + 4\lambda_x g_x g .
\end{aligned}
$$

Lastly, to derive the results, note that $G(t,x,y) = \lambda(t,x)y^2$. Therefore,

$$
G_t = \lambda_t y^2;\ G_x = \lambda_x y^2;\ G_y = 2\lambda y;\ G_{xx} = \lambda_{xx} y^2;\ G_{xy} = 2\lambda_x y;\ G_{yy} = 2\lambda .
$$

Therefore, from the extended HJB equation derived above and the fact that the partial derivatives of the function G are evaluated at the point $(t,x,g(t,x))$, we note the following:

1.

$$
\begin{aligned}
V_t - f_s - G_t =& f_t + f_s + 2gg_t\lambda + g^2\lambda_t - f_s - g^2\lambda_t \\
=& f_t + 2gg_t\lambda .
\end{aligned}
$$

2.

$$
\begin{aligned}
V_x - f_y - G_x =& f_x + f_y + 2\lambda gg_x + g^2\lambda_x - f_y - \lambda_x g^2 \\
=& f_x + 2\lambda gg_x .
\end{aligned}
$$

3.

$$
\begin{aligned}
V_{xx} - 2f_{xy} - f_{yy} - G_{xx} - 2g_x G_{xy} - g_x^2 G_{yy} =& f_{xx} + 2f_{xy} + f_{yy} + 2\lambda[gg_{xx} + g_x^2] \\
& + 4\lambda_x g_x g + g^2\lambda_{xx} - 2f_{xy} - f_{yy} \\
& - g^2\lambda_{xx} - 4g_x\lambda_x g - 2g_x^2\lambda \\
=& f_{xx} + 2\lambda gg_{xx} .
\end{aligned}
$$

Therefore, our extended HJB equation can be simplified to:

$$
(f_t + 2gg_t\lambda) + \sup_{u\in\mathcal{U}}\left\{\tilde{\mu}[f_x + 2\lambda gg_x] + \frac{\tilde{\sigma}^2}{2}[f_{xx} + 2\lambda gg_{xx}]\right\} = 0,
$$

where the terminal condition is given by $V(T,x)=0$.

Now, as a reminder $\tilde{\mu}=r+(\mu-r)u-\frac{\sigma^2u^2}{2}$ and $\tilde{\sigma}=\sigma u$ (for a fixed $u\in\mathcal{U}$). Moreover, let us define the following functions:

$$\Phi(t,x,t,x):=f_x+2\lambda g g_x \quad \text{and} \quad \Psi(t,x,t,x):=f_{xx}+2\lambda g g_{xx}.$$

In the above, the partial derivatives of f are evaluated at (t,x,t,x) while those of g are evaluated at (t,x). Also, for the calculations that follow, we will use the following shorthand notation: $\Phi:=\Phi(t,x,t,x)$ and $\Psi:=\Psi(t,x,t,x)$.

So, our extended HJB becomes:

$$f_t+2gg_t\lambda+r(f_x+2\lambda g g_x)+\sup_{u\in\mathcal{U}}\left\{(\mu-r)\Phi u-\frac{\sigma^2}{2}\Phi u^2+\frac{\sigma^2}{2}\Psi u^2\right\}=0$$

$$\implies \qquad f_t+rf_x+2\lambda g[g_t+rg_x]+\sup_{u\in\mathcal{U}}\left\{\frac{\sigma^2}{2}(\Psi-\Phi)u^2+(\mu-r)\Phi u\right\}=0. \tag{$\star$}$$

Notice that the expression inside $\sup_u\{.\}$ above is a simple quadratic equation. Also, since we are considering the supremum, we know that the problem will only make sense and be economically feasible if the coefficient of u^2 is negative. Therefore, we require that:

$$\Psi-\Phi<0 \iff f_{xx}-f_x+2\lambda g(g_{xx}-g_x)<0.$$

The equilibrium control is then easily calculated, by first order condition, to be:

$$\hat{\boldsymbol{u}}(t,x)=\frac{(\mu-r)\Phi}{\sigma^2(\Phi-\Psi)}=\frac{(\mu-r)}{\sigma^2}\left[\frac{f_x+2\lambda g_x g}{f_x-f_{xx}+2\lambda g(g_x-g_{xx})}\right].$$

Notice that this is well-defined since the denominator is non-zero as we have assumed that $[f_x-f_{xx}+2\lambda g(g_x-g_{xx})]>0$. It should also be clear that we are no closer to finding a useful equilibrium control because the control above is in feedback form. Indeed, we will need to solve a set of partial differential equations – if indeed they can be solved – to find the solution we are looking for. These PDEs are found from the additional conditions in the extended HJB system.

So, we also have the following sets of equations:

1.

$$\mathcal{A}^{\hat{\boldsymbol{u}}} g(t,x) = 0.$$

$$\implies \qquad g_t + g_x\left[r + (\mu - r)\hat{u} - \frac{\sigma^2}{2}\hat{u}^2\right] + \frac{\sigma^2}{2} g_{xx}\hat{u}^2 = 0$$

$$\implies \qquad g_t + r g_x + (\mu - r) g_x \hat{u} + \frac{\sigma^2}{2}(g_{xx} - g_x)\hat{u}^2 = 0. \qquad (1)$$

The terminal condition is given by $g(t,x) = x$.

2.

$$\mathcal{A}^{\hat{\boldsymbol{u}}} f^{tx}(t,x) = 0$$

$$\implies \qquad f_t + f_x\left[r + (\mu - r)\hat{u} - \frac{\sigma^2}{2}\hat{u}^2\right] + \frac{\sigma^2}{2} f_{xx}\hat{u}^2 = 0$$

$$\implies \qquad f_t + r f_x + (\mu - r) f_x \hat{u} + \frac{\sigma^2}{2}(f_{xx} - f_x)\hat{u}^2 = 0. \qquad (2)$$

The terminal condition is given by $f^{sy}(T,x) = (T-s)x - \lambda(s,y)x^2$.

So, by plugging in the equilibrium control we found, the extended HJB equation $(\star)$ is further simplified to:

$$f_t + r f_x + 2\lambda g[g_t + r g_x] + (\mu - r)\Phi\hat{u} - \frac{\sigma^2}{2}(\Phi - \Psi)\hat{u}^2 = 0.$$

And therefore,

$$f_t + r f_x + 2\lambda g[g_t + r g_x] + (\mu - r)[f_x + 2\lambda g g_x]\hat{u} - \frac{\sigma^2}{2}(f_x + 2\lambda g g_x - f_{xx} - 2\lambda g g_{xx})\hat{u}^2 = 0.$$

The above can be simplified to get:

$$f_t + r f_x + (\mu - r) f_x \hat{u} + \frac{\sigma^2 (f_{xx} - f_x)\hat{u}^2}{2} + 2\lambda g\left[g_t + r g_x + (\mu - r) g_x \hat{u} + \frac{\sigma^2 (g_{xx} - g_x)\hat{u}^2}{2}\right] = 0.$$

Notice therefore that the extended HJB equation is equal to $[2\lambda g \times (1)] + (2)$, where $(1) = (2) = 0$. Therefore, we only need to solve equations (1) and (2) to solve the whole system and find our equilibrium control law.

To summarise, the equilibrium control law $\hat{\boldsymbol{u}}(t,x)$ can be found explicitly as follows:

$$\hat{\boldsymbol{u}}(t,x) = \frac{(\mu - r)}{\sigma^2}\left[\frac{f_x + 2\lambda g_x g}{f_x - f_{xx} + 2\lambda g(g_x - g_{xx})}\right],$$

where $f(t,x,t,x)$ and $g(t,x)$ are solved by the following equations:

$$f_t + rf_x + (\mu - r)f_x\hat{u} + \frac{\sigma^2}{2}(f_{xx} - f_x)\hat{u}^2 = 0,$$

$$g_t + rg_x + (\mu - r)g_x\hat{u} + \frac{\sigma^2}{2}(g_{xx} - g_x)\hat{u}^2 = 0,$$

with boundary conditions being:

$$f(T,x,s,y) = F(s,y,x) = (T-s)x - \lambda(s,y)x^2 \quad \text{and} \quad g(T,x) = x.$$

Then, the equilibrium value function is given by:

$$V(t,x) = f(t,x,t,x) + \lambda(t,x)[g(t,x)]^2.$$

Notice that in general we cannot always solve for the equilibrium control law from the PDEs above. In fact, we do not even know whether the equilibrium control actually exists for the problem above. Worse still, even if an equilibrium control law exists (together with its value function), we can only **conjecture** – **not prove** – that it will solve the above extended HJB system given some regularity condition. All that we know for certain is that **if** somehow we manage to find a solution to the above extended HJB system, i.e. we can find the functions f, g and V such that the above equations hold, then we have an equilibrium control law $\hat{\boldsymbol{u}}(t,x)$ which is given by the supremum of the V-equation in the extended HJB system.

Having derived a general result for our problem, we will now look at more concrete applications by choosing a specific form for the risk aversion index.

5.1 Constant risk aversion index

One of the most obvious choices for the risk aversion index is a constant. Let us denote this constant by $\eta > 0$. Notice that we will still choose a positive risk aversion index to make it explicit that the investor is risk averse.

While it seems quite simplistic to choose $\lambda(t,x) = \eta$, we must realise that in reality and over an investment period $[0,T]$, how risk-averse an investor is generally does not change. Indeed, people who do not like to take risks will most likely be consciously choosing lower-yield investments because they are not very risky and vice versa. Therefore, how risk-averse an investor is boils down to the investment personality of the latter which does not change, if at all, much over time.

On a more mathematical note, it makes sense to choose $\lambda(t,x)$ to be a constant because our problem consists an optimisation problem involving the **annualised log-return** of our investment. What this means is that, by taking log-returns, we are already removing the dependence on our current wealth. This is because a return of \$100 out of a starting capital of \$1,000 corresponds to the same percentage return – of 10% – as a return of \$100,000 out of a starting capital of \$1,000,000. So, our initial wealth level does not matter. Moreover, by further annualising our log-return, we should be indifferent about the length of the investment period $[t,T]$ and thus of our starting point t. For example, a return of 20% over a period 2 years is equivalent to a return of 10% over a 1 year period. Viewed from this perspective then, by choosing to formulate our problem with the annualised log-return, we should in theory be indifferent to what our wealth is and where we start from. Thus, it is reasonable to assume that $\lambda(t,x) = \eta$.

So, our problem will be of the following form:

$$\max_{u \in \mathcal{U}} \; \mathbb{E}_{t,x}[(T-t)X(T) - \eta X^2(T)] + \eta \; \left(\mathbb{E}_{t,x}[X(T)]\right)^2,$$

given that:

$$\begin{aligned} dX(t) &= \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2} \right] dt + [\sigma u_t]\, dW_t \\ &= \tilde{\mu} dt + \tilde{\sigma} dW_t. \end{aligned}$$

So, from the above formulation, it is clear that:

$$\begin{aligned} & F(t,x,X_T^{\boldsymbol{u}}) = (T-t)X_T^{\boldsymbol{u}} - \eta (X_T^{\boldsymbol{u}})^2 \\ \implies \quad & F(t,x,y) = (T-t)y - \eta y^2. \end{aligned}$$

Moreover,

$$\begin{aligned} & G(t,x,\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]) = \eta \left(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]\right)^2 \\ \implies \quad & G(t,x,y) = \eta y^2. \end{aligned}$$

Notice that F does not depend on the variable x; while G depends only on the variable y. Now, from the definition of the extended HJB system and our problem formulation, the following holds:

$$f(t,x,s,y) = \mathbb{E}_{t,x}\left[F(s,y,X_T^{\hat{\boldsymbol{u}}})\right] = \mathbb{E}_{t,x}\left[F(s,X_T^{\hat{\boldsymbol{u}}})\right] = f(t,x,s).$$

Therefore, $f_y(t,x,t,x) = f_{xy}(t,x,t,x) = f_{yy}(t,x,t,x) = 0$.

Moreover, $\lambda_t(t,x) = \lambda_x(t,x) = \lambda_{xx}(t,x) = 0$. Also, $G_t(t,x,y) = G_x(t,x,y) = G_{xy}(t,x,y) = G_{xx}(t,x,y) = 0$ and $G_{yy} = 2\eta$.

So, using the results we derived previously, our extended HJB equation becomes:

$$(V_t - f_s) + \sup_{u\in\mathcal{U}} \left\{ V_x\tilde{\mu} + \frac{\tilde{\sigma}^2}{2}[V_{xx} - 2\eta g_x^2] \right\} = 0$$

$$\implies \quad (V_t - f_s + rV_x) + \sup_{u\in\mathcal{U}} \left\{ (\mu - r)V_x u - \frac{\sigma^2 u^2}{2} V_x + \frac{\sigma^2 u^2}{2}[V_{xx} - 2\eta g_x^2] \right\} = 0$$

$$\implies \quad (V_t - f_s + rV_x) + \sup_{u\in\mathcal{U}} \left\{ [(\mu - r)V_x]u + \frac{\sigma^2}{2}\left[V_{xx} - 2\eta g_x^2 - V_x\right] u^2 \right\} = 0.$$

Notice that for the problem to be well-posed, we need $V_{xx} < 2\eta g_x^2 + V_x$. Assuming that this holds true, then by first-oder conditions, the equilibrium control is given by:

$$\hat{\boldsymbol{u}}(t,x) = \frac{(\mu - r)V_x}{\sigma^2\left[2\eta g_x^2 + V_x - V_{xx}\right]} \ . \tag{$\star\star$}$$

To solve the above, problem, let us suppose that $\hat{\boldsymbol{u}}(t,x)$ is a deterministic function of time only i.e. $\hat{\boldsymbol{u}}(t,x) = \hat{\boldsymbol{u}}(t)$. If that is true, then from section 3, we know that:

$$X^{\hat{\boldsymbol{u}}}(T) \mid_{X(t)=x} \ \sim \ \mathcal{N}\left(x + \int_t^T r + (\mu - r)\hat{\boldsymbol{u}}(s) - \frac{\sigma^2\hat{\boldsymbol{u}}^2(s)}{2} ds, \int_t^T \sigma^2\hat{\boldsymbol{u}}^2(s) ds \right)$$

Therefore, $\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]$ is of the form $\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right] = g(t,x) = x + b(t)$, where $b : [0,T] \to \mathbb{R}$ is a deterministic function given by:

$$b(t) := \int_t^T r + (\mu - r)\hat{\boldsymbol{u}}(s) - \frac{\sigma^2\hat{\boldsymbol{u}}^2(s)}{2} \ ds.$$

Moreover, $\mathrm{Var}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right] = a(t)$, where $a : [0,T] \to \mathbb{R}$ is a deterministic function given by:

$$a(t) := \int_t^T \sigma^2\hat{\boldsymbol{u}}^2(s) \ ds.$$

Now, notice that:

$$\begin{aligned}
V(t,x) &= (T-t)\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right] - \eta\mathrm{Var}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right] \\
&= (T-t) \times [x + b(t)] - \eta a(t) \\
&= (T-t)x + [(T-t)b(t) - \eta a(t)] \\
&= (T-t)x + c(t), \quad \text{where } c(t) := (T-t)b(t) - \eta a(t).
\end{aligned}$$

Therefore, $V_x(t,x) = (T-t), V_{xx}(t,x) = 0$ and $V_t(t,x) = -x + c'(t)$. Moreover, $g_x(t,x) = 1, g_{xx}(t,x) = 0$ and $g_t(t,x) = b'(t)$.

Hence, we find from ($\star\star$) that our equilibrium control is as follows:

$$\hat{\boldsymbol{u}}(t,x) = \frac{(\mu - r)(T-t)}{\sigma^2\left[2\eta + (T-t)\right]} \; .$$

The above represents the proportion of our wealth that we should invest in the risky asset at any point in time. Notice that our equilibrium control depends only on t and not of our log-wealth x. This, of course, is not very realistic in practice. However, notice that if instead of using an ansatz where $\hat{\boldsymbol{u}}(t,x)$ is a deterministic function of time only i.e. we use one which depends on (t,x), then our problem becomes significantly harder to solve. Indeed, finding the above equilibrium control hinges on the fact that if $\hat{\boldsymbol{u}}(t,x)$ is a deterministic function in time, then we know the form of the functions $V(.,.)$ and $g(.,.)$. Without that assumption, it becomes almost impossible to make a reasonable guess for $V(.,.)$ and $g(.,.)$.

While our initial reaction to the equilibrium control is that it is rather simplistic; by conducting some analysis – as we will do shortly – we can, however, assure ourselves that the control is indeed reasonable and makes economic sense.

Recall that $X^{\hat{\boldsymbol{u}}}(T) = \log[Y^{\hat{\boldsymbol{u}}}(T)]$, where $Y^{\hat{\boldsymbol{u}}}(T)$ is our wealth at time T where we have used the equilibrium control. So,

$$\begin{aligned} \mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right] &= \mathbb{E}_{t,x}\left[\log\left[Y^{\hat{\boldsymbol{u}}}(T)\right]\right] \\ &\leq \log\left[\mathbb{E}_{t,x}\left[Y^{\hat{\boldsymbol{u}}}(T)\right]\right], \text{by Jensen's inequality.} \end{aligned}$$

Therefore,

$$\mathbb{E}_{t,x}\left[Y^{\hat{\boldsymbol{u}}}(T)\right] \geq \exp\left\{\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]\right\} = e^{x+b(t)} = ye^{b(t)},$$

where $y = e^x$ is our wealth at time t.

Now, $b(t)$ is found as follows:

$$\begin{aligned} b(t) &= \int_t^T r + (\mu - r)\hat{\boldsymbol{u}}(s) - \frac{\sigma^2 \hat{\boldsymbol{u}}^2(s)}{2} ds \\ &= \int_t^T r + \frac{(\mu - r)^2 (T-s)}{\sigma^2[2\eta + (T-s)]} - \frac{(\mu - r)^2 (T-s)^2}{2\sigma^2[2\eta + (T-s)]^2} ds \\ &= r(T-t) + \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 (T-t) - \left(\frac{\mu - r}{\sigma}\right)^2 \frac{\eta(T-t)}{2\eta + T - t}. \end{aligned}$$

The last line follows by using standard integration techniques and partial fractions.

Therefore, we have:

$$\begin{aligned}\mathbb{E}_{t,x}\left[Y^{\hat{\boldsymbol{u}}}(T)\right] &\geq y\exp\left\{r(T-t)+\frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2(T-t)-\left(\frac{\mu-r}{\sigma}\right)^2\frac{\eta(T-t)}{2\eta+T-t}\right\}\\ &> ye^{r(T-t)},\ \text{for } \eta>0 \text{ and } t<T.\end{aligned}$$

The last line follows from the fact that the RHS of the inequality above is decreasing in η, which can easily be proved using standard calculus.

Moreover, we can calculate $\mathrm{Var}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]=a(t)$ as follows:

$$\begin{aligned}a(t) &= \int_t^T \sigma^2\hat{\boldsymbol{u}}^2(s)ds\\ &= \left(\frac{\mu-r}{\sigma}\right)^2\int_t^T \frac{(T-s)^2}{[2\eta+(T-s)]^2}ds\\ &= \left(\frac{\mu-r}{\sigma}\right)^2\left\{(T-t)+4\eta\log\left(\frac{2\eta}{2\eta+T-t}\right)+\frac{2\eta(T-t)}{2\eta+T-t}\right\}.\end{aligned}$$

Now, since $\lim_{\eta\to\infty} b(t)=r(T-t)$, we have that $\lim_{\eta\to\infty}\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]=x+r(T-t)$. Moreover, $\lim_{\eta\to\infty}\mathrm{Var}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]=\lim_{\eta\to\infty}a(t)=0$, using L'Hopital's rule. Therefore, we can conclude that $X^{\hat{\boldsymbol{u}}}(T)\xrightarrow{p} x+r(T-t)$ as $\eta\to\infty$. So, using the Continuous Mapping Theorem [47], we have that:

$$Y^{\hat{\boldsymbol{u}}}(T)=e^{X^{\hat{\boldsymbol{u}}}(T)}\xrightarrow{p} e^{x+r(T-t)}=ye^{r(T-t)}.$$

What this says is that if we are become increasingly risk averse, the growth of our wealth will converge towards the risk-free rate (in probability). This should also be obvious from the fact that the equilibrium control tends to zero, almost surely, as $\eta\to\infty$, which means that we will not invest at all in the risky asset (except on a set of measure zero). Moreover, $\mathbb{E}_{t,x}\left[X^{\hat{\boldsymbol{u}}}(T)\right]$ is decreasing in η. This just says that to get a greater expected log-return, we need to be less risk-averse – we need to take more risk. Thus, from the above, it is clear that the properties of our equilibrium control does indeed make economic sense.

Moreover, we find that our equilibrium control, while quite simplistic, does indeed provide us with a terminal **expected** return greater than the risk-free rate.

Therefore, the use of such a strategy can surely be justified.

<u>Note:</u> Notice that while we have derived a general result for the equilibrium control, it is much simpler in the above example to proceed using the unsimplified version of the extended HJB equation. This is because in the above example, we could find an appropriate ansatz for V and g which greatly simplifies the problem. This, of course, is not always possible for other problems.

5.2 Modified Basak-Chabakauri problem

Let us now consider a different function for $\lambda(t, x)$. The function we will use is $\lambda(t, x) = \eta(T - t), \eta > 0$. Notice that our choice of $\lambda(t, x)$, which as we discussed before represents how risk averse we are, is decreasing in t. The reason for this choice is that as we get closer and closer to the end of our investment period, we might be more willing to invest in a risky asset as we do not expect the risky asset to move drastically in such a short period of time. So, even if we face a downside move in the risky asset, we would not expect it to be very substantial. Thus, we might be tempted to take on more risk and to indicate this behaviour, our risk-aversion must decrease as we move forward in time – hence explaining our choice for $\lambda(t, x)$.

At this point, notice that it is easier to simplify our objective problem rather than to calculate all the derivatives as presented in the general case. Of course, this trick does not work in all cases and therefore, the problem can be significantly harder, or even unsolvable, for some other function $\lambda(t, x)$.

So, choosing $\lambda(t, x) = \eta(T - t)$, then our problem becomes:

$$\max_{u \in \mathcal{U}} \mathbb{E}_{t,x}[(T - t)X(T) - \eta(T - t)X^2(T)] + \eta(T - t)\left(\mathbb{E}_{t,x}[X(T)]\right)^2.$$

This, of course, is mathematically equivalent to the following problem:

$$\max_{u \in \mathcal{U}} \mathbb{E}_{t,x}[X(T) - \eta X^2(T)] + \eta\left(\mathbb{E}_{t,x}[X(T)]\right)^2,$$

with the dynamics of the controlled process being given by:

$$\begin{aligned} dX(t) &= \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right] dt + [\sigma u_t]\, dW_t \\ &= \tilde{\mu} dt + \tilde{\sigma} dW_t, \end{aligned}$$

where $\tilde{\mu} = r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}$ and $\tilde{\sigma} = \sigma u_t$.

Notice that with our choice of $\lambda(t,x) = \eta(T-t)$, our problems closely resembles the Basak-Chabakauri [48] problem. However, instead of maximising the expected terminal wealth given a certain risk level; our problem reduces to maximising the expected terminal log-wealth given a certain risk level. Moreover, in our problem, the dynamics of the controlled process differs from the Basak-Chabakauri one.

Proceeding with our problem, we have $F(t,x,X_T^{\boldsymbol{u}}) = X_T^{\boldsymbol{u}} - \eta\,(X_T^{\boldsymbol{u}})^2$. Clearly, the function only depends on the terminal value. Therefore, $F(t,x,X_T^{\boldsymbol{u}}) = F(X_T^{\boldsymbol{u}}) = X_T^{\boldsymbol{u}} - \eta X_T^{\boldsymbol{u}}$. On the other hand, $G(t,x,\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]) = G(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}]) = \eta(\mathbb{E}_{t,x}[X_T^{\boldsymbol{u}}])^2$.

Moreover, $f^{sy}(t,x) = \mathbb{E}_{t,x}[F(s,y,X_T^{\hat{\boldsymbol{u}}})] = \mathbb{E}_{t,x}[F(X_T^{\hat{\boldsymbol{u}}})]$ in our case. Therefore, $f^{sy}(t,x) = f(t,x,s,y) = f(t,x)$. Therefore, in our extended HJB equation $\mathcal{A}^u f^{tx}(t,x) - \mathcal{A}^u f(t,x) = 0$. Thus, our extended HJB equation can be simplified to:

$$\sup_{u\in\mathcal{U}} \{(\mathcal{A}^u V)(t,x) - \mathcal{A}^u(G\circ g)(t,x) + (\mathcal{H}^u g)(t,x)\} = 0.$$

with the boundary condition being given by:

$$V(T,x) = F(T,x,x) + G(T,x,x) = F(x) + G(x) = x - \eta x^2 + \eta x^2 = x.$$

We will now calculate all the terms inside the extended HJB equation. First,

$$(\mathcal{A}^u V)(t,x) = V_t + \tilde{\mu} V_x + \frac{\tilde{\sigma}^2}{2} V_{xx}.$$

As for the second term of the extended HJB, we have:

$$\begin{aligned} d[G\circ g(t,x)] &= d[\eta g^2] \\ &= \eta\left\{2gg_t dt + 2gg_x[\tilde{\mu}dt + \tilde{\sigma}dW_t] + [gg_{xx} + g_x^2]\tilde{\sigma}^2 dt\right\} \\ \implies \mathcal{A}^u G\circ g(t,x) &= \eta \times \left\{2gg_t + 2\tilde{\mu}gg_x + \tilde{\sigma}^2[gg_{xx} + g_x^2]\right\}. \end{aligned}$$

For the third term, we can calculate it as follows:

$$\begin{aligned} \mathcal{H}^u g(t,x) &= G_y(g(t,x)) \times \mathcal{A}^u g(t,x) \\ &= 2\eta g \times \left\{ g_t + \tilde{\mu} g_x + \frac{\tilde{\sigma}^2}{2} g_{xx} \right\}. \end{aligned}$$

Therefore, the extended HJB equation can be simplified to as:

$$\sup_{u\in\mathcal{U}}\left\{V_t+\tilde{\mu}V_x+\frac{\tilde{\sigma}^2}{2}V_{xx}-\eta\tilde{\sigma}^2g_x^2\right\}=0$$

$$\Longrightarrow \qquad \sup_{u\in\mathcal{U}}\left\{V_t+\tilde{\mu}V_x+\frac{\tilde{\sigma}^2}{2}[V_{xx}-2\eta g_x^2]\right\}=0$$

$$\Longrightarrow \qquad \sup_{u\in\mathcal{U}}\left\{V_t+V_x\left[r+(\mu-r)u-\frac{\sigma^2u^2}{2}\right]+\frac{\sigma^2u^2}{2}[V_{xx}-2\eta g_x^2]\right\}=0$$

$$\Longrightarrow \qquad V_t+rV_x+\sup_{u\in\mathcal{U}}\left\{(\mu-r)V_xu+\frac{\sigma^2}{2}[V_{xx}-V_x-2\eta g_x^2]u^2\right\}=0. \tag{1}$$

From equation (1) above, we see that we have to find the maximum of a simple quadratic function. So, for the above problem to make sense, we need $V_{xx}-V_x-2\eta g_x^2<0 \iff V_{xx}<V_x+2\eta g_x^2$.

The optimisation problem yields an equilibrium control in feedback form. We will denote it by $\hat{\boldsymbol{u}}(t)$, where again we assume the control is deterministic in t. So,

$$\hat{\boldsymbol{u}}(t)=\frac{(\mu-r)V_x}{\sigma^2[V_x+2\eta g_x^2-V_{xx}]} \quad .$$

Plugging the above equilibrium control into our extended HJB equation, we get:

$$V_t+rV_x+\frac{(\mu-r)^2V_x^2}{\sigma^2[V_x+2\eta g_x^2-V_{xx}]}-\left(\frac{\sigma^2(V_x+2\eta g_x^2-V_{xx})\times(\mu-r)^2V_x^2}{2\times\sigma^4(V_x+2\eta g_x^2-V_{xx})^2}\right) = 0$$

$$\Longrightarrow \qquad V_t+rV_x+\frac{(\mu-r)^2V_x^2}{\sigma^2(V_x+2\eta g_x^2-V_{xx})}-\frac{(\mu-r)^2V_x^2}{2\sigma^2(V_x+2\eta g_x^2-V_{xx})} = 0$$

$$\Longrightarrow \qquad V_t+rV_x+\frac{(\mu-r)^2V_x^2}{2\sigma^2(V_x+2\eta g_x^2-V_{xx})} = 0.$$

Notice that $V(T,x)=x$. Therefore, we will use the ansatz $V(t,x)=A(t)x+B(t)$, with $A(T)=1$ and $B(T)=0$. So, $V_t(t,x)=A'(t)x+B'(t)$; $V_x(t,x)=A(t)$ and $V_{xx}(t,x)=0$. Moreover, we also know that $g(T,x)=x$. Therefore, we will use the ansatz $g(t,x)=a(t)x+b(t)$, where $a(T)=1$ and $b(T)=0$.

Remark: The ansatz proposed above follows immediately from the fact that $\hat{\boldsymbol{u}}(t)$ is a deterministic function of time only – which is the approach we used in the previous example. The intention here, however, is to make the reader realise that in a general case, we do not necessarily know the form of either V or g, and therefore have to resort to guesswork to further simplify the PDEs involved. In this case, we guess the form of V or g based on their terminal value.

So, we can simplify our extended HJB equation as follows:

$$A'(t)x + B'(t) + rA(t) + \frac{(\mu - r)^2 A(t)}{2\sigma^2[A(t) + 2\eta a^2(t)]} = 0.$$

Notice that since the above holds for every $x \in \mathbb{R}$, then the following equations must hold:

$$\begin{aligned} A'(t) &= \ 0. \\ B'(t) + rA(t) + \frac{(\mu - r)^2 A(t)}{2\sigma^2[A(t) + 2\eta a^2(t)]} &= \ 0. \end{aligned}$$

By simple calculus, $A'(t) = 0$ and $A(T) = 1$ implies $A(t) = 1$.

The second equation therefore becomes:

$$B'(t) + r + \frac{(\mu - r)^2}{2\sigma^2[1 + 2\eta a^2(t)]} = \ 0, \tag{$\dagger$}$$

and our equilibrium control can be further simplified to:

$$\hat{\boldsymbol{u}}(t) = \frac{(\mu - r)}{\sigma^2[1 + 2\eta a^2(t)]}.$$

We will now use the other conditions of the extended HJB system. So, first of all, we have $\mathcal{A}^{\hat{\boldsymbol{u}}} g(t,x) = 0$. Let us first denote $\hat{\mu} := \left[r + (\mu - r)\hat{\boldsymbol{u}}(t) - \frac{\sigma^2 \hat{\boldsymbol{u}}^2(t)}{2}\right]$ and $\hat{\sigma} := [\sigma \hat{\boldsymbol{u}}(t)]$.

Therefore,

$$\begin{aligned} & & \mathcal{A}^{\hat{\boldsymbol{u}}} g(t,x) = 0 \\ \implies & & g_t + \hat{\mu} g_x + \frac{\hat{\sigma}^2}{2} g_{xx} = 0 \\ \implies & & a'(t)x + b'(t) + a(t)\left[r + \frac{(\mu - r)^2}{\sigma^2[1 + 2\eta a^2(t)]} - \frac{\sigma^2(\mu - r)^2}{2\sigma^4[1 + 2\eta a^2(t)]^2}\right] = 0 \\ \implies & & a'(t)x + b'(t) + a(t)\left[r + \frac{(\mu - r)^2}{\sigma^2[1 + 2\eta a^2(t)]} - \frac{(\mu - r)^2}{2\sigma^2[1 + 2\eta a^2(t)]^2}\right] = 0. \end{aligned}$$

Again, the above holds for every $x \in \mathbb{R}_{\geq 0}$. So, the following equations must hold:

$$a'(t) = \ 0.$$

$$b'(t) + a(t)\left[r + \frac{(\mu - r)^2}{\sigma^2[1 + 2\eta a^2(t)]} - \frac{(\mu - r)^2}{2\sigma^2[1 + 2\eta a^2(t)]^2}\right] = \ 0. \qquad (\ddagger)$$

We know that $a(T) = 1$, and this together with the ODE $a'(t) = 0$ implies that $a(t) = 1$.

Now, we can simplify the equation (†) as follows:

$$B'(t) + r + \frac{(\mu - r)^2}{2\sigma^2[1 + 2\eta a^2(t)]} = \ 0$$

$$\implies \qquad B'(t) + r + \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)} = \ 0$$

$$\implies \qquad B(t) - B(T) = -\left[r + \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)}\right](t - T)$$

$$\implies \qquad B(t) = \left[r + \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)}\right](T - t), \text{ as } B(T) = 0.$$

Moreover, from (‡) and terminal condition $b(T) = 0$, we have:

$$b'(t) + a(t)\left[r + \frac{(\mu - r)^2}{\sigma^2[1 + 2\eta a^2(t)]} - \frac{(\mu - r)^2}{2\sigma^2[1 + 2\eta a^2(t)]^2}\right] = \ 0$$

$$\implies \qquad b'(t) + \left[r + \frac{(\mu - r)^2}{\sigma^2(1 + 2\eta)} - \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)^2}\right] = \ 0.$$

Therefore,

$$b(t) = \left[r + \frac{(\mu - r)^2}{\sigma^2(1 + 2\eta)} - \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)^2}\right](T - t).$$

Finally, combining the results we have obtained above, we get:

$$V(t, x) = x + \left[r + \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)}\right](T - t).$$

$$g(t, x) = x + \left[r + \frac{(\mu - r)^2}{\sigma^2(1 + 2\eta)} - \frac{(\mu - r)^2}{2\sigma^2(1 + 2\eta)^2}\right](T - t).$$

$$\hat{\boldsymbol{u}}(t) = \frac{(\mu - r)}{\sigma^2[1 + 2\eta]}.$$

This time our equilibrium control is a positive constant. So, we will hold a constant proportion of our wealth in the risky asset. Moreover, this is a long-only portfolio. Again, we find that our control is simplistic; however, preliminary analysis shows that it does indeed make sense. For example, our equilibrium control decreases as we increases η – meaning that we will hold less of the risky asset if we become more risk averse.

Furthermore, notice that $g(t,x) = \mathbb{E}_{t,x}[X_T^{\hat{\boldsymbol{u}}}] = x + \left[r + \frac{(\mu-r)^2}{\sigma^2(1+2\eta)} - \frac{(\mu-r)^2}{2\sigma^2(1+2\eta)^2} \right](T-t)$ is a decreasing function of η. So, $g(t,x) > x + r(T-t)$. Had we only invested into the risk-free asset, then our expected log-wealth would have been only $x + r(T-t)$. Hence, because our expected terminal log-wealth is greater than what it would have been had we invested only into the risk-free asset, it makes sense to use the equilibrium control. Of course, this argument holds because we know that the equilibrium control has taken into account our risk preferences.

Remark: Notice that by using a more "complicated" risk-aversion index (i.e. $\lambda(t,x) = \eta(T-t)$ instead of $\lambda(t,x) = \eta$), we get a much simpler equilibrium control. In fact, one can argue that in the case where $\lambda(t,x) = \eta$, the equilibrium control we get is more meaningful – given its dependence on time – than the one above where we simply hold a constant proportion of wealth in stocks.

6 Equivalent time consistent formulation

One interesting property of a time-inconsistent problem in the present framework presented is that there always exists an equivalent time consistent problem [49]. While this does not have many practical applications, we will still provide the equivalent time-consistent formulations to our examples above for the sake of completeness.

Remember that the normal HJB equation is of the form:

$$\sup_{u \in \mathcal{U}} \left\{ V_t + \tilde{\mu} V_x + \frac{\tilde{\sigma}^2}{2} V_{xx} + C(t,x,u) \right\} = 0, \quad V(T,x) = F(x),$$

which is used to solve a problem of the form:

$$V(t,x) := \max_{u \in \mathcal{U}} \ \mathbb{E}_{t,x} \left[\int_t^T C(s, X_s, u_s) ds + F(X_T^u) \right].$$

Moreover, the controlled process is governed by the SDE:

$$dX_t^u = \mu(t, X_t^u, u_t) dt + \sigma(t, X_t^u, u_t) dW_t, \quad X_t = x.$$

In the above, we have used the following notation: $\tilde{\mu} := \mu(t, X_t^u, u_t)$ and $\tilde{\sigma} := \sigma(t, X_t^u, u_t)$. In our case, $\tilde{\mu} = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right]$ and $\tilde{\sigma} = [\sigma u_t]$.

Notice that in the case where $\lambda(t,x) = \eta$, our extended HJB equation is given by:

$$\sup_{u \in \mathcal{U}} \left\{ V_t + V_x \tilde{\mu} + \frac{\tilde{\sigma}^2}{2} V_{xx} - \tilde{\sigma}^2 \eta g_x^2 - f_s \right\} = 0, \quad V(T,x) = 0.$$

Since $-\tilde{\sigma}^2 \eta g_x^2(t,x) - f_s(t,x,t) = -\sigma^2 \eta u_t^2 g_x^2(t,x) - f_s(t,x,t)$ depends only on (t,x,u), then it is clear that if we define $C(t,x,u) := -\sigma^2 \eta u_t^2 g_x^2(t,x) - f_s(t,x,t) = -\sigma^2 \eta u_t^2 - f_s(t,x,t) = -\sigma^2 \eta u_t^2 - x - b(t)$ and $F(X_T^u) = 0$, we get the equivalent time-consistent problem.

Remark: We have used the same notations and solutions as presented when solving our problem when $\lambda(t,x) = \eta$. So, refer to section 5.1 for the definitions of $b(t)$ and g_x.

So, the time-consistent problem when $\lambda(t,x) = \eta$ is given by:

$$V(t,x) := \max_{u \in \mathcal{U}} \mathbb{E}_{t,x} \left[-\int_t^T \sigma^2 \eta u_s^2 + X_s^u + b(s) ds \right], \quad V(T,x) = 0.$$

Moreover, the controlled process is governed by the SDE given by:

$$dX_t^u = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2} \right] dt + [\sigma u_t] \, dW_t, \quad X_t = x.$$

Similarly, for the case where $\lambda(t,x) = \eta(T-t)$, the extended HJB is given by:

$$\sup_{u \in \mathcal{U}} \left\{ V_t + \tilde{\mu} V_x + \frac{\tilde{\sigma}^2}{2} V_{xx} - \eta \tilde{\sigma}^2 g_x^2 \right\} = 0, \quad V(T,x) = x.$$

Since $-\eta \tilde{\sigma}^2 g_x^2 = -\eta \sigma^2 u_t^2 g_x^2(t,x)$ depends only on (t,x,u), then it is clear that if we define $C(t,x,u) := -\eta \sigma^2 u_t^2 g_x^2(t,x) = -\eta \sigma^2 u_t^2$, we get the equivalent problem.

Therefore, we can define the equivalent time-consistent problem as:

$$\max_{u \in \mathcal{U}} \mathbb{E}_{t,x} \left[X_T^u - \int_t^T \eta \sigma^2 u_s^2 \, ds \right].$$

given that $dX_t^u = \left[r + (\mu - r)u_t - \frac{\sigma^2 u_t^2}{2}\right] dt + [\sigma u_t]\, dW_t, \quad X_t = x.$

Remark: Notice that we first need to solve the extended HJB system fully to be able to formulate its equivalent time consistent problem.

7 Conclusion

The Markowitz formulation of finding the optimal weights which maximise the final expected return of a portfolio given a pre-determined level of risk is without doubt a brilliant idea. However, it is also fraught with several weaknesses. To begin with, the problem is static. This means that once we find the optimal weights at time $t = 0$ to build our portfolio, we can no longer adjust the weights again in the future. This is clearly a very big weakness since in practice, investment managers consistently adjust the weights of the securities in their portfolios.

So, to remediate this issue, we have formulated our problem in a continuous and dynamic setting whereby at every point in time, we will be allowed to rebalance our portfolio. Once we move to the continuous time setting, however, it becomes crucial to change our definition of returns as used in the MPT – which is why we have used log-returns. Additionally, we will annualise those returns to enable comparison over different investment periods. Our objective, then, much like in the standard Markowitz formulation will be to maximise the final expected annualised return of our portfolio given a certain risk level at every point in time.

This approach, however, has a drawback in the sense that it is a time-inconsistent problem. What this means is that if we were at time t with log-wealth x and found the optimal weights to hold at every subsequent time and wealth, then as we move along, we will actually find that those weights will actually no longer be optimal. Clearly then, the standard concept of optimality becomes very problematic.

This is why we have used a different approach and introduced the concept of equilibrium control. Broadly speaking, this means that we view our problem as a series of games being played at every point in time, and where the player at time t can choose the weights of the risky asset that he can hold. The key then to solve the problem is, loosely speaking, to find the best control now given that the players after us will choose a control to maximise their own objective function.

We have seen that by using this approach, we can indeed get a sensible equilibrium control. However, it is also obvious that because our equilibrium control is a function of time only, it is too simplistic. Of course, we would have preferred our equilibrium control to depend on our wealth also, but as we have discussed, such a control – although it might exist – is incredibly hard to find.

On the bright side, however, the equilibrium control law we found in our example gives us an expected final return greater than the risk free rate. This, combined with the fact that the equilibrium control has taken into consideration our changing risk preferences in the future, enables us to conclude that it does make sense to use them when managing our portfolio. Moreover, we have also seen that the equilibrium control laws we derived make economic sense – they behave in a way that we would expect them to based on economic theory.

One final point to bear in mind is that an equilibrium control law may not be unique. Indeed, there may be other functions which satisfy the extended HJB system but which will not yield the same equilibrium control. So, uniqueness of the control is certainly an issue. With non-uniqueness of an equilibrium control, we obviously would want to know which equilibrium control is best and makes the most economic sense. However, this turns out to be an incredibly complex problem for which no meaningful results have been provided so far. Unfortunately, the same is also true when it comes to proving that a solution to the extended HJB system actually exists.

References

[1] Harry Markowitz. Portfolio Selection. The Journal of Finance, Vol. 7, No. 1. (Mar., 1952).

[2] Hans Wolfgang Brachinger. From Variance to Value at Risk: A Unified Perspective on Standardized Risk Measures. Seminar für Statistik, Wirtschafts- und Sozialwissenschaftliche Fakultät, Universität Fribourg. 1999.

[3] Enrique Sentana. Mean-Variance portfolio allocation with a value at risk constraint. Revista de ECONOMÍA FINANCIERA. 2003.

[4] Peter Byrne and Stephen Lee. Different Risk Measures: Different Portfolio Compositions? Centre for Real Estate Research. The University of Reading Business School. 2004.

[5] Nils H. Hakansson, Capital Growth and the Mean-Variance Approach to Portfolio Selection, Journal of Financial and Quantitative Analysis, 1971, vol. 6, issue 01, 517-557.

[6] Paul Samuelson, Lifetime Portfolio Selection by Dynamic Stochastic Programming, The Review of Economics and Statistics, 1969, vol. 51, issue 3, 239-46.

[7] Robert Merton, Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case, The Review of Economics and Statistics, 1969, vol. 51, issue 3, 247-57.

[8] Isabelle Bajeux-Besnainou and Roland Portait, Dynamic Asset Allocation in a Mean-Variance Framework, Management Science, 44(11-part-2):S79?S95, 1998.

[9] Ekeland and Lazrak, Being serious about non-commitment: subgame perfect equilibrium in continuous time, 2006.

[10] Ekeland and Privu, Investment and consumption without commitment, 2007.

[11] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[12] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[13] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[14] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[15] Dr. Kempthorne, Portfolio Theory, MIT, 2013.

[16] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[17] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[18] Daniel Wilson-Nunn, An Introduction to Mathematical Finance, The University of Warwick, 2014.

[19] Simon, Carl and Lawrence Blume. Mathematics for Economists (Student ed.). Viva Norton. p. 363.

[20] Dr. Kempthorne, Portfolio Theory, MIT, 2013.

[21] Protter, Philip E., Stochastic Integration and Differential Equations (2nd ed.), Springer, 2004

[22] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[23] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[24] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[25] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[26] Chandrasekhar Karnam, Jin Ma and Jianfeng Zhang, Dynamic Approaches for Some Time Inconsistent Problems, 2016.

[27] Chandrasekhar Karnam, Jin Ma and Jianfeng Zhang, Dynamic Approaches for Some Time Inconsistent Problems, 2016.

[28] Bellman, R.E. 1957. Dynamic Programming. Princeton University Press, Princeton, NJ.

[29] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[30] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[31] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[32] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[33] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[34] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[35] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[36] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[37] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[38] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[39] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[40] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[41] Ekeland and Lazrak, Being serious about non-commitment: subgame perfect equilibrium in continuous time, 2006.

[42] Ekeland and Privu, Investment and consumption without commitment, 2007.

[43] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[44] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[45] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[46] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[47] Billingsley, Patrick. Convergence of Probability Measures. John Wiley & Sons. 1969.

[48] Suleyman Basak and Georgy Chabakauri, Dynamic Mean-Variance Asset Allocation, London Business School; Centre for Economic Policy Research (CEPR), 2007.

[49] Björk and Murgoci, A General Theory of Markovian Time Inconsistent Stochastic Control Problems, 2010.

[50] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.

[51] Björk and Khapko, Time Inconsistent Stochastic Control in Continuous Time: Theory and Examples, 2016.

[52] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.

[53] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.

[54] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.

[55] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.

[56] Björk, Khapko and Murgoci, On time-inconsistent stochastic control in continuous time, 2017.