OPTIMAL MULTIPLE STOPPING: THEORY AND APPLICATIONS

by

Gary Wayne Crosby

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2017

Approved by:

____

Dr. Mingxin Xu

____

Dr. Jaya Bishwal

____

Dr. Oleg Safronov

____

Dr. Dmitry Shapiro

ABSTRACT

GARY WAYNE CROSBY. Optimal multiple stopping: theory and applications.
(Under the direction of Dr. Mingxin Xu)

The classical secretary problem was an optimal selection thought experiment for a decision process where candidates with independent and identically distributed values to the observer appear in a random order and the observer must attempt to choose the best candidate with limited knowledge of the overall system. For each observation (interview) the observer must choose to either permanently dismiss the candidate or hire the candidate without knowing any information on the remaining candidates beyond their distribution. We sought to extend this problem into one of sequential events where we examine continuous payoff processes of a function of continuous stochastic processes. With the classical problem the goal was to maximize the probability of a desired occurrence. Here, we are interested in maximizing the expectation of integrated functions of stochastic processes. Further, our problem is not one of observing and discarding, but rather one where we have a job or activity that must remain filled by some candidate for as long as it is profitable to do so. After posing the basic problem we then examine several specific cases with a single stochastic process providing explicit solutions in the infinite horizon using PDE and change of numeraire approaches and providing limited solutions and Monte Carlo simulations in the finite horizon, and finally we examine the two process switching case in both finite and infinite horizon.

As our general model will include supremum of the expected value of integrated stochastic processes, we will make use of techniques that allow us to rewrite the problem into a form without integrals. In the infinite horizon cases we will make use of a method developed by Cissé, Patie, and Tanré [4] and change of numeraire before taking standard PDE-approaches to solving the resulting variational inequality. In finite

horizon cases, we will use a portfolio method developed by Večeř [16] to rewrite the problem. For optimal stopping problems with integrated processes, the resulting PDE before transforming the problem will have two dimensions for every integral process. The strength of Večeř's approach is that it effectively reduces the dimension of the problem, although it does introduce a portfolio term that has complicated dynamics. This makes the approach unsuited to finding closed form solutions in general, but it does offer advantages in Monte Carlo simulations. In a general model, a numerical simulation of the integrals will require simulations of the variables themselves, then integration, then examinations over all possible (discrete) stopping times in the limits of integration. With the portfolio approach, however, we only need to simulate each variable and the portfolio itself.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In this work, we examine extensions to the Secretary Problem into something more resembling the Asian financial option. We begin by considering single candidate formulations and try to extend this work to multiple candidates.

1.1 Classical Secretary Problem

The classical Secretary Problem is formulated as follows [7]: There are N total candidates, each with an independent and identically distributed value to the interviewer with candidates taking distinct values (no ties) according to the underlying probability distribution. That is, after all interviews are completed, the candidates can be ranked from best to worst. Further, they arrive in a random order with all possible orderings equally likely. These candidates are interviewed as they arrive. The interviewer must decide to hire or permanently dismiss on the candidate and interview the next candidate. If the interviewer reaches the end of the sequence, he must hire the final candidate. The interviewer, naturally, wants to hire the best but they must make their choice of hiring or dismissing with only the information gathered by the candidates that have already appeared, including the one currently being interviewed. If the interviewer chooses to pass on the candidate, they cannot be recalled and hired later. The optimal strategy to this problem, the one that maximizes the probability of selecting the best candidate, is to observe but preemptively dismiss the first $K < N$ candidates and hire the first candidate among the $K + 1$ remaining with a value that exceeds those candidates that have already been seen. Notice that if the best candidate is among the first K, the best candidate is not only not hired, but that only the final candidate can be hired regardless of that candidate's overall rank. With this problem structured thus, the optimal strategy is to choose K to be N/e, or approximately 37% of the total candidates, and such a K will maximize the probability of choosing the best candidate. In the infinite candidate case, this converges to $1/e$, or again approximately 37% [6].

The purpose of the Secretary Problem is not to develop a reasonable or useful hiring strategy, but it is rather a thought experiment to develop a strategy to maximize the value to a decision maker who must either choose to accept or permanently reject without perfect information. It is a discrete-time optimal selection problem and all candidates take their value from a common underlying distribution.

1.2 Extensions to the Classical Problem

We sought to investigate a problem that is structurally similar to the Secretary Problem, but where the N candidates take their value from a stochastic process $X_{i,t}$. The candidates' values will continue to fluctuate in accordance with their individual underlying stochastic processes even after their interview. We surmise that there will be threshold strategies for the hiring of candidates, as well as the potential firing or switching between successive candidates. Further, we wanted to investigate a problem with a hiring strategy where there was always one candidate working. As is the case in the classical secretary problem, our examinations were not necessarily intended for use as a potential hiring strategy. Rather, we sought to examine a value process that is the sum of continuous payoff different stochastic processes that had no overlapping activity. Possible applications include valuation of real options in which there are competing processes and only one can be active at any time. Another possible application is in the valuation of American-type Asian exchange options. The general form of such a model, in the case of n such stochastic processes each with initial value

x_i, $i = 1, \ldots, n$ would take the form

$$\Phi(x_1, x_2, \ldots, x_n) = \sup_{\substack{\tau_1,\ldots,\tau_n \\ \sigma_1,\ldots,\sigma_n}} \mathbb{E}_{x_1,\ldots,x_n} \left\{ \sum_{i=1}^{n} \int_{\tau_i}^{\sigma_i} e^{-rs} f(X_{i,s})\, ds \right\}$$

where the supremum is taken over all starting and stopping times of process $X_{i,t}$, denoted respectively as τ_i and σ_i, and the function f is a deterministic function of $X_{i,t}$ yielding the instantaneous payoff of the stochastic process to the observer.

1.3 Prior Work

To formulate solution strategies for our problems, we required theory and techniques that have been previously employed to solve similarly structured problems. While our problem has not been directly solved within existing literature, the aforementioned techniques were instrumental in finding solutions to our problem under certain conditions.

The first major work in multiple stopping problems was done by Gus W. Haggstrom of the University of California at Berkley in 1967 [8]. Presenting solutions in the discrete-time case and for sums of stochastic processes, he was able to extend the theory of optimal one- and two-stopping problems to allow for problems where $r > 2$ stops were possible.

The work of René Carmona and Nizar Touzi in 2008 extended the optimal multiple stopping theory to include valuation procedures for swing options [2]. In energy markets, option contracts exist that allow energy companies to buy excess energy from other companies. Such swing options typically have multiple exercise rights, but the same underlying stochastic process and a minimal time between successive exercises, called refraction time.

Eric Dahlgren and Time Leung in 2015 examined optimal multiple stopping in the context of real options, such as those requiring infrastructure investments, and examine the effect of lead time and project lifetime under this context [5]. Their technique illustrates the potential benefits to several smaller-scale shorter lifetime projects over those that require significantly more investment in infrastructure. The solution technique provided in the paper is one of backward iteration. We conjecture that extension of our problem into the general case for $N > 2$ will require a similar strategy.

Kjell Arne Brekke and Bernt Øksendal in 1994 studied the problem of finding optimal sequences of starting and stopping times in production processes with multiple activities [1]. Included considerations are the costs associated with opening, maintaining, and eventual ending of activities. Their problem was one of a single integral process.

As we desired a numerical implementation for our single candidate results, a natural choice was the least-squares Monte Carlo method developed by Longstaff and Schwartz [10]. Their work began with a standard Monte Carlo simulation on the variable for the purpose of valuing American options. First, they calculated cash flow in the final time period as if the option were European. Looking at the previous (next-to-last) time period, they compare the value of exercising in the next-to-last period with the value of the discounted cash flow of the expected value (calculated via regression) of continuing, but only on those paths that were in-the-money. If the expected value of continuing exceeded the current exercise value, the model chooses to continue. Then they examined the next previous time period and proceed as before. We adapted this technique to our problem, choosing as our continuation criteria as “continue only in the case that expected future cash flows are nonnegative.”

It is our intention with this work to examine optimal stopping problems where there are different stochastic integrals whose starting and stopping times affect each other. To this end, in the sections that follow we build the necessary background for a discussion of our work.

1.4 Brownian Motion, Markov Process, Itô Process

The stochastic processes that we will use in our formulations are all assumed to be geometric Brownian motions. A Brownian motion is a stochastic process with the following properties [14]:

- It has independent increments,
- the increments are Gaussian random variables, and
- the motion is continuous.

Or, more precisely, we have:

Definition 1.1 (Brownian Motion, [9])**.** A *(standard, one-dimensional) Brownian motion* is a continuous, adapted process $B = \{B_t, \mathscr{F}_t; 0 \leq t < \infty\}$, defined on some probability space $(\Omega, \mathscr{F}, P)$, with properties that $B_0 = 0$ a.s. For $0 \leq s < t$, the increment $B_t - B_s$ is independent of $\mathscr{F}_s$ and is normally distributed with mean zero and variance $t - s$.

Definition 1.2 (d-dimensional Brownian motion, [9])**.** Let d be a positive integer and μ a probability measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Let $B = \{B_t, \mathscr{F}_t; t \geq 0\}$ be a continuous, adapted process with values in $\mathbb{R}^d$, defined on some probability space $(\Omega, \mathscr{F}, P)$. This process is a *d-dimensional Brownian motion with initial distribution* μ, if

(i) $P[B_0 \in \Gamma] = \mu(\Gamma), \ \forall \Gamma \in \mathscr{B}(\mathbb{R}^d)$

(ii) for $0 \leq s < t$, the increment $B_t - B_s$ is independent of $\mathscr{F}_s$ and is normally distributed with mean zero and covariance matrix equal to $(t-s)I_d$ where I_d is the $(d \times d)$ identity matrix.

If μ assigns measure one to some singleton $\{x\}$, we say that B is a *d-dimensional Brownian motion starting at* x.

Definition 1.3 (Markov Process, [9])**.** Let d be a positive integer and μ a probability measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. An adapted, d-dimensional process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on some probability space $(\Omega, \mathscr{F}, P^\mu)$ is said to be a *Markov process with initial distribution* μ if

(i) $P^\mu[X_0 \in \Gamma] = \mu(\Gamma), \ \forall \Gamma \in \mathscr{B}(\mathbb{R}^d)$;

(ii) for $s, t \geq 0$ and $\Gamma \in \mathscr{B}(\mathbb{R}^d)$,

$$P^\mu[X_{t+s} \in \Gamma | \mathscr{F}_s] = P^\mu[X_{t+s} \in \Gamma | X_s], \ P^\mu\text{-a.s.}$$

Definition 1.4 (Markov Family, [9])**.** Let d be a positive integer. A d-dimensional *Markov family* is an adapted process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on some $(\Omega, \mathscr{F})$, together with a family of probability measures $\{P^x\}_{x \in \mathbb{R}^d}$ on $(\Omega, \mathscr{F})$ such that

(a) for each $F \in \mathscr{F}$, the mapping $x \mapsto P^x(F)$ is universally measurable;

(b) $P^x[X_0 = x] = 1, \ \forall x \in \mathbb{R}^d$;

(c) for $x \in \mathbb{R}^d, \ s, t \geq 0$ and $\Gamma \in \mathscr{B}(\mathbb{R}^d)$,

$$P^x[X_{t+s} \in \Gamma | \mathscr{F}_s] = P^x[X_{t+s} \in \Gamma | X_s], \ P^x\text{-a.s.};$$

(d) for $x \in \mathbb{R}^d, \ s, t \geq 0$ and $\Gamma \in \mathscr{B}(\mathbb{R}^d)$,

$$P^x[X_{t+s} \in \Gamma | X_s = y] = P^y[X_t \in \Gamma], \ P^x X_s^{-1}\text{-a.e. } y.$$

Definition 1.5 (Strong Markov Process, [9])**.** Let d be a positive integer and μ a probability measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. A progressively measurable, d-dimensional process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on some $(\Omega, \mathscr{F}, P^\mu)$ is said to be a *strong Markov process with initial distribution* μ if

(i) $P^\mu[X_0 \in \Gamma] = \mu(\Gamma)$, $\forall \Gamma \in \mathscr{B}(\mathbb{R}^d)$;

(ii) for any optional time S of $\{\mathscr{F}_t\}$, $t \geq 0$ and $\Gamma \in \mathscr{B}(\mathbb{R}^d)$,

$$P^\mu[X_{S+t} \in \Gamma | \mathscr{F}_{S+}] = P^\mu[X_{S+t} \in \Gamma | X_S], \; P^\mu\text{-a.s. on } \{S < \infty\}.$$

Definition 1.6 (Strong Markov Family, [9])**.** Let d be a positive integer. A d-dimensional *strong Markov family* is a progressively measurable process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on some $(\Omega, \mathscr{F})$, together with a family of probability measures $\{P^x\}_{x \in \mathbb{R}^d}$ on $(\Omega, \mathscr{F})$, such that:

(a) for each $F \in \mathscr{F}$, the mapping $x \mapsto P^x(F)$ is universally measurable;

(b) $P^x[X_0 = x] = 1, \; \forall x \in \mathbb{R}^d$;

(c) for $x \in \mathbb{R}^d, \; t \geq 0, \; \Gamma \in \mathscr{B}(\mathbb{R}^d)$, and any optional time S of $\{\mathscr{F}_t\}$,

$$P^x[X_{S+t} \in \Gamma | \mathscr{F}_{S+}] = P^x[X_{S+t} \in \Gamma | X_S], \; P^x\text{-a.s. on } \{S < \infty\};$$

(d) for $x \in \mathbb{R}^d, \; t \geq 0, \; \Gamma \in \mathscr{B}(\mathbb{R}^d)$, and any optional time S of $\{\mathscr{F}_t\}$,

$$P^x[X_{S+t} \in \Gamma | X_S = y] = P^y[X_t \in \Gamma], \; P^x X_S^{-1}\text{-a.e. } y.$$

For the strong Markov property, we will often make use of what is called a shift operator. For a given process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on a measurable space $(\Omega, \mathscr{F})$, we may sometimes construct a family of *shift operators* $\theta_s : \Omega \to \Omega$, $s \geq 0$ such that each θ_s is $\mathscr{F}/\mathscr{F}$-measurable and

$$X_{s+t}(\omega) = X_t(\theta_s(\omega)); \; \forall \omega \in \Omega, \; s, t \geq 0. \tag{1.1}$$

We may use shift operators to write a generalized Markov property as follows. If $H = H(\omega)$ is a bounded (or nonnegative) $\mathscr{F}$-measurable function, then for any initial distribution π and for any $t \geq 0$ and $x \in E$ we have

$$\mathbb{E}_\pi \left[H \circ \theta_t | \mathscr{F}_t^X \right] (\omega) = \mathbb{E}_{X_t(\omega)} H \quad \mathbb{P}_x - \text{a.s.} \tag{1.2}$$

However, for any stopping time τ in the family of all stopping times, we have the strong Markov property:

$$\mathbb{E}_\pi \left[H \circ \theta_\tau | \mathscr{F}_\tau^X \right] (\omega) = \mathbb{E}_{X_\tau(\omega)} H \quad \mathbb{P}_x - \text{a.s.} \tag{1.3}$$

Definition 1.7 (Itô Process, [15])**.** Let W_t, $t \geq 0$, be a Brownian motion, and let $\mathscr{F}_t$, $t \geq 0$, be an associated filtration. An *Itô process* is a stochastic process of the form

$$X_t = X_0 + \int_0^t \Delta_u dW_u + \int_0^t \Theta_u du, \tag{1.4}$$

where X_0 is nonrandom and Δ_u and Θ_u are adapted stochastic processes.

All of our stochastic processes will be geometric Brownian motions with drift $\alpha_i > 0$, volatility $\beta_i > 0$, and initial value $X_{i,0} = x_i > 0$, they are Itô processes satisfying

$$X_{i,t} = X_{i,0} + \int_0^t \beta_i dW_{i,u} + \int_0^t \alpha_i du,$$

or equivalently,

$$dX_{i,t} = X_{i,t} \left(\alpha_i dt + \beta_i dW_{i,t} \right).$$

For $X = (X_t)_{t \geq 0}$ a continuous semimartingale satisfying $dX_t = b(X_t)\, dt + \sigma(X_t) dB_t$, $\mathbb{P}(X_s = c(s)) = 0$ for $s \in (0, t]$, $c : \mathbb{R}_+ \to \mathbb{R}$ a continuous function of bounded variation and $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ satisfying

$$F \text{ is } C^{1,2} \text{ on } \overline{C}_1, C_1 = \{(t,x) \in \mathbb{R}_+ \times \mathbb{R} : x > c(t)\},$$

$$F \text{ is } C^{1,2} \text{ on } \overline{C}_2, C_2 = \{(t,x) \in \mathbb{R}_+ \times \mathbb{R} : x < c(t)\},$$

and the infinitessimal generator $\mathbb{L}_X f = bf_x + (\sigma^2/2)F_{xx}$, then we may write Itô's formula as [13]

$$\begin{aligned} F(t, X_t) = F(0, X_0) &+ \int_0^t (F_t + \mathbb{L}_X F)(s, X_s) I(X_s \neq c(s))\ ds \\ &+ \int_0^t F_x(s, X_s)\sigma(X_s) I(X_s \neq c(s)) dB_s \\ &+ \tfrac{1}{2} \int_0^t (F_x(s, X_s+) - F_x(s, X_s-)) I(X_s = c(s)) d\ell_s^c(X), \end{aligned} \tag{1.5}$$

where $\ell_s^c(X)$ is the local time of X at the curve c given by

$$\ell_s^c(X) = \mathbb{P}\text{-}\lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^s I(c(r) - \varepsilon < X_r < c(r) + \varepsilon) d\langle X, X \rangle_r, \tag{1.6}$$

and $\mathbb{P}$- lim indicates limit in probability.

1.5 Optimal Stopping, Martingale Approach

There are two major approaches to continuous-time optimal stopping problems: the martingale approach, and the Markov approach. Both approaches will be outlined below. The main theorems from each approach are from Peskir & Shiryaev, but with added details to their proofs.

Definition 1.8 (Martingale, Submartingale, Supermartingale, [15]). Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathscr{F}_t$, $0 \leq t \leq T$, be a filtration of sub-σ-algebras of $\mathscr{F}$. Consider an adapted stochastic process M_t, $0 \leq t \leq T$.

i) If $\mathbb{E}\{M_t | \mathscr{F}_s\} = M_s$ for all $0, \leq s \leq t \leq T$, we say this process is a *martingale.*

ii) If $\mathbb{E}\{M_t | \mathscr{F}_s\} \geq M_s$ for all $0, \leq s \leq t \leq T$, we say this process is a *submartingale.*

iii) If $\mathbb{E}\{M_t | \mathscr{F}_s\} \leq M_s$ for all $0, \leq s \leq t \leq T$, we say this process is a *supermartingale.*

As we need the result of the Optional Sampling, it is stated below. There are many different modifications on this, but we state J. Doob's stopping time theorem from Peskir & Shiryaev [13].

Theorem 1.1 (Optional Sampling: Doob's Stopping Time Theorem, [13])**.** *Suppose that $X = (X_t, \mathscr{F}_t)_{t\geq 0}$ is a submartingale (martingale) and τ is a Markov time. Then the "stopped" process $X^\tau = (X_{t\wedge\tau}, \mathscr{F}_t)_{t\geq 0}$ is also a submartingale (martingale).*

Below we build the assumptions necessary for what follows. Let $G = (G_t)_{t\geq 0}$ be a stochastic process on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}), P)$ that is adapted to the filtration $(\mathscr{F}_t)_{t\geq 0}$ in the sense that each G_t is $\mathscr{F}_t$-measurable.

Definition 1.9 ([13])**.** A random variable $\tau : \Omega \to [0, \infty]$ is called a *Markov time* if $\{\tau \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. A Markov time is called a *stopping time* if $\tau < \infty$ P-a.s.

Assume G is right-continuous and left-continuous over stopping times. That is, if τ_n and τ are stopping time such that $\tau_n \uparrow \tau$ as $n \to \infty$, then $G_{\tau_n} \to G_\tau$ as $n \to \infty$. Further assume that $G_T = 0$ when $T = \infty$, and that

$$\mathbb{E}\left\{\sup_{0\leq t\leq T} |G_t|\right\} < \infty. \tag{1.7}$$

The existence of the right-continuous modification of a supermartingale is a result of $(\mathscr{F}_t)_{t\geq 0}$ being a right-continuous filtration, and that each $\mathscr{F}_t$ contains all P-null sets from $\mathscr{F}$.

We define the family of all stopping times τ to be those stopping times satisfying $\tau \geq t$. In the case of $T < \infty$, the family of all stopping times τ satisfies $t \leq \tau \leq T$. Consider the optimal stopping problem

$$V_t^T = \sup_{t\leq\tau\leq T} \mathbb{E}\, G_\tau. \tag{1.8}$$

We have two methods for solving this problem.

i) Begin with the discrete-time problem by replacing time interval $[0, T]$ with $D_n = \{t_0^n, t_1^n, \ldots, t_n^n\}$ where $D_n \uparrow D$ as $n \to \infty$ and D is a countable, dense subset of $[0, T]$. Then we use backward induction methods and take limits. This method is particularly useful for numerical approximations.

ii) Use the method of essential supremum.

For discrete time problems of finite horizon, we have a method for determining an explicit solution from with backwards induction via the Bellman equation. For infinite horizon problems, the Bellman equation method involves an initial guess from which we calculate from the fixed point of the guess and then check our answer. However, in the case of continuous time calculations, there is no time directly next to any fixed time t, and so there is no iterative method we can use. Further, for both finite and infinite horizon problems, there are an uncountably infinite number of times t. Infinite horizon problems are easier as there is more likely to be a closed form solution to the problem. This not necessarily the case for finite horizon problems [13].

We will not treat finite and infinite horizon problems differently. This allows us to simplify notation to $V_t = V_t^T$. Consider the process $S = (S_t)_{t \geq 0}$ defined by

$$S_t = \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}\left\{ G_\tau | \mathscr{F}_t \right\} \tag{1.9}$$

where τ is a stopping time. We call S the right modification of G. In cases of finite horizon, we require $\tau \leq T$. The process S is called the *Snell envelope* of G.

Consider the stopping time

$$\tau_t = \inf\{s \geq t : S_s = G_s\} \tag{1.10}$$

for $t \geq 0$ where $\inf \emptyset = \infty$. In the case of finite horizon, we require in Equation (1.10) that $s \leq T$.

We will require the concluding statements of the following lemma in the proof of the main theorem for the martingale approach, stated here for completeness.

Lemma 1.2 (Essential Supremum, [13]). *Let* $\{Z_\alpha : \alpha \in I\}$ *be a family of random variables defined on* $(\Omega, \mathscr{G}, P)$ *where the index set* I *can be arbitrary. Then there exists a countable subset* J *of* I *such that the random variable* $Z^* : \Omega \to \bar{\mathbb{R}}$ *defined by*

$$Z^* = \sup_{\alpha \in J} Z_\alpha \tag{1.11}$$

satisfies the following two properties:

$$P(Z_\alpha \leq Z^*) = 1 \ \forall \alpha \in I. \tag{1.12}$$

$$\text{If } \tilde{Z} : \Omega \to \bar{\mathbb{R}} \text{ is another random variable satisfying} \tag{1.13}$$

$$\text{Equation (1.12) in place of } Z^*, \text{ then } P(Z^* \leq \tilde{Z}) = 1.$$

The random variable Z^* *is called the* essential supremum *of* $\{Z_\alpha : \alpha \in I\}$ *relative to* P *and is denoted by* $Z^* = \operatorname{ess\,sup}_{\alpha \in I} Z_\alpha$. *It is determined by the two properties above uniquely up to a* P*-null set. Moreover, if the family* $\{Z_\alpha : \alpha \in I\}$ *is upwards directed in the sense that*

$$\forall \alpha, \beta \in I \ \exists \gamma \in I \ni Z_\alpha \vee Z_\beta \leq Z_\gamma \ P\text{-a.s}, \tag{1.14}$$

then the countable set $J = \{\alpha_n : n \geq 1\}$ *can be chosen so that*

$$Z^* = \lim_{n \to \infty} Z_{\alpha_n} \ P\text{-a.s.}, \tag{1.15}$$

where $Z_{\alpha_1} \leq Z_{\alpha_2} \leq \ldots$ P*-a.s.*

We may now state the main theorem of the martingale approach. The theorem is Theorem 2.2 from Peskir & Shiryaev, but we add details to their proof.

Theorem 1.3 (Martingale Approach, [13]). *Consider the optimal stopping problem Equation* (1.8)

$$V_t^T = \sup_{0\leq\tau\leq T} \mathbb{E}\, G_\tau$$

upon assuming that the condition Equation (1.7)

$$\mathbb{E}\left\{\sup_{0\leq t\leq T} |G_t|\right\} < \infty$$

holds. Assume moreover when required below that

$$P(\tau_t < \infty) = 1 \tag{1.16}$$

where $t \geq 0$ *and* τ_t *is the stopping time* $\inf\{s \geq t : S_t = G_s\}$*. (This condition is automatically satisfied when the horizon* T *is finite). Then for all* $t \geq 0$ *we have:*

$$S_t \geq \mathbb{E}\left\{G_\tau | \mathscr{F}_t\right\} \text{ for each } \tau \text{ in the set of all stopping times,} \tag{1.17}$$

$$S_t = \mathbb{E}\left\{G_{\tau_t} | \mathscr{F}_t\right\}. \tag{1.18}$$

Moreover, if $t \geq 0$ *is given as fixed then we have:*

(a) The stopping time τ_t *is optimal in Equation* (1.8).

(b) If τ_* *is an optimal stopping time in Equation* (1.8) *then* $\tau_t \leq \tau_*$ *P-a.s.*

(c) The process $(S_s)_{s\geq t}$ *is the smallest right-continuous supermartingale which dominates* $\{G_s\}_{s\geq t}$.

(d) The stopped process $(S_{s\wedge\tau_t})_{s\geq t}$ *is a right-continuous martingale.*

Finally, if the condition in Equation (1.16) *fails so that* $P(\tau_t = \infty) > 0$*, then with probability 1 there is no optimal stopping time in Equation* (1.8).

Proof:

We first establish that $S = (S_t)_{t\geq 0} = \left(\operatorname{ess\,sup}_{\tau\geq t} \mathbb{E}\left\{ G_\tau \middle| \mathscr{F}_t\right\}\right)_{t\geq 0}$ is a supermartingale. To this end, fix $t \geq 0$ and we will show that the family $\{\mathbb{E}\{ G_\tau | \mathscr{F}_t\} : \tau \geq t\}$ is upwards directed in the sense that Equation (1.14) is satisfied. Note that for $\sigma_1, \sigma_2 \geq t$and $\sigma_3 = \sigma_1 I_A + \sigma_2 I_{A^C}$, where $A = \{\mathbb{E}\{ G_{\sigma_1} | \mathscr{F}_t\} \geq \mathbb{E}\{ G_{\sigma_2} | \mathscr{F}_t\}\}$, then $\sigma_3 \geq t$ is in the family of stopping timesand

$$\begin{aligned} \mathbb{E}\{ G_{\sigma_3} | \mathscr{F}_t\} &= \mathbb{E}\{ G_{\sigma_1} I_A + G_{\sigma_2} I_{A^C} | \mathscr{F}_t\} \\ &= I_A \mathbb{E}\{ G_{\sigma_1} | \mathscr{F}_t\} + I_{A^C}\mathbb{E}\{ G_{\sigma_2} | \mathscr{F}_t\} \\ &= \mathbb{E}\{ G_{\sigma_1} | \mathscr{F}_t\} \vee \mathbb{E}\{ G_{\sigma_2} | \mathscr{F}_t\}, \end{aligned} \tag{1.19}$$

which establishes that the family is upwards directed in the sense of Equation (1.14). Since (1.14) is satisfied, we may appeal to Equation (1.15) for the existence of a countable set $J = \{\sigma_k : k \geq 1\}$ that is a subset of all stopping times greater than or equal to t such that

$$\operatorname*{ess\,sup}_{\tau\ \geq t} \mathbb{E}\{ G_\tau | \mathscr{F}_t\} = \lim_{k\to\infty} \mathbb{E}\{ G_{\sigma_k} | \mathscr{F}_t\}, \tag{1.20}$$

where $\mathbb{E}\{ G_{\sigma_1} | \mathscr{F}_t\} \leq \mathbb{E}\{ G_{\sigma_2} | \mathscr{F}_t\} \leq \ldots$ P-a.s.

Since $S_t = \operatorname{ess\,sup}_{\tau\geq t} \mathbb{E}\{ G_\tau | \mathscr{F}_t\}$ and the conditional monotone convergence theorem and by condition (1.7)

$$\mathbb{E}\left\{ \sup_{0\leq t\leq T} |G_t| \right\} < \infty,$$

we must have that for all $s \in [0, t]$

$$\begin{aligned} \mathbb{E}\{ S_t | \mathscr{F}_s\} &= \mathbb{E}\left\{ \lim_{k\to\infty} \mathbb{E}\{ G_{\sigma_k} | \mathscr{F}_t\} \middle| \mathscr{F}_s \right\} \\ &= \lim_{k\to\infty} \mathbb{E}\{\mathbb{E}\{ G_{\sigma_k} | \mathscr{F}_t\} | \mathscr{F}_s\} \end{aligned} \tag{1.21}$$

$$= \lim_{k\to\infty} \mathbb{E}\{G_{\sigma_k} | \mathscr{F}_s\}$$
$$\leq S_s \text{ by the definition of } S,$$

and hence $(S_t)_{t\geq 0}$ is a martingale. Note also that the definition of S and Equation (1.20) imply that

$$\mathbb{E}\ S_t = \sup_{\tau\geq t} \mathbb{E}\ G_\tau, \tag{1.22}$$

where τ is a stopping time and $t \geq 0$.

Next we establish that the supermartingale S admits a right-continuous modification $\widetilde{S} = (\widetilde{S}_t)_{t\geq 0}$. A right-continuous modification of a supermartingale is possible if and only if

$$t \mapsto \mathbb{E}\ S_t \text{ is right-continuous on } \mathbb{R}_+ \tag{1.23}$$

is satisfied [11]. We have by the supermartingale property of S that $\mathbb{E}\ S_t \geq \cdots \geq \mathbb{E}\ S_{t_2} \geq \mathbb{E}\ S_{t_1}$, i.e. $\mathbb{E}\ S_{t_n}$ is an increasing sequence of numbers. Define $L := \lim_{n\to\infty} \mathbb{E}\ S_{t_n}$, which must exist by the supermartingale property and as it is an increasing sequence bounded above. Further, we have that $\mathbb{E}\ S_t \geq L$ for given and fixed t_n such that $t_n \downarrow t$ as $n \to \infty$. Fix $\varepsilon > 0$ and by Equation (1.22) choose $\sigma \geq t$ such that

$$\mathbb{E}\ G_\sigma \geq \mathbb{E}\ G_t - \varepsilon. \tag{1.24}$$

Fix $\delta > 0$. Note that we are under no restriction to assume that $t_n \in [t, t+\delta]$ for all $n \geq 1$. Define stopping time σ_n by setting

$$\sigma_n = \begin{cases} \sigma & \text{if} \quad \sigma > t_n, \\ t+\delta & \text{if} \quad \sigma \leq t_n \end{cases} \tag{1.25}$$

for $n \geq 1$. For all $n \geq 1$ we have

$$\mathbb{E}\ G_{\sigma_n} = \mathbb{E}\ G_\sigma I(\sigma > t_n) + \mathbb{E}\ G_{t+\delta} I(\sigma \leq t_n) \leq \mathbb{E}\ S_{t_n} \tag{1.26}$$

since $\sigma_n \geq t_n$ and Equation (1.22) holds. As $n \to \infty$ in Equation (1.26) and using condition (1.7) we have by the Bounded Convergence Theorem

$$\mathbb{E}\ G_\sigma I(\sigma > t) + \mathbb{E}\ G_{t+\delta} I(\sigma = t) \leq L \tag{1.27}$$

for all $\delta > 0$. Letting $\delta \downarrow 0$ and by virtue of G being right-continuous, we have

$$\mathbb{E}\ G_\sigma I(\sigma > t) + \mathbb{E}\ G_t I(\sigma = t) = \mathbb{E}\ G_\sigma \leq L. \tag{1.28}$$

That is, $L \geq \mathbb{E}\ S_t - \varepsilon$ for all $\varepsilon > 0$, and hence $L \geq \mathbb{E}\ S_t$. Thus $L = \mathbb{E}\ S_t$ and statement (1.23) holds, and therefore S does indeed admit a right continuous modification $\widetilde{S} = (\widetilde{S}_t)_{t \geq 0}$. To simplify notation, we will denote the the right-continuous modification as S for the remainder of the proof.

We may now establish statement (c). Denote $\hat{S} = (\hat{S}_s)_{s \geq t}$ be another right-continuous supermartingale dominating $G = (G_s)_{s \geq t}$. By the optional sampling theorem and using condition (1.7) we have

$$\hat{S}_s \geq \mathbb{E} \left\{ \hat{S}_\tau \middle| \mathscr{F}_s \right\} \geq \mathbb{E} \left\{ G_\tau | \mathscr{F}_s \right\} \tag{1.29}$$

for all $\tau \geq s$when $s \geq t$. By the definition of S_s we have that $S_s \leq \hat{S}_s$ P-a.s. for all $s \geq t$. By the right-continuity of S and $\hat{S}$, this further establishes the claim that $P(S_s \leq \hat{S}_s \text{ for all } s \geq t) = 1$. We may now establish Equations (1.17) and (1.18).

By the definition of S_t, Equation (1.17) follows immediately. To establish (1.18), we consider cases. For the first case, consider $G_t \geq 0$ for all $t \geq 0$. Then for each

$\lambda \in (0,1)$, we define the stopping time

$$\tau_t^\lambda = \inf\{s \geq t : \lambda S_s \leq G_s\}, \tag{1.30}$$

where $t \geq 0$ is given and fixed.

Note that by the right-continuity of S and G we have for all $\lambda \in (0,1)$

$$\lambda S_{\tau_t^\lambda} \leq G_{\tau_t^\lambda}, \tag{1.31}$$

$$\tau_{t+}^\lambda = \tau_t^\lambda, \tag{1.32}$$

$$S_{\tau_t} = G_{\tau_t}, \tag{1.33}$$

$$\tau_{t+} = \tau_t, \tag{1.34}$$

where $\tau_t = \inf\{s \geq t : S_s = G_s\}$, as defined earlier. The optional sampling theorem and condition (1.7) implies

$$S_t \geq \mathbb{E}\left\{ S_{\tau_t^\lambda} \middle| \mathscr{F}_t \right\} \tag{1.35}$$

since τ_t^λ is a stopping time greater than or equal to t. To show that the reverse inequality hold, let us consider the process

$$R_t = \mathbb{E}\left\{ S_{\tau_t^\lambda} \middle| \mathscr{F}_t \right\} \tag{1.36}$$

for $t \geq 0$.

For $s < t$ we have

$$\mathbb{E}\left\{ R_t \middle| \mathscr{F}_s \right\} = \mathbb{E}\left\{ \mathbb{E}\left\{ S_{\tau_t^\lambda} \middle| \mathscr{F}_t \right\} \middle| \mathscr{F}_s \right\} = \mathbb{E}\left\{ S_{\tau_t^\lambda} \middle| \mathscr{F}_s \right\} \leq \mathbb{E}\left\{ S_{\tau_s^\lambda} \middle| \mathscr{F}_s \right\} = R_s, \tag{1.37}$$

where the inequality is a result of the optional sampling theorem and using condition (1.7), since $\tau_t^\lambda \geq \tau_s^\lambda$ when $s < t$. Thus R is a supermartingale. Therefore $\mathbb{E}\ R_{t+h}$ increases as h decreases and $\lim_{h \downarrow 0} \mathbb{E}\ R_{t+h} \leq \mathbb{E}\ R_t$. But by Fatou's lemma, condition

(1.7), the fact that τ^λ_{t+h} decreases as h decreases, and S is right-continuous we have

$$\lim_{h\downarrow 0} \mathbb{E}\ R_{t+h} = \lim_{h\downarrow 0} \mathbb{E}\ S_{\tau^\lambda_{t+h}} \geq \mathbb{E}\ S_{\tau^\lambda_t} = \mathbb{E}\ R_t. \tag{1.38}$$

Thus $t \mapsto \mathbb{E}\ R_t$ is right-continuous on $\mathbb{R}_+$, and hence R admits a right-continuous modification. We are therefore under no restriction to make a further assumption that R is right-continuous. To finish showing the reverse inequality, i.e. that $S_t \leq R_t = \mathbb{E}\left\{ S_{\tau^\lambda_t} \middle| \mathscr{F}_t \right\}$, consider the right-continuous supermartingale

$$L_t = \lambda S_t + (1-\lambda) R_t \tag{1.39}$$

for $t \geq 0$.

To proceed further, we require the following claim:

$$L_t \geq G_t \ \mathbb{P}\text{-a.s.} \tag{1.40}$$

for all $t \geq 0$. However, since

$$\begin{aligned}
L_t &= \lambda S_t + (1-\lambda) R_t \\
&= \lambda S_t + (1-\lambda) R_t I(\tau^\lambda_t = t) + (1-\lambda) R_t I(\tau^\lambda_t > t) \\
&= \lambda S_t + (1-\lambda)\mathbb{E}\left\{ S_t I(\tau^\lambda_t = t) \middle| \mathscr{F}_t \right\} + (1-\lambda) R_t I(\tau^\lambda_t > t) \\
&= \lambda S_t I(\tau^\lambda_t = t) + (1-\lambda) S_t I(\tau^\lambda_t = t) + \lambda S_t I(\tau^\lambda_t > t) \\
&\quad + (1-\lambda) R_t I(\tau^\lambda_t > t) \\
&\geq S_t I(\tau^\lambda_t = t) + \lambda S_t I(\tau^\lambda_t > t), \text{ as } R_t \geq 0 \\
&\geq G_t I(\tau^\lambda_t = t) + G_t I(\tau^\lambda_t > t), \text{ by definition of } \tau^\lambda_t \\
&= G_t.
\end{aligned} \tag{1.41}$$

This establishes the claim. Since S is the smallest right-continuous supermartingale dominating G, we see that Equation (1.41) implies

$$S_t \leq L_t \ \mathbb{P}\text{-a.s.}, \tag{1.42}$$

and so by (1.39) we may conclude that $S_t \leq R_t$ P-a.s. Thus the reverse inequality holds and

$$S_t = \mathbb{E}\left\{ S_{\tau_t^\lambda} \middle| \mathscr{F}_t \right\} \tag{1.43}$$

for all $\lambda \in (0,1)$. Thus we must have

$$S_t \leq \frac{1}{\lambda}\mathbb{E}\left\{ G_{\tau_t^\lambda} \middle| \mathscr{F}_t \right\} \tag{1.44}$$

for all $\lambda \in (0,1)$. By letting $\lambda \uparrow 1$ and using the conditional Fatou's lemma, condition (1.7) and the fact that G is left-continuous over stopping times, we obtain

$$S_t \leq \mathbb{E}\left\{ G_{\tau_t^1} \middle| \mathscr{F}_t \right\} \tag{1.45}$$

where τ_t^1 is a stopping time given by $\lim_{\lambda\uparrow 1} \tau_t^\lambda$. By Equation (1.9), the reverse inequality of (1.45) is always satisfied and we may conclude that

$$S_t = \mathbb{E}\left\{ G_{\tau_t^1} \middle| \mathscr{F}_t \right\} \tag{1.46}$$

for all $t \geq 0$. To complete establishing (1.18) it is enough to verify that $\tau_t^1 = \tau_t$. Since $\tau_t^\lambda \leq \tau_t$ for all $\lambda \in (0,1)$, we have $\tau_t^1 \leq \tau_t$. If $\tau_t(\omega) = t$, $\tau_t^1 = \tau_t$ is obviously true. If $\tau_t(\omega) > t$, then there exists $\varepsilon > 0$ such that $\tau_t(\omega) - \varepsilon > t$ and $S_{\tau_t(\omega)-\varepsilon} > G_{\tau_t(\omega)-\varepsilon} \geq 0$. Hence we can find $\lambda \in (0,1)$ and close enough to 1 such that $\lambda S_{\tau_t(\omega)-\varepsilon} \geq G_{\tau_t(\omega)-\varepsilon}$ showing that $\tau_t^\lambda(\omega) \geq \tau_t(\omega) - \varepsilon$. If we let $\lambda \uparrow 1$ and then let $\varepsilon \downarrow 0$ we conclude $\tau_t^1 \geq \tau_t$. Thus $\tau_t^1 = \tau_t$, and the case of $G_t \geq 0$ is proven.

Next we consider G in general satisfying condition (1.7)). Set $H = \inf_{t\geq 0} G_t$ and introduce the right-continuous martingale $M_t = \mathbb{E}\{H|\mathscr{F}_t\}$ for $t \geq 0$ so as to replace the initial gain process G with $\widetilde{G} = (\widetilde{G}_t)_{t\geq 0}$ defined by $\widetilde{G}_t = G_t - M_t$ for $t \geq 0$. Note that $\widetilde{G}$ need not satisfy (1.7), but M is uniformly integrable since $H \in L^1(P)$. $\widetilde{G}$ is right-continuous and not necessarily left-continuous over stopping times due to the existence of M, but M itself is a uniformly integrable martingale so that the optional sampling theorem applies. Clearly $\widetilde{G}_t \geq 0$ and the optional sampling theorem implies

$$\widetilde{S}_t = \operatorname*{ess\,sup}_{\tau\geq t} \mathbb{E}\left\{\widetilde{G}_\tau \middle| \mathscr{F}_t\right\} = \operatorname*{ess\,sup}_{\tau\geq t} \mathbb{E}\{G_\tau - M_\tau|\mathscr{F}_t\} = S_t - M_t \tag{1.47}$$

for all $t \geq 0$. The same arguments for justifying the case of $G_t \geq 0$ may be applied to $\widetilde{G}$ and $\widetilde{S}$ to imply Equation (1.18) and the general case is proven.

Now we establish parts (a) and (b). Part (a) follows by taking expectations in Equation (1.18) and using (1.22). To establish (b), we claim that the optimality of τ_* implies that $S_{\tau_*} = G_{\tau_*}$ P-a.s. If the claim were false, then we would have $S_{\tau_*} \geq G_{\tau_*}$ P-a.s. with $P(S_{\tau_*} > G_{\tau_*}) > 0$, and thus $\mathbb{E}\ G_{\tau_*} < \mathbb{E}\ S_{\tau_*} \leq \mathbb{E}\ S_t = V_t$ where the second inequality follows from the optional sampling theorem and the supermartingale property of $(S_s)_{s\geq t}$, while the final inequality is directly from Equation (1.22). The strict inequality contradicts the fact that τ_* is optimal. Hence we must have that $S_{\tau_*} = G_{\tau_*}$ P-a.s.and the claim has been proven. That $\tau_t \leq \tau_*$ P-a.s. follows from the definition of τ_t.

Next we establish part (d). It is enough to show that

$$\mathbb{E}\ S_{\sigma\wedge\tau_t} = \mathbb{E}\ S_t \tag{1.48}$$

for all bounded stopping times $\sigma \geq t$. To this end, note that the optional sampling theorem and condition (1.7) imply $\mathbb{E}\ S_{\sigma\wedge\tau_t} \leq \mathbb{E}\ S_t$. However, by Equations (1.18) and (1.33) we have

$$\mathbb{E}\ S_t = \mathbb{E}\ G_{\tau_t} = \mathbb{E}\ S_{\tau_t} \leq \mathbb{E}\ S_{\sigma\wedge\tau_t}. \tag{1.49}$$

Thus we see that Equation (1.48) holds and thus $(S_{s\wedge\tau_t})_{s\geq t}$ is a right-continuous martingale (right-continuity from (c)), and (d) is established. □

1.6 Optimal Stopping, Markov Approach

While we have formally defined Markov process above, the following definition provides a better intuition into the process and its defining property.

Definition 1.10 (Markov Process, [15])**.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathscr{F}_t$, $0 \leq t \leq T$, be a filtration of sub-σ-algebras of $\mathscr{F}$. Consider an adapted stochastic process X_t, $0 \leq t \leq T$ and for every nonnegative, Borel-measurable function f, then there is another Borel-measurable function g such that

$$\mathbb{E}\left\{ f(X_t) \middle| \mathscr{F}_s \right\} = g(X_s). \tag{1.50}$$

Then we say that X is a *Markov process.*

As we will require the strong Markov property at several points in the discussion that follows, we state without proof a theorem on the Markov Property for Itô diffusions.

Theorem 1.4 (Strong Markov Property for Itô diffusions, [12])**.** *Let f be a bounded Borel function on $\mathbb{R}^n$, τ a stopping time with respect to $\mathscr{F}_t$, $\tau < \infty$ a.s. Then,*

$$\mathbb{E}_x \left\{ f\left(X_{\tau+h}\right) \middle| \mathscr{F}_\tau \right\} = \mathbb{E}_{X_\tau}\ f(X_h) \qquad \textit{for all } h \geq 0. \tag{1.51}$$

Consider a strong Markov process $X = (X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P}_x)$ and taking values in a measurable space $(E, \mathscr{B})$. For simplicity we assume $E = \mathbb{R}^d$, $d \geq 1$, and $\mathscr{B} = \mathscr{B}(\mathbb{R}^d)$ is the Borel σ-algebra on $\mathbb{R}^d$. We assume that the process X starts at x at time zero under $\mathbb{P}_x$ for $x \in E$ and that the sample paths of X are both right- and left-continuous over stopping times. That is, for stopping times $\tau_n \uparrow \tau$ then $X_{\tau_n} \to X_\tau$ $\mathbb{P}_x$-a.s. as $n \to \infty$. It is also assumed that the filtration $(\mathscr{F}_t)_{t\geq 0}$ is right-continuous. Further assume that the mapping $x \mapsto \mathbb{P}_x(F)$ is measurable for each $F \in \mathscr{F}$, and hence $x \mapsto \mathbb{E}_x\, Z$ is measurable for each (bounded or non-negative) random variable Z. Without loss of generality we assume $(\Omega, \mathscr{F}) = (E^{[0,\infty)}, \mathscr{B}^{[0,\infty)})$ so that the shift operator $\theta_t : \Omega \to \Omega$ is well defined by $\theta_t(\omega)(s) = \omega(t+s)$ for $\omega = (\omega(s))_{s\geq 0} \in \Omega$ and $t, s \geq 0$.

For a given measurable function $G : E \to \mathbb{R}$ satisfying $G(X_T) = 0$ for $T = \infty$ and

$$\mathbb{E}_x \left\{ \sup_{0\leq t\leq T} |G(X_t)| \right\} < \infty \tag{1.52}$$

for all $x \in E$, we consider the optimal stopping problem

$$V(x) = \sup_{\tau\geq 0} \mathbb{E}_x\, G(X_\tau) \tag{1.53}$$

where $x \in E$ and the supremum is taken over all stopping times τ of X.

To consider the optimal stopping problem Equation (1.53) when $T = \infty$, we introduce the continuation set

$$C = \{x \in E : V(x) > G(x)\} \tag{1.54}$$

and the stopping set

$$D = \{x \in E : V(x) = G(x)\}. \tag{1.55}$$

If V is lower semicontinuous and G is upper semicontinuous, then C is an open set and D is closed. Introduce the first entry time τ_D of X into D by setting

$$\tau_D = \inf\{t \geq 0 : X_t \in D\}. \tag{1.56}$$

Then τ_D is a stopping (Markov) time with respect to $(\mathscr{F}_t)_{t\geq 0}$ when D is closed since both X and $(\mathscr{F}_t)_{t\geq 0}$ are right-continuous.

Superharmonic functions are important to solving the optimal stopping problem.

Definition 1.11 (Superharmonic, [13])**.** A measurable function $F : E \to \mathbb{R}$ is said to be *superharmonic* if

$$\mathbb{E}_x\ F(X_\sigma) \leq F(x) \tag{1.57}$$

for all stopping times σ and all $x \in E$.

It will be verified in the proof of the next theorem that superharmonic functions have the following property whenever F is lower semicontinuous and $(F(X_t))_{t\geq 0}$ is uniformly integrable:

F is superharmonic if and only if $(F(X_t))_{t\geq 0}$ is a right-continuous supermartingale under $\mathbb{P}_x$ for every $x \in E$. (1.58)

This theorem presents the necessary conditions for the existence of an optimal stopping time, quoted from Theorem 2.4 of Peskir & Shiryaev. We provide proof with additional details for the sake of completeness.

Theorem 1.5 (Existence of optimal stopping, [13])**.** *Let us assume that there exists an optimal stopping time τ_* in Equation* (1.53)*,*

$$V(x) = \sup_\tau \mathbb{E}_x\ G(X_\tau).$$

That is, and let τ_* be such that

$$V(x) = \mathbb{E}_x\ G(X_{\tau_*}) \quad \forall x \in E. \tag{1.59}$$

Then we have:

$$\textit{The value function } V \textit{ is the smallest superharmonic function which dominates the gain function } G \textit{ on } E. \tag{1.60}$$

Let us in addition to Equation (1.59) *assume that* V *is lower semicontinuous and* G *is upper semicontinuous. Then we have:*

$$\textit{The optimal stopping time } \tau_D \textit{ satisfies } \tau_D \leq \tau_* \ \mathbb{P}_x\textit{-a.s. for all } x \in E \textit{ and is optimal in } (1.53). \tag{1.61}$$

$$\textit{The stopped process } (V(X_{t\wedge\tau_D}))_{t\geq 0} \textit{ is a right-continuous martingale under } \mathbb{P}_x \textit{ for every } x \in E. \tag{1.62}$$

Proof:

First we establish Equation (1.60). Let $x \in E$, and let σ be a stopping time. Then we have,

$$\begin{aligned}
\mathbb{E}_x\, V(X_\sigma) &= \mathbb{E}_x \mathbb{E}_{X_\sigma}\, G(X_{\tau_*}) \quad \text{by plugging in } V(X_\sigma).\\
&= \mathbb{E}_x \mathbb{E}_x \left\{ G\left(X_{\tau_*} \circ \theta_\sigma\right) \middle| \mathscr{F}_\sigma \right\} \quad \text{by strong Markov}\\
&= \mathbb{E}_x\, G\left(X_{\sigma + \tau_* \circ \theta_\sigma}\right)\\
&\leq \sup_\tau \mathbb{E}_x\, G(X_\tau)\\
&= V(x),
\end{aligned}$$

and hence V is superharmonic. Let F be a superharmonic function dominating G on E. Then we have

$$\mathbb{E}_x\, G(X_\tau) \leq \mathbb{E}_x\, F(X_\tau) \leq F(x)$$

for $x \in E$ and all stopping times τ. Taking the supremum over all stopping times τ, we see that

$$\sup_{\tau} \mathbb{E}_x\ G(X_\tau) = V(x) \leq F(x).$$

Hence V is the smallest superharmonic function dominating G on E.

Next we establish Equation (1.61). Proceed by making the following claim: $V(X_{\tau_*}) = G(X_{\tau_*})$ $\mathbb{P}_x$-a.s. for all $x \in E$.

If $\mathbb{P}_x(V(X_{\tau_*}) > G(X_{\tau_*})) > 0$ for some $x \in E$, then $\mathbb{E}_x\ G(X_{\tau_*}) < \mathbb{E}_x\ V(X_{\tau_*}) \leq V(x)$ as V is superharmonic, which contradicts the optimality of τ_*. Thus the claim is verified.

It follows that $\tau_D \leq \tau_*$ $\mathbb{P}_x$-a.s. for all $x \in E$.

We have that $V(x) \geq \mathbb{E}_x\ V(X_\tau)$ since V is superharmonic. By setting $\sigma \equiv s$ in Equation (1.57), we see that

$$\begin{aligned} V(X_t) &\geq \mathbb{E}_{X_t}\ V(X_s) \\ &= \mathbb{E}_x \left\{ V\left(X_{t+s}\right) \middle| \mathscr{F}_t \right\} \quad \text{by Markov property} \end{aligned}$$

for all $t, x \geq 0$ and all $x \in E$. Since $V(X_t) \geq \mathbb{E}_{X_t} \left\{ V\left(X_{t+s}\right) \middle| \mathscr{F}_t \right\}$, we have that $(V(X_t))_{t\geq 0}$ is a supermartingale under $\mathbb{P}_x$ for all $x \in E$. Since V is lower semicontinuous and $(V(X_t))_{t\geq 0}$ is a supermartingale, we have that $(V(X_t))_{t\geq 0}$ is right-continuous by Proposition 2.5, which is stated below. Thus we have that $\tau_D \leq \tau_*$ $\mathbb{P}_x$-a.s. for all $x \in E$ and is optimal in Equation (1.53).

Now we establish Equation (1.62). Let $x \in E$, $0 \leq s \leq t$. By the strong Markov property,

$$\begin{aligned} \mathbb{E}_x \left\{ V\left(X_{t\wedge\tau_D}\right) \middle| \mathscr{F}_{s\wedge\tau_D} \right\} &= \mathbb{E}_x \left\{ \mathbb{E}_{X_{t\wedge\tau_D}} \left\{ G\left(X_{\tau_D}\right) \middle| \mathscr{F}_{s\wedge\tau_D} \right\} \right\} \\ &= \mathbb{E}_x \left\{ \mathbb{E}_x \left\{ G\left(X_{\tau_D}\right) \circ \theta_{t\wedge\tau_D} \middle| \mathscr{F}_{t\wedge\tau_D} \right\} \middle| \mathscr{F}_{s\wedge\tau_D} \right\} \end{aligned}$$

$$= \mathbb{E}_x \left\{ \mathbb{E}_x \left\{ G\left(X_{\tau_D}\right) \middle| \mathscr{F}_{t\wedge\tau_D} \right\} \middle| \mathscr{F}_{s\wedge\tau_D} \right\}$$

$$= \mathbb{E}_{X_{s\wedge\tau_D}} \ G\left(X_{\tau_D}\right)$$

$$= V(X_{s\wedge\tau_D}).$$

Thus $V(X_{t\wedge\tau_D})$ is a martingale. The right-continuity of $V(X_{t\wedge\tau_D})$ follows by the right-continuity of $(V(X_t))_{t\geq 0}$. □

We will require the statement of the following proposition for the proof of the next theorem.

Proposition 1.6 ([13]). *If a superharmonic function $F : E \to \mathbb{R}$ is lower semicontinuous, then the superharmonic $(F(X_t))_{t\geq 0}$ is right-continuous $\mathbb{P}_x$=a.s. for every $x \in E$.*

We now turn our attention to the main theorem of this section, quoted from Theorem 2.7 of Peskir & Shiryaev. We provide proof with added details.

Theorem 1.7 (Markov Approach, [13]). *Consider the optimal stopping problem Equation* (1.53),

$$V(x) = \sup_{\tau} \mathbb{E}_x \ G\left(X_\tau\right),$$

upon assuming that the condition Equation (1.52),

$$\mathbb{E}_x \left\{ \sup_{0\leq t\leq T} |G(X_t)| \right\} < \infty,$$

is satisfied. Let us assume that there exists the smallest superharmonic function $\hat{V}$ which dominates the gain function G on E. Let us in addition assume that $\hat{V}$ is lower semicontinuous and G is upper semicontinuous. Set $D = \{x \in E : \hat{V}(x) = G(x)\}$ and let τ_D be defined by Equation (1.56),

$$\tau_D = \inf\{t \geq 0 : X_t \in D\}.$$

We then have:

$$\text{If } \mathbb{P}_x(\tau_D < \infty) = 1 \text{ for all } x \in E, \text{ then } \hat{V} = V \text{ and } \tau_D \text{ is optimal in Equation (1.53).} \tag{1.63}$$

$$\text{If } \mathbb{P}_x(\tau_D < \infty) = 1 \text{ for some } x \in E, \text{ then there is no optimal stopping time OME in Equation (1.53).} \tag{1.64}$$

Proof:

Since $\hat{V}$ is superharmonic, we have

$$\mathbb{E}_x\, G\,(X_\tau) \le \mathbb{E}_x\, \hat{V}\,(X_\tau) \le \hat{V}(x) \tag{1.65}$$

for all stopping times τ and all $x \in E$. Taking the supremum over all τ of both sides of $\mathbb{E}_x\, V(X_\tau) \le \mathbb{E}_x\, V(X_\sigma)$, for stopping times σ and τ such that $\sigma \le \tau$ $\mathbb{P}_x$-a.s. with $x \in E$, we find that

$$G(x) \le V(x) \le \hat{V}(x) \tag{1.66}$$

for all $x \in E$.

To establish Equation (1.63), we assume that $P_x(\tau_D < \infty) = 1$ for all $x \in E$, and that G is bounded. Then for given and fixed $\varepsilon > 0$, consider the sets

$$C_\varepsilon = \{x \in E : \hat{V}(x) > G(x) + \varepsilon\}, \tag{1.67}$$

$$D_\varepsilon = \{x \in E : \hat{V}(x) \le G(x) + \varepsilon\}. \tag{1.68}$$

Since $\hat{V}$ is lower semicontinuous and G is upper semicontinuous, we have that C_ε is open and D_ε is closed. Further, we also have that $C_\varepsilon \uparrow C$ and $D_\varepsilon \downarrow D$ as $\varepsilon \downarrow 0$, where C and D are defined by Equations (1.54) and (1.55), respectively.

Define the stopping time

$$\tau_{D_\varepsilon} = \inf\{t \geq 0 : X_t \in D_\varepsilon\}. \tag{1.69}$$

Since $D \subseteq D_\varepsilon$ and $\mathbb{P}_x(\tau_D < \infty) = 1$ for all $x \in E$, we see that $\mathbb{P}_x(\tau_{D_\varepsilon} < \infty) = 1$ for all $x \in E$. To show that $\mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) = \hat{V}(x)$ for all $x \in E$, we must first show that $G(x) \leq \mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right)$ for all $x \in E$. To this end, we set

$$c = \sup_{x\in E} \left(G(x) - \mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right)\right) \tag{1.70}$$

and note that

$$G(x) \leq c + \mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \tag{1.71}$$

for all $x \in E$. Further note that c is finite as G is bounded, and hence $\hat{V}$ is bounded. By the strong Markov property we have

$$\begin{aligned}
\mathbb{E}_x \left\{\mathbb{E}_{X_\sigma}\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right)\right\} &= \mathbb{E}_x \left\{\mathbb{E}_x \left\{\hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \circ \theta_\sigma\right\} \Big|\, \mathscr{F}_\sigma\right\} \\
&= \mathbb{E}_x \left\{\mathbb{E}_x \left\{\hat{V}\left(X_{\sigma+\tau_{D_\varepsilon}\circ\theta_\sigma}\right)\right\} \Big|\, \mathscr{F}_\sigma\right\} \\
&= \mathbb{E}_x\, \hat{V}\left(X_{\sigma+\tau_{D_\varepsilon}\circ\theta_\sigma}\right) \\
&\leq \mathbb{E}_x\, \hat{V}(X_{\tau_{D_\varepsilon}})
\end{aligned} \tag{1.72}$$

using the fact that $\hat{V}$ is superharmonic and lower semicontinuous from the above proposition, and that $\sigma + \tau_{D_\varepsilon} \circ \theta_\sigma \geq \tau_{D_\varepsilon}$ since τ_{D_ε} is the first entry time to a set. This shows that the function

$$x \mapsto \mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \quad \text{is superharmonic} \tag{1.73}$$

from E to $\mathbb{R}$. Hence $c + \mathbb{E}_x\, \hat{V}\left(X_{\tau_{D_\varepsilon}}\right)$ is also superharmonic and we may conclude by the definition of $\hat{V}$ that

$$\hat{V}(x) \leq c + \mathbb{E}_x \; \hat{V}(X_{\tau_{D_\varepsilon}}) \tag{1.74}$$

for all $x \in E$.

Given $0 < \delta \leq \varepsilon$ choose $x_\delta \in E$ such that

$$G(x_\delta) - \mathbb{E}_{x_\delta} \; \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \geq c - \delta. \tag{1.75}$$

Then by Equations (1.74) and (1.75) we get

$$\hat{V}(x_\delta) \leq c + \mathbb{E}_{x_\delta} \; \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \leq G(x_\delta) + \delta \leq G(x_\delta) + \varepsilon. \tag{1.76}$$

This shows that $x_\delta \in D_\varepsilon$ and thus $\tau_{D_\varepsilon} \equiv 0$ under $\mathbb{P}_{x_\delta}$. Inserting this conclusion into Equation (1.75) we have

$$c - \delta \leq G(x_\delta) - \hat{V}(x_\delta) \leq 0. \tag{1.77}$$

Letting $\delta \downarrow 0$ we see that $c \leq 0$, thus establishing $G(x) \leq \mathbb{E}_x \; \hat{V}(X_{\tau_{D_\varepsilon}})$ for all $x \in E$. Using the definition of $\hat{V}$ and Equation (1.73), we immediately see that $\mathbb{E}_x \; \hat{V}(X_{\tau_{D_\varepsilon}}) = \hat{V}(x)$ for all $x \in E$. And from this result, we get

$$\hat{V}(x) = \mathbb{E}_x \; \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) \leq \mathbb{E}_x \; G\left(X_{\tau_{D_\varepsilon}}\right) + \varepsilon \leq V(x) + \varepsilon \tag{1.78}$$

for all $x \in E$ upon using the fact that $\hat{V}(X_{\tau_{D_\varepsilon}}) \leq G(X_{\tau_{D_\varepsilon}}) + \varepsilon$ since $\hat{V}$ is lower semicontinuous and G is upper semicontinuous. Letting $\varepsilon \downarrow 0$ in Equation (1.78) we see that $\hat{V} \leq V$ and thus by Equation (1.66) we can conclude that $\hat{V} = V$. From (1.78) we also have that

$$V(x) \leq \mathbb{E}_x \; G\left(X_{\tau_{D_\varepsilon}}\right) + \varepsilon \tag{1.79}$$

for all $x \in E$. Letting $\varepsilon \downarrow 0$ and using that $D_\varepsilon \downarrow D$ we see that $\tau_{D_\varepsilon} \uparrow \tau_0$ where τ_0 is a stopping time satisfying $\tau_0 \leq \tau_D$. Since V is lower semicontinuous and G is upper semicontinuous we see from the definition of τ_{D_ε} that $V(X_{\tau_{D_\varepsilon}}) \leq G(X_{\tau_{D_\varepsilon}}) + \varepsilon$ for all $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$ and using that X is left-continuous over stopping times, it follows that $V(X_{\tau_0}) \leq G(X_{\tau_0})$ since V is lower semicontinuous and G is upper semicontinuous. This shows that $V(X_{\tau_)}) = G(X_{\tau_0})$ and therefore $\tau_D \leq \tau_0$, showing that $\tau_0 = \tau_D$. Thus $\tau_{D_\varepsilon} \uparrow \tau_D$ as $\varepsilon \downarrow 0$. Making use of the latter fact in $\mathbb{E}_x \hat{V}\left(X_{\tau_{D_\varepsilon}}\right) = \hat{V}(x)$ after letting $\varepsilon \downarrow 0$ and applying Fatou's lemma, we have

$$\begin{aligned} V(x) &\leq \limsup_{\varepsilon \downarrow 0} \mathbb{E}_x \, G\left(X_{\tau_{D_\varepsilon}}\right) && (1.80) \\ &\leq \mathbb{E}_x \limsup_{\varepsilon \downarrow 0} \, G\left(X_{\tau_{D_\varepsilon}}\right) \\ &\leq \mathbb{E}_x \, G\left(\limsup_{\varepsilon \downarrow 0} X_{\tau_{D_\varepsilon}}\right) \\ &= \mathbb{E}_x \, G\left(X_{\tau_{D_\varepsilon}}\right) && (1.81) \end{aligned}$$

using that G is upper semicontinuous. This shows that τ_D is optimal in the case where G is bounded. □

1.7 Cissé-Patie-Tanré Method

A technique we will make use of later developed by Cissé et. al in 2012 [4]. will allow us to transform these integral problems into ones without integrals. Consider the optimal stopping problem

$$\begin{aligned} \Phi(x) &= \sup_{\sigma \geq 0} \mathbb{E}_x \left\{ e^{-r\sigma} g(X_\sigma) - C_\sigma \right\}, \\ C_t &= \int_0^t e^{-rs} c(X_s)\, ds, \end{aligned}$$

where $r > 0$, $c(x)$ and $g(x)$ are deterministic functions, and X_t a geometric Brownian Motion starting at x. Denote $\delta(x) = \mathbb{E}_x\left\{C_\infty\right\}$. For a stopping time σ, we have the identity $C_\infty = C_\sigma + e^{-r\sigma} C_\infty \circ \theta_\sigma$, where θ denotes the shift operator. Note that for $s = u + \sigma$, $ds = du$ and $X_{u+\sigma} = X_u \circ \theta_\sigma$ for a stopping time σ. The justification of this identity follows directly from application of properties of shift operators:

$$\begin{aligned} C_\infty &= \int_0^\infty e^{-rs} c(X_s)\, ds \\ &= \int_0^\sigma e^{-rs} c(X_s)\, ds + \int_\sigma^\infty e^{-rs} c(X_s)\, ds \\ &= \int_0^\sigma e^{-rs} c(X_s)\, ds + \int_0^\infty e^{-r(u+\sigma)} c(X_u \circ \theta_\sigma)\, du \\ &= \int_0^\sigma e^{-rs} c(X_s)\, ds + e^{-r\sigma} \left(\int_0^\infty e^{-ru} c(X_u)\, du \right) \circ \theta_\sigma \\ &= C_\sigma + e^{-r\sigma} C_\infty \circ \theta_\sigma. \end{aligned}$$

The following lemma allows us to transform our problems later. For ease of discussion, for the purposes of this paper we will denote this as the CPT approach.

Lemma 1.8 ([4]). *If $\delta(x)$ is finite on the domain E, then for any $x \in E$ we have*

$$\sup_{\sigma \geq 0} \mathbb{E}_x \left\{ e^{-r\sigma} g(X_\sigma) - C_\sigma \right\} = \sup_{\sigma \geq 0} \mathbb{E}_x \left\{ e^{-r\sigma} \left(g(X_\sigma) - \delta(X_\sigma) \right) \right\} - \delta(x), \tag{1.82}$$

were σ is any stopping time in the set of all stopping times.

Proof:

As $C_\infty = C_\sigma + e^{-r\sigma} C_\infty \circ \theta_\sigma$, we have $-C_\sigma = e^{-r\sigma} C_\infty \circ \theta_\sigma - C_\infty$. Hence by iterated conditioning and the strong Markov property,

$$\begin{aligned} &\mathbb{E}_x \left\{ \mathbb{E}_x \left\{ e^{-r\sigma} g(X_\sigma) - C_\sigma \middle| \mathscr{F}_\sigma \right\} \right\} \\ &\quad = \mathbb{E}_x \left\{ e^{-r\sigma} g(X_\sigma) + \mathbb{E}_x \left\{ e^{-r\sigma} C_\infty \circ \theta_\sigma - C_\infty \middle| \mathscr{F}_\sigma \right\} \right\} \\ &\quad = \mathbb{E}_x \left\{ e^{-r\sigma} g(X_\sigma) + e^{-r\sigma} \mathbb{E}_{X_\sigma} C_\infty \right\} - \mathbb{E}_x\, C_\infty \end{aligned}$$

$$= \mathbb{E}_x \left\{ e^{-r\sigma} \left(g(X_\sigma) + \delta(X_\sigma) \right) \right\} - \delta(x)$$

as desired. □

CHAPTER 2: SINGLE VARIABLE FORMULATIONS

2.1 Problem Formulation

We begin our discussion of our extensions by examining single candidate cases with infinite horizon. We have, in general, two approaches for the infinite horizon problem: PDE and change of numeraire. While we use both in the two variable case, we will only use the PDE method and the Cissé-Patie-Tanré method, which is also PDE-based. First, we examine the case in which the single candidate arrives and is hired immediately at time $t = 0$. Hence, for this case, our problem will result in an integral from zero to stopping time σ. The value function is given by

$$\Phi(x) = \sup_{\sigma \geq 0} \mathbb{E}_x \left\{ \int_0^\sigma e^{-rs} f(X_s) ds \right\}. \tag{2.1}$$

Here Φ denotes the value function that maximizes the expected value over all stopping times σ, X_t is the candidate's innate value, and f is the value of hte candidate to the observer. The supremum is taken over all possible stopping times $\sigma > 0$. The stochastic process X_t is a geometric Brownian motion with dynamics

$$dX_t = X_t \left(\alpha \ dt + \beta \ dW_t \right), \tag{2.2}$$

where W_t is a one-dimensional Brownian motion. We assume throughout that $r > 0$, $\alpha > 0$, $\beta > 0$, and $X_0 = x_0 > 0$. The following remark allows us to begin our search for solutions.

Remark 2.9 (Øksendal [12])**.** Let $h \in C^2(\mathbb{R}^n)$ and $\mathcal{A}$ be the characteristic operator for X. Let h^* be the optimal reward function for the optimal stopping problem $\sup_\tau \mathbb{E}\{h(X_\tau)\}$. Define the continuation region to be

$$\mathcal{C} = \{x : h(x) < h^*(x)\} \subset \mathbb{R}^n.$$

Then for

$$\mathcal{U} = \{x : \mathcal{A}_X h(x) > 0\},$$

we observe that $\mathcal{U} \subseteq \mathcal{C}$, and it is never optimal to exercise at any $x \in \mathcal{U}$. It is only optimal to exercise upon exiting $\mathcal{C}$. However, as it may be the case that $\mathcal{U} \neq \mathcal{C}$, it may be optimal to continue beyond the boundary of $\mathcal{U}$ and exercise by exiting $\mathcal{C} \setminus \mathcal{U}$.

As the integral of a one-dimensional Markov process is not itself Markovian, we instead consider the following two-dimensional Markov process:

$$\begin{aligned} dZ_t &= \begin{bmatrix} dX_t \\ dY_t \end{bmatrix} := \begin{bmatrix} \alpha X_t \\ e^{-rt} f(X_t) \end{bmatrix} dt + \begin{bmatrix} \beta X_t \\ 0 \end{bmatrix} dW_t, \\ Z_0 &= z_0 = (X_0, Y_0), \end{aligned}$$

where $Y_t = \int_0^t e^{-rs} f(X_s) ds$. Then,

$$\begin{aligned} \Phi(x_0) &= \sup_\sigma \mathbb{E}_{x_0} \{Y_\sigma\} \\ &= \sup_\sigma \mathbb{E}_{x_0,0} \{Y_\sigma + g(X_\sigma)\}, \end{aligned}$$

where $g(x) \equiv 0$.

Then we have

$$\Phi(x_0) = \sup_{\sigma} \mathbb{E}_{x_0,0} \{Y_\sigma + g(X_\sigma)\}$$
$$= \sup_{\sigma} \mathbb{E}_{x_0,0} \{\tilde{g}(Z_\sigma)\},$$

where $\tilde{g}(z) := y+g(x)$. The characteristic operator of X is $\mathcal{A}_X = \frac{1}{2}\beta^2 x^2 \partial_{xx} + \alpha x \partial_x + \partial_t$. Then the characteristic operator $\mathcal{A}_Z$ of Z_t acting on a function ϕ is

$$\mathcal{A}_Z \phi(z) = \mathcal{A}_X \phi(t,x) + e^{-rt} f(x) \frac{\partial \phi}{\partial y}, \tag{2.3}$$

and we begin by examining a subset of the continuation region where

$$\mathcal{U} = \left\{(x,t) \in \mathbb{R}^+ \times \mathbb{R}^+ : \mathcal{A}_Z \tilde{g} > 0\right\}. \tag{2.4}$$

Notice that, in general,

$$\mathcal{A}_Z \tilde{g} = \mathcal{A}_X \tilde{g}(x) + e^{-rt} f(x) \frac{\partial \tilde{g}}{\partial y}$$
$$= 0 + e^{-rt} f(x),$$

and as $e^{-rt} > 0$ for all t, we will primarily be investigating $f(x) > 0$ in our examination of $\mathcal{U}$. We can now examine several specific cases for $f(x)$.

2.2 The case of $f(x) = x$

In the case where the candidate's instantaneous value is $f(x) = x$, and we have discounting inside the integral, so the subset of the continuation region where we begin our investigation is

$$\mathcal{U} = \left\{(x,t) \in \mathbb{R}^+ \times \mathbb{R}^+ : \mathcal{A}_X g(x) + e^{-rt} f(x) > 0\right\}$$

$$= \left\{ (x,t) \in \mathbb{R}^+ \times \mathbb{R}^+ : x > 0 \right\}.$$

For candidate ability X_t a Geometric Brownian Motion with annual drift α and annual volatility β, this subset of the continuation region is the entire domain and therefore the continuation region is the entire domain. That is, there is no finite stopping time, and $\sigma = \infty$.

Since $\sigma = \infty$ we may evaluate $\Phi(x)$ directly. In the case where $\alpha \neq r$,

$$\begin{aligned}
\Phi(x_0) &= \lim_{t\to\infty} \int_0^t e^{-rs}\mathbb{E}_{x_0}\left\{X_s\right\} ds \\
&= \lim_{t\to\infty} \int_0^t x_0 e^{(\alpha-r)s} ds \\
&= \lim_{t\to\infty} \frac{x_0}{\alpha - r} e^{(\alpha-r)s}\bigg|_0^t \\
&= \begin{cases} \infty & \text{for} \quad \alpha > r \\ \frac{x_0}{r-\alpha} & \text{for} \quad \alpha < r \end{cases}.
\end{aligned}$$

However, in the case where $\alpha = r$ we have

$$\begin{aligned}
\Phi(x_0) &= \lim_{t\to\infty} \int_0^t x_0 ds \\
&= \infty.
\end{aligned}$$

Hence our complete solution to Equation (2.1) in the case of $f(x) = x$ is given by

$$\Phi(x) = \begin{cases} \infty & \text{for } \alpha \geq r \\ \frac{x}{r-\alpha} & \text{for } \alpha < r \end{cases}. \tag{2.5}$$

2.3 The case of $f(x) = x - K, K > 0$

For $f(x) = x - K$ where $K > 0$ is some known constant, we have the subset of the continuation region as

$$\begin{aligned}\mathcal{U} &= \left\{(x,t) \in \mathbb{R}^+ \times \mathbb{R}^+ : \mathcal{A}_X g(x) + e^{-rt} f(x) > 0\right\} \\ &= \{(x,t) : x > K\}.\end{aligned}$$

Thus we seek a continuation region of the form

$$\mathcal{C} = \{(x,t) : x > d\}$$

for some $0 \leq d < K$. Using the $\mathcal{A}_Z$, we see that the partial differential equation is

$$\tfrac{1}{2}\beta^2 x^2 \phi_{xx} + \alpha x \phi_x + \phi_t + e^{-rt}(x - K) = 0 \tag{2.6}$$

for $x > d$ and 0 otherwise. In this later domain, $0 \leq x \leq d$, we stop immediately and have $\int_0^0 e^{-rs} f(X_s) ds$. That is, $\sigma = 0$. We seek an overall solution of the form $\phi(x,t) = e^{-rt}\psi(x)$. Then the PDE reduces to the ODE

$$\tfrac{1}{2}\beta^2 x^2 \psi'' + \alpha x \psi' - r\psi + (x - K) = 0, \tag{2.7}$$

which is a nonhomogeneous Cauchy-Euler equation. The homogeneous solution is of the form $\psi_h(x) = C_1 x^{\gamma_1} + C_2 x^{\gamma_2}$ where

$$\begin{aligned}\gamma_1 &= \beta^{-2}\left[\tfrac{1}{2}\beta^2 - \alpha + \sqrt{\left(\tfrac{1}{2}\beta^2 - \alpha\right)^2 + 2r\beta^2}\right] > 0, \\ \gamma_2 &= \beta^{-2}\left[\tfrac{1}{2}\beta^2 - \alpha - \sqrt{\left(\tfrac{1}{2}\beta^2 - \alpha\right)^2 + 2r\beta^2}\right] < 0.\end{aligned} \tag{2.8}$$

As we assume that our solution will be continuous and smooth over the boundary d, we use the method of Variation of Parameters to find a particular solution. In standard form, the ODE becomes

$$\psi'' + \tfrac{2\alpha}{\beta^2 x}\psi' - \tfrac{2r}{\beta^2 x^2}\psi = -\frac{2(x-K)}{\beta^2 x^2}$$

Let $y_1(x) = x^{\gamma_1}$, $y_2(x) = x^{\gamma_2}$, and $g(x) = -2(x-K)/(\beta^2 x^2)$. Hence the Wronskian is

$$W[y_1 y_2](x) = (\gamma_2 - \gamma_1)\, x^{\gamma_1+\gamma_2-1}.$$

Denote $\sqrt{\cdot} = \sqrt{\left(\frac{1}{2}\beta^2 - \alpha\right)^2 + 2r\beta^2}$, and assume $r > \alpha$. The assumption of $r \neq \alpha$ guarantees that $\gamma_1 \neq 1$, and hence the following integrals do not result in logarithmic functions.

$$\begin{aligned}
v_1(x) &= -\int \frac{g(x)y_2(x)}{W[y_1y_2](x)}dx \\
&= \frac{-1}{\sqrt{\cdot}}\left[\frac{x^{-\gamma_1+1}}{1-\gamma_1} - K\frac{x^{-\gamma_1}}{-\gamma_1}\right] \\
v_1(x)y_1(x) &= \frac{-1}{\sqrt{\cdot}}\left[\frac{x}{1-\gamma_1} + \frac{K}{\gamma_1}\right] \\
v_2(x) &= \int \frac{g(x)y_1(x)}{W[y_1y_2](x)}dx \\
&= \frac{1}{\sqrt{\cdot}}\left[\frac{x^{-\gamma_2+1}}{1-\gamma_2} - K\frac{x^{-\gamma_2}}{-\gamma_2}\right] \\
v_2(x)y_2(x) &= \frac{1}{\sqrt{\cdot}}\left[\frac{x}{1-\gamma_2} + \frac{K}{\gamma_2}\right]
\end{aligned}$$

Thus the particular solution is

$$\psi_p(x) = v_1 y_1(x) + v_2 y_2(x) = \frac{x}{r-\alpha} - \frac{K}{r}, \tag{2.9}$$

and the general solution to the ODE is

$$\psi(x) = \begin{cases} C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \frac{x}{r-\alpha} - \frac{K}{r} & \text{for} \quad x > d \\ 0 & \text{for} \quad 0 \le x \le d \end{cases} \tag{2.10}$$

If we suppose that there is no finite stopping time, that is $\sigma = \infty$, then $\Phi(x) = \psi_p(x)$. Since we would expect Φ to asymptotically approach this function as $\sigma \to \infty$, we may consider that $C_1 \equiv 0$.

As we have assumed that $\psi(x)$ be continuous and smooth at the boundary d, we must have that

$$\begin{aligned} \psi(x) &= \begin{cases} C_2 x^{\gamma_2} + \frac{x}{r-\alpha} - \frac{K}{r} & \text{for} \quad x > d \\ 0 & \text{for} \quad 0 \le x \le d \end{cases} \\ C_2 &= -\frac{d^{-\gamma_2+1}}{\gamma_2(r-\alpha)} \\ d &= K\frac{\gamma_2(r-\alpha)}{r(\gamma_2-1)} \end{aligned}$$

Thus for $r > \alpha$,

$$\phi(x,t) = \begin{cases} e^{-rt}\left(\frac{-d}{\gamma_2(r-\alpha)}\left(\frac{x}{d}\right)^{\gamma_2} + \frac{x}{r-\alpha} - \frac{K}{r}\right) & \text{for} \quad x > d \\ 0 & \text{for} \quad 0 \le x \le d \end{cases}. \tag{2.11}$$

So at $(x,t) = (x_0, 0)$ we have

$$\phi(x_0) = \begin{cases} \left(\frac{-d}{\gamma_2(r-\alpha)}\left(\frac{x_0}{d}\right)^{\gamma_2} + \frac{x_0}{r-\alpha} - \frac{K}{r}\right) & \text{for} \quad x_0 > d \\ 0 & \text{for} \quad 0 \le x_0 \le d \end{cases}. \tag{2.12}$$

2.4 CPT Approach

For $\Phi(x) = \sup_\sigma \mathbb{E}_x \left\{ \int_0^\sigma e^{-rs} f(X_s) ds \right\}$, rewriting this into the CPT form we have $g \equiv 0$ and $c(x) = -f(x)$. Then for a general f we have

$$\delta(x) = \mathbb{E}_x \left\{ C_\infty \right\} = \mathbb{E}_x \left\{ -\int_0^\infty e^{-rs} f(X_s)\ ds \right\}, \tag{2.13}$$

and we may rewrite the problem according to Lemma 1.8 as the following optimal stopping problem:

$$\Phi(x) = \sup_\sigma \mathbb{E}_x \left\{ e^{-r\sigma} \delta\left(X_\sigma\right) \right\} - \delta(x). \tag{2.14}$$

In the case of $f(x) = x$, we have

$$\begin{aligned} \delta(x) &= -\int_0^\infty e^{(\alpha - r)s} X_s\ ds \\ &= \begin{cases} -\frac{x}{r-\alpha} & \text{for } r > \alpha \\ -\infty & \text{for } r \leq \alpha \end{cases}. \end{aligned}$$

So as before, we shall see that the problem becomes trivial. When $r \leq \alpha$, we have $\Phi(x) \equiv \infty$. When $r < \alpha$ our problem can be rewritten as

$$\Phi(x) = \sup_\sigma \mathbb{E}_x \left\{ -e^{-r\sigma} \frac{X_\sigma}{r-\alpha} \right\} - \left(-\frac{x}{r-\alpha} \right). \tag{2.15}$$

Notice that the problem becomes trivial as we are taking the supremum of a strictly negative process. That is, the optimal stopping time will be ∞ in order to have the discounting reduce δ to zero, and our final value of $\Phi(x)$ will be $x/(r-\alpha)$ as seen in Equation (2.5).

In the case of $f(x) = x - K$, $K > 0$, and $r > \alpha$, we find

$$C_t = \int_0^t e^{-rs} \left(K - X_s\right) ds$$

$$\begin{aligned}
\delta(x) &= \mathbb{E}_x\ C_\infty = \frac{K}{r} - \frac{x}{r-\alpha} \\
\Phi(x) &= \sup_\sigma \mathbb{E}_x \left\{ e^{-r\sigma} \left(\frac{K}{r} - \frac{X_\sigma}{r-\alpha} \right) \right\} - \frac{K}{r} + \frac{x}{r-\alpha} \\
&= \frac{1}{r-\alpha} \sup_\sigma \mathbb{E}_x \left\{ e^{-r\sigma} \left(K \left(\frac{r-\alpha}{r} \right) - X_\sigma \right) \right\} - \frac{K}{r} + \frac{x}{r-\alpha}
\end{aligned}$$

Recall that the infinitessimal generator of X_t is $\mathbb{L}_x f(x) = \frac{1}{2}\beta^2 x^2 f''(x) + \alpha x f'(x)$. By examining the ODE

$$\psi(x) \text{ satisfies } \begin{cases} \mathbb{L}_x \psi(x) - r\psi(x) = 0 & \text{for } x \geq x_0 \\ \psi(x) = K\left(\frac{r-\alpha}{r}\right) - x & \text{for } 0 \leq x < x_0 \end{cases}, \tag{2.16}$$

we guess that the function ψ will take the form Cx^γ and we require that it be continuous and smooth at $x = x_0$. Plugging in this test function, we have that γ is the root of the quadratic

$$\tfrac{1}{2}\beta^2\gamma^2 + \left(\alpha - \tfrac{1}{2}\beta^2\right)\gamma - r = 0.$$

Then γ is as in Equation (2.8). As we seek a solution that approaches $\delta(x)$ asymptotically as $x \to \infty$, we choose γ_2 from Equation (2.8). Continuity at $x = x_0$ implies,

$$\begin{aligned}
Cx_0^{\gamma_2} &= K\left(\frac{r-\alpha}{r}\right) - x_0, \\
C &= x_0^{-\gamma_2}\left[K\left(\frac{r-\alpha}{r}\right) - x_0\right].
\end{aligned}$$

Smoothness at $x = x_0$ implies,

$$\begin{aligned}
C\gamma_2 x_0^{\gamma_2 - 1} &= -1, \\
C &= -\frac{x_0^{1-\gamma_2}}{\gamma_2}.
\end{aligned}$$

Combining these, we see that

$$\begin{aligned} x_0^{-\gamma_2}\left[K\left(\frac{r-\alpha}{r}\right)-x_0\right] &= -\frac{x_0^{1-\gamma_2}}{\gamma_2}, \\ K\left(\frac{r-\alpha}{r}\right) &= x_0\left(\frac{\gamma_2-1}{\gamma_2}\right) \\ x_0 &= K\left(\frac{\gamma_2}{\gamma_2-1}\right)\left(\frac{r-\alpha}{r}\right). \end{aligned}$$

Then our overall solution will be

$$\begin{aligned} \phi(x) &= \frac{1}{r-\alpha}\psi(x) - \frac{K}{r} + \frac{x}{r-\alpha} \\ &= \begin{cases} \frac{1}{r-\alpha}\left(-\frac{x_0^{1-\gamma_2}}{\gamma_2}\right)x^{\gamma_2} - \frac{K}{r} + \frac{x}{r-\alpha} & \text{for } x \geq x_0 \\ \frac{1}{r-\alpha}\left[K\left(\frac{r-\alpha}{r}\right) - x\right] - \frac{K}{r} + \frac{x}{r-\alpha} & \text{for } 0 \leq x < x_0 \end{cases} \\ &= \begin{cases} \frac{-x_0}{\gamma_2(r-\alpha)}\left(\frac{x}{x_0}\right)^{\gamma_2} - \frac{K}{r} + \frac{x}{r-\alpha} & \text{for } x \geq x_0 \\ 0 & \text{for } 0 \leq x < x_0 \end{cases}, \end{aligned} \tag{2.17}$$

which is precisely the solution and boundary as seen in Equation (2.12).

2.5 Optimal Hiring Time and Random Arrival Time Effects

As stated in the general model, the observer has the right to not hire immediately but rather at some starting time $\tau \geq 0$. For the above two cases, $f(x) = x$ and $f(x) = x - K$, $\tau \equiv 0$. The rationale for this is as follows: Assume the candidate arrives at time $t = 0$. As X_t is a nonnegative process, when $f(x) = x$ the value function $\Phi(x) = \sup_{\tau \geq 0, \sigma > 0} \mathbb{E}_x \left\{ \int_\tau^\sigma e^{-rs} X_s \, ds \right\}$ is always accruing positive return so we must have $\tau = 0$ and $\sigma = \infty$ at the supremum. In the case of $f(x) = x - K$, if the starting value is below the threshold derived above, we stop immediately and $\tau = \sigma = 0$. If instead we begin at X_0 inside the continuation region, the integral process is accruing positive value for (t, X_t) for the duration of time that X_t is in the continuation region. As such $\tau = 0$ and $\sigma = \inf\{t > \tau : (t,x) \notin \mathcal{C}\}$.

When we examine more general cases of a single candidate in which the candidate does not arrive at time $t = 0$, but assume that the candidate's arrival is a Poisson distributed arrival time ρ. When considering the instantaneous payoff functions above, we have the following result:

$$\begin{aligned}
\Phi_\rho(x) &= \sup_\sigma \mathbb{E}_x \left[\int_\rho^\sigma e^{-rs} f(X_s)\, ds \right] \\
&= \sup_\sigma \mathbb{E}_x \left[e^{-r\rho} \mathbb{E} \left[\int_\rho^\sigma e^{-r(s-\rho)} f(X_s)\, ds \middle| \mathscr{F}_\rho \right] \right] \\
&= \sup_\sigma \mathbb{E}_x \left[e^{-r\rho} \mathbb{E} \left[\int_0^\sigma e^{-r(u)} f(X_u \circ \theta_\rho)\, du \middle| \mathscr{F}_\rho \right] \right] \\
&= \sup_\sigma \mathbb{E}_x \left[e^{-r\rho} \mathbb{E}_{X_\rho} \left[\int_0^\sigma e^{-r(u)} f(X_u)\, du \right] \right] \\
&= \sup_\sigma \mathbb{E}_x \left[e^{-r\rho} \Phi(X_\rho) \right],
\end{aligned}$$

using the change of variable $s = u + \rho$, and the Strong Markov property. Therefore,

$$\Phi_\rho(x) = \sup_\sigma \mathbb{E}_x \left[e^{-r\rho} \Phi(X_\rho) \right]. \tag{2.18}$$

2.6 Finite Horizon and Portfolio Approach

If instead we wish to examine the problem with horizon $T < \infty$, then another approach becomes necessary. For finite horizon problems, we wish to follow Večeř's [16] technique and rewrite the problem in the form

$$\sup_{\sigma \geq 0} \mathbb{E}_x \left\{ e^{-r\sigma} \overline{X}_\sigma \right\},$$

for some new stochastic process $\overline{X}$. To determine the dynamics of this new process, we use portfolio arguments. Let $\Delta_t = \Delta(t)$ be a function measuring the amount of the portfolio in the "asset" of interest to us. Then for a general function f, we have

$$d\overline{X}_t = \Delta_t d\left(f(X_t)\right) + \left[\overline{X}_t - \Delta_t f(X_t)\right] rdt. \tag{2.19}$$

Let us consider the case of $f(x) = x$. Then we have the self-financing portfolio,

$$d\overline{X}_t = \Delta_t dX_t + \left[\overline{X}_t - \Delta_t X_t\right] rdt. \tag{2.20}$$

Then since $d\left(e^{-rt}\Delta_t X_t\right) - e^{-rt}X_t d\Delta_t = e^{-rt}\left(-r\Delta_t X_t dt + \Delta_t dX_t\right)$, we have

$$\begin{aligned} d\left(e^{-rt}\overline{X}_t\right) &= -re^{-rt}\overline{X}_t dt + e^{-rt}d\overline{X}_t \\ &= -re^{-rt}\overline{X}_t dt + e{-}rt\left[\Delta_t dX_t + \left(\overline{X}_t - \Delta_t X_t\right) rdt\right] \\ &= e^{-rt}\left(\Delta_t dX_t - r\Delta_t X_t dt\right) \\ &= d\left(e^{-rt}\Delta_t X_t\right) - e^{-rt}X_t d\Delta_t. \end{aligned}$$

If we then integrate both sides from 0 to t we obtain,

$$\begin{aligned} \int_0^t d\left(e^{-rs}\overline{X}_s\right) &= \int_0^t d\left(e^{-rs}\Delta_s X_s\right) - \int_0^t e^{-rs}X_s d\Delta_s \\ e^{-rt}\overline{X}_t - \overline{X}_0 &= e^{-rt}\Delta_t X_t - \Delta_0 X_0 - \int_0^t e^{-rs}X_s d\Delta_s \\ e^{-rt}\left(\overline{X}_t - \Delta_t X_t\right) &= \overline{X}_0 - \Delta_0 X_0 - \int_0^t e^{-rs}X_s d\Delta_s. \end{aligned}$$

Letting the initial wealth of the portfolio be $\overline{X}_0 = \Delta_0 X_0$, we obtain

$$e^{-rt}\left(\overline{X}_t - \Delta_t X_t\right) = -\int_0^t e^{-rs}X_s d\Delta_s. \tag{2.21}$$

Recall that the function $\Delta_t = \Delta(t)$ is measuring the amount of an "asset" of interest. By examining different choices of Δ we can construct different integral problems. For example, in the case of $\Delta_t = T - t$ for fixed $T > 0$, $\Delta_0 = T$ and

$d\Delta_t = -dt$, and we have

$$e^{-rt}\left[\overline{X}_t - (T-t)X_t\right] = \int_0^t e^{-rs}X_s ds, \tag{2.22}$$

which is precisely the integral of interest to us in the single candidate case of $f(x) = x$. Such a choice of $\Delta(t)$ and $\overline{X}_0$ allows us to express the continuous discounted payoff without the integral. But if we instead choose $\Delta_t = 1 - t/T$ for fixed $T > 0$, $\Delta_0 = 1$ and $d\Delta_t = -(1/T)dt$, and we have

$$e^{-rt}\left[\overline{X}_t - \left(1 - \frac{t}{T}\right)X_t\right] = \frac{1}{T}\int_0^t e^{-rs}X_s ds, \tag{2.23}$$

which gives us an integral similar to the Asian Option. Here, we have expressed something along the lines of a continuous average up to time T where we have the right to discontinue our process at $t < T$, however the average is still taken over $[0, T]$. That is, $X_t \equiv 0$ on $[t, T]$.

As we have seen from investigation of the subset of the continuation region $\mathcal{U}$ the solution in the case of $f(x) = x$ is trivial, that is $\sigma = \infty$ in the infinite horizon problem, we would expect a similar result here. Intuitively this makes sense as X_t is a nonnegative process and as such its integral must be accruing positive value over time. In fact, from direct calculation we find

$$\begin{aligned}\Phi(x_0) &= \mathbb{E}_{x_0}\left\{\int_0^T e^{-rs}X_s\, ds\right\} \\ &= \int_0^T x_0 e^{(\alpha-r)s}\, ds \\ &= \frac{x_0}{\alpha - r}\left(e^{(\alpha-r)T} - 1\right).\end{aligned} \tag{2.24}$$

In Table 2.1 we compare the results of the least-squares Monte Carlo approach to the solution derived in Equation (2.24). The base Monte Carlo simulation is one of 500 independent paths of a geometric Brownian motion with $\alpha = 0.09$, $\beta = 0.2$, and

$X_0 = 1$. The interest rate is $r = 0.1$, and each year is divided into 12 time periods. The time horizon T is in years. Notice that for all but two of the simulations, the least-squares Monte Carlo approach does yield values fairly close to those from Equation (2.24). Furthermore, it appears from the data that given a sufficient number of paths and long enough time horizon, these simulations will approach the solution to the infinite horizon problem.

Table 2.1: Least-Squares Monte Carlo for $f(x) = x$.

T	LSM Prediction	Solution
1	0.997	0.995
10	9.489	9.516
20	18.501	18.127
30	26.910	25.918
40	30.516	32.968
50	38.953	39.347
60	45.896	45.119
70	48.667	50.341
80	64.847	55.067
90	46.689	59.343
100	63.487	63.212
∞	-	100

However, we may also perform a simulation on the expression containing the portfolio $\overline{X}$: $e^{-rt}\left[\overline{X}_t - (T-t)f(X_t)\right]$. This portfolio simulation approach was handled in two ways, with results of the first method yielding virtually identical results to the simulated integral. The first portfolio method we used took the Monte Carlo simulation for X_t and calculated of $\int_0^t e^{-rs} f(X_s)\, ds$ in the expression, then constructed $\overline{X}_t$ with the expression

$$\overline{X}_t = (T-t)f(X_t) + e^{rt}\int_0^t e^{-rs} f(X_s)\, ds. \qquad (2.25)$$

The second involved using the expression for $d\overline{X}_t$ to create a Monte Carlo simulation for $\overline{X}_t$ more directly via the formula

$$\begin{aligned}\overline{X}_{n+1} &= \overline{X}_n + d\overline{X}_n \\ &= \overline{X}_n + \Delta_n(f(X_{n+1}) - f(X_n)) + \left[\overline{X}_n - \Delta_n f(X_n)\right] \frac{r}{N}\end{aligned} \tag{2.26}$$

where N is the number of time periods per year. That is, $dt = 1/N$. In the following table we report both portfolio approaches along with direct simulation of the integral and the actual solution value.

Table 2.2: Least-Squares Monte Carlo for $f(x) = x$ using both direct simulation and portfolio simulations.

T	LSM Direct	LSM Portfolio 1	LSM Portfolio 2	Solution
1	0.9897	0.9897	0.9887	0.9950
10	9.3141	9.3141	9.2892	9.5163
20	19.4155	19.4155	19.3337	18.1269
30	25.1279	25.1279	24.9529	25.9182
40	32.5869	32.5869	32.2858	32.9680
50	40.6267	40.6267	40.1693	39.3469
60	41.0449	41.0449	40.4512	45.1188
70	50.6306	50.6306	49.7910	50.3415
80	54.3896	54.3896	53.3559	55.0671
90	61.3689	61.3689	60.0702	59.3430
100	65.9508	65.9508	64.3293	63.2121

In Table 2.2 we compare the simulations for both the direct approach and portfolio approaches, with the final column again containing the calculation from the derived solution formula. Column three (Portfolio 1) uses $\overline{X}_t$ generated from the method in Equation (2.25), while column four (Portfolio 2) uses $\overline{X}_t$ generated with the method from Equation (2.26). All parameters were the same between the two approaches, and as reported in the description of Table 2.1.

We can make one immediate observation by comparing columns two and three of Table 2.2: They are identical. As Portfolio 1 uses an algebraically equivalent statement to the simulated integral process, Portfolio 1 will therefore yield the same value as the simulated integral process.

Now let us state the problem for a general function $f(x)$.

As differentiation of $d\left(e^{-rt}\Delta_t f(X_t)\right)$ yields the relationship

$$d\left(e^{-rt}\Delta_t f(X_t)\right) - e^{-rt} f(X_t) d\Delta_t = e^{-rt}\left[\Delta_t df(X_t) - r\Delta_t f(X_t)dt\right],$$

we therefore have for a general function $f(x)$ that

$$\begin{aligned} d\left(e^{-rt}\overline{X}_t\right) &= -re^{-rt}\overline{X}_t dt + e^{-rt} d\overline{X}_t \\ &= e{-}rt\left[\Delta_t df(X_t) - r\Delta_t f(X_t)dt\right] \\ &= d\left(e^{-rt}\Delta_t f(X_t)\right) - e^{-rt} f(X_t) d\Delta_t. \end{aligned}$$

By integrating both sides of this expression from 0 to t, we see that

$$\begin{aligned} \int_0^t d\left(e^{-rs}\overline{X}_s\right) &= \int_0^t d\left(e^{-rs}\Delta_s f(X_s)\right) - \int_0^t e^{-rs} f(X_s) d\Delta_s \\ e^{-rt}\overline{X_t} - \overline{X_0} &= e^{-rt}\Delta_t f(X_t) - \Delta_0 f(X_0) - \int_0^t e^{-rs} f(X_s) d\Delta_s. \end{aligned}$$

Choosing $\overline{X}_0 = \Delta_0 f(X_0)$ and $\Delta_t = T - t$, we then have that $d\Delta_s = -ds$ and

$$\int_0^t e^{-rs} f(X_s) ds = e^{-rt}\left[\overline{X}_t - (T-t) f(X_t)\right], \tag{2.27}$$

and hence $\Phi(x) = \sup_\sigma \mathbb{E}_x\left\{\int_0^\sigma e^{-rs} f(X_s) ds\right\}$ is equivalent to the problem

$$\overline{\Phi}(\overline{x}, x) = \sup_\sigma \mathbb{E}_{\overline{x},x}\left\{e^{-r\sigma}\left[\overline{X}_\sigma - (T-\sigma) f(X_\sigma)\right]\right\}. \tag{2.28}$$

Notice that we have effectively increased the dimension of the problem. While we do not pursue this line of investigation very far, we believe that investigation of cases with more general functions $f(x)$ in finite horizon problems may be instructive, particularly in the case of numerical simulations. In addition, the ability to choose

$\Delta_t = \Delta(t)$, as shown earlier, can yield strategies for formulating problems with quite different interpretations.

However, using our Least Squares Monte Carlo approach from earlier, we are able to numerically gather some information on the behavior of Φ in the case of $f(x) = x - K$ in the finite horizon case. As before, we report the results for both direct simulation of the problem and for the portfolio problem.

Table 2.3: Least-Squares Monte Carlo simulations for $f(x) = x - K$, using both direct simulation and portfolio simulation.

T	LSM Direct	LSM Portfolio 1	LSM Portfolio 2
1	52.4373	52.4373	52.5302
20	1412.2776	1412.2776	1410.1104
40	2805.0327	2805.0327	2793.9662
60	4085.5901	4085.5901	4059.7381
80	4817.0127	4817.0127	4773.6173
100	5858.2174	5858.2174	5791.4500

In Table 2.3, we have a Monte Carlo simulation with 500 paths, 24 time divisions per year, $\alpha = 0.09$, $\beta = 0.1$, $r = 0.1$, $X_0 = x = 100$, and $K = 50$. T is the time horizon. As before, column two gives data for the simulated integral, column three gives the portfolio simulation via Equation (2.25), and column three gives the portfolio simulation via Equation (2.26). The infinite horizon problem in this scenario is given by

$$\phi(x) = \begin{cases} \frac{-d}{\gamma_2(r-\alpha)} \left(\frac{x}{d}\right)^{\gamma_2} + \frac{x}{r-\alpha} - \frac{K}{r} & \text{for } x \geq d \\ 0 & \text{for } 0 \leq x < d \end{cases}, \tag{2.29}$$

where $d = 4.7383$ and $\gamma_2 = -18.1047$. $\Phi(X_0)$ then has the value 9500 for $X_0 = 100$, $\alpha = 0.09$, $\beta = 0.2$, $r = 0.1$, and $K = 50$. As was the case for $f(x) = x$, we believe that the prediction is converging to the infinite horizon value as $T \to \infty$.

CHAPTER 3: TWO VARIABLE SWITCHING

3.1 Finite Horizon: Portfolio Approach

In the two candidate case, we consider their instantaneous value modeled by the dynamics

$$dX_{1,t} = X_{1,t}\left(\alpha_1 dt + \beta_1 dW_{1,t}\right), X_{1,0} = x_1,$$
$$dX_{2,t} = X_{2,t}\left(\alpha_2 dt + \beta_2 dW_{2,t}\right), X_{2,0} = x_2,$$

where $W_{1,t}, W_{2,t}$ are two Brownian motions and $dW_{1,t}dW_{2,t} = \rho dt$. First we consider

$$\Phi(x_1, x_2) = \sup_{\sigma} \mathbb{E}_{x_1,x_2}\left\{\int_0^{\sigma} e^{-rs}X_{1,s}\ ds + \int_{\sigma}^{T} e^{-rs}X_{2,s}\ ds\right\}, \tag{3.1}$$

and construct the self-financing portfolio $\overline{X}_t$ with $\Delta_{i,t}$ indicating the amount of $X_{i,t}$ in the portfolio at time t, $0 \le t \le T$. We construct the portfolio's dynamics by

$$d\overline{X}_t = \Delta_{1,t}dX_{1,t} + \Delta_{2,t}dX_{2,t} + \left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right) rdt. \tag{3.2}$$

Since

$$d\left(e^{-rt}\Delta_{i,t}X_{i,t}\right) = -re^{-rt}\Delta_{i,t}X_{i,t}dt + e^{-rt}X_{i,t}d\Delta_{i,t} + e^{-rt}\Delta_{i,t}dX_{i,t}$$
$$\Rightarrow d\left(e^{-rt}\Delta_{i,t}X_{i,t}\right) - e^{-rt}X_{i,t}d\Delta_{i,t} = -re^{-rt}\Delta_{i,t}X_{i,t}dt + e^{-rt}\Delta_{i,t}dX_{i,t},$$

and we have

$$d\left(e^{-rt}\overline{X}_t\right) = d\left(e^{-rt}\Delta_{1,t}X_{1,t}\right) - e^{-rt}X_{1,t}d\Delta_{1,t} + d\left(e^{-rt}\Delta_{2,t}X_{2,t}\right) - e^{-rt}X_{2,t}d\Delta_{2,t}. \quad (3.3)$$

By integrating from 0 to t, we obtain

$$\begin{aligned}\int_0^t d\left(e^{-rs}\overline{X}_s\right) &= \int_0^t d\left(e^{-rt}\Delta_{1,t}X_{1,t}\right) - \int_0^t e^{-rt}X_{1,t}d\Delta_{1,t} \\ &\quad + \int_0^t d\left(e^{-rt}\Delta_{2,t}X_{2,t}\right) - \int_0^t e^{-rt}X_{2,t}d\Delta_{2,t}, \\ e^{-rt}\overline{X}_t - \overline{X}_0 &= e^{-rt}\Delta_{1,t}X_{1,t} - \Delta_{1,0}X_{1,0} + e^{-rt}\Delta_{2,t}X_{2,t} - \Delta_{2,0}X_{2,0} \\ &\quad - \int_0^t e^{-rt}X_{1,t}d\Delta_{1,t} - \int_0^t e^{-rt}X_{2,t}d\Delta_{2,t}.\end{aligned}$$

By setting $\overline{X}_0 = \Delta_{1,0}X_{1,0} + \Delta_{2,0}X_{2,0}$ and rearranging terms, we have

$$e^{-rt}\left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right) = -\int_0^t e^{-rs}X_{1,s}\ d\Delta_{1,s} - \int_0^t e^{-rs}X_{2,s}\ d\Delta_{2,s}. \quad (3.4)$$

Notice that since

$$\begin{aligned}\Phi(x_1,x_2) &= \mathbb{E}_{x_1,x_2}\left\{\int_0^t e^{-rs}X_{1,s}\ ds + \int_t^T e^{-rs}X_{2,s}\ ds\right\} \\ &= \mathbb{E}_{x_1,x_2}\left\{\int_0^t e^{-rs}X_{1,s}\ ds - \int_0^t e^{-rs}X_{2,s}\ ds + \int_0^T e^{-rs}X_{2,s}\right\} \\ &= \frac{X_{2,0}}{\alpha_2 - r}\left(e^{(\alpha_2-r)T} - 1\right) + \mathbb{E}_{x_1,x_2}\left\{\int_0^t e^{-rs}X_{1,s}\ ds - \int_0^t e^{-rs}X_{2,s}\ ds\right\},\end{aligned}$$

we may choose $\Delta_{1,t} = T - t$ and $\Delta_{2,t} = t - T$ to obtain

$$\begin{aligned}&\mathbb{E}_{x_1,x_2}\left\{-\int_0^t e^{-rs}X_{1,s}\ d\Delta_{1,s} - \int_0^t e^{-rs}X_{2,s}\ d\Delta_{2,s}\right\} \\ &\quad = \mathbb{E}_{x_1,x_2}\left\{\int_0^t e^{-rs}X_{1,s}\ ds - \int_0^t e^{-rs}X_{2,s}\ ds\right\},\end{aligned}$$

and thus, denoting $\overline{x} = \overline{X}_0$, we have

$$\begin{aligned}\overline{\Phi}(\overline{x}, x_1, x_2) &= \mathbb{E}_{\overline{x},x_1,x_2}\left\{e^{-rt}\left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right)\right\} \\ &= \Phi(x_1, x_2) - \frac{x_2}{\alpha_2 - r}\left(e^{(\alpha_2 - r)T} - 1\right). \end{aligned} \tag{3.5}$$

To proceed, we require the following theorem.

Theorem 3.10 ([15])**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let Z be an almost surely nonnegative random variable with $\mathbb{E}Z = 1$. For $A \in \mathscr{F}$, define*

$$\widetilde{\mathbb{P}}(A) = \int_Z Z(\omega) d\mathbb{P}(\omega). \tag{3.6}$$

Then $\widetilde{\mathbb{P}}$ is a probability measure. Furthermore, if X is a nonnegative random variable, then

$$\widetilde{\mathbb{E}}X = \mathbb{E}\left[XZ\right]. \tag{3.7}$$

If Z is almost surely strictly positive, we also have

$$\mathbb{E}Y = \widetilde{\mathbb{E}}\left[\frac{Y}{Z}\right] \tag{3.8}$$

for every nonnegative random variable Y.

As $d\left(X_{1,t}X_{2,t}\right) = X_{1,t}X_{2,t}\left[(\alpha_1 + \alpha_2 + \beta_1\beta_2\rho)dt + \beta_1 dW_{1,t} + \beta_2 dW_{2,t}\right]$, letting

$$Z_t = \frac{e^{-(\alpha_1+\alpha_2+\beta_1\beta_2\rho)t}X_{1,t}X_{2,t}}{X_{1,0}X_{2,0}}, \tag{3.9}$$

yields a martingale starting at 1.

Definition 3.12 (Radon-Nikodým Derivative [15])**.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $\widetilde{\mathbb{P}}$ be another probability measure on $(\Omega, \mathscr{F})$ that is equivalent to $\mathbb{P}$, and let Z be an almost surely positive random variable that relates $\mathbb{P}$ to $\widetilde{\mathbb{P}}$ via (3.6). Then

Z is called the *Radon-Nikodým derivative* of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$, and we write

$$Z = \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}.$$

Thus our choice of Z_t is a Radon-Nikodým derivative, and we have for $0 \le s \le t$,

$$\begin{aligned}
&\mathbb{E}_{\overline{x},x_1,x_2}\left\{ e^{-rt}\left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right) \middle| \mathscr{F}_s \right\} \\
&= \mathbb{E}^X_{\overline{x},x_1,x_2}\left\{ \frac{1/Z_t}{1/Z_s} e^{-rt}\left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right) \middle| \mathscr{F}_s \right\} \\
&= e^{-(\alpha_1+\alpha_2+\beta_1\beta_2\rho)s} X_{1,s}X_{2,s} \\
&\quad \cdot \mathbb{E}^X_{\overline{x},x_1,x_2}\left\{ e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)t}\left(\frac{\overline{X}_t}{X_{1,t}X_{2,t}} - \frac{\Delta_{1,t}}{X_{2,t}} - \frac{\Delta_{2,t}}{X_{1,t}}\right) \middle| \mathscr{F}_s \right\}.
\end{aligned}$$

where $\mathbb{E}^X$ indicates expectation with respect to the new probability measure $\mathbb{P}^X$. So for $s = 0$, we have

$$\begin{aligned}
&\mathbb{E}_{\overline{x},x_1,x_2}\left\{ e^{-rt}\left(\overline{X}_t - \Delta_{1,t}X_{1,t} - \Delta_{2,t}X_{2,t}\right) \right\} \\
&\quad = X_{1,0}X_{2,0}\mathbb{E}^X_{\overline{x},x_1,x_2}\left\{ e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)t}\left(\frac{\overline{X}_t}{X_{1,t}X_{2,t}} - \frac{\Delta_{1,t}}{X_{2,t}} - \frac{\Delta_{2,t}}{X_{1,t}}\right) \right\}.
\end{aligned}$$

We may rewrite the main problem as

$$\Phi^X(y, x_1, x_2) = \sup_\sigma \mathbb{E}^X_{y,x_1,x_2}\left\{ e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)\sigma}\left(\frac{\overline{X}_\sigma}{X_{1,\sigma}X_{2,\sigma}} - \frac{\Delta_{1,\sigma}}{X_{2,\sigma}} - \frac{\Delta_{2,\sigma}}{X_{1,\sigma}}\right) \right\}, \quad (3.10)$$

where $Y_t = \dfrac{\overline{X}_t}{X_{1,t}X_{2,t}}$, $y = \dfrac{\overline{x}}{x_1x_2} = (T-t)\dfrac{x_1 - x_2}{x_1x_2} = (T-t)\left(\dfrac{1}{x_2} - \dfrac{1}{x_1}\right)$, and $\Phi^X(y, x_1, x_2) = \dfrac{1}{x_1x_2}\overline{\Phi}(\overline{x}, x_1, x_2) = \dfrac{1}{x_1x_2}\left(\Phi(x_1, x_2) - \dfrac{x_2}{\alpha_2 - r}\left(e^{(\alpha_2-r)T} - 1\right)\right)$.

To adjust our Brownian motion terms to our new probability measure, we require Girsanov's Theorem.

Theorem 3.11 (Girsanov's Theorem [15]). *Let T be a fixed positive time, and let*

$\Theta(t) = (\Theta_1(t), \ldots, \Theta_d(t))$ be a d-dimensional adapted process. Define

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)\cdot dW(u) - \tfrac{1}{2}\int_0^t \|\Theta(u)\|^2\, du\right\}, \tag{3.11}$$

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)du, \tag{3.12}$$

and assume that

$$\mathbb{E}\int_0^T \|\Theta(u)\|^2 Z^2(u)du < \infty. \tag{3.13}$$

Set $Z = Z(T)$. Then $\mathbb{E}Z = 1$, and under the probability measure $\widetilde{\mathbb{P}}$ given by

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)d\mathbb{P}(\omega) \text{ for all } A \in \mathscr{F},$$

the process $\widetilde{W}(t)$ is a d-dimensional Brownian motion.

Rewriting Z_t with the independent Brownian motions $B_{i,t}$ where $W_{1,t} = B_{1,t}$ and $W_{2,t} = \rho B_{1,t} + \sqrt{1-\rho^2}B_{2,t}$, we have

$$\begin{aligned}
Z_t &= \exp\left\{-\int_0^t -\beta_1 dW_{1,s} - \int_0^t -\beta_2 dW_{2,s} - \tfrac{1}{2}\int_0^t \left(\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho\right) ds\right\} \\
&= \exp\left\{-\int_0^t -\left(\beta_1 + \beta_2\rho\right) dB_{1,s} - \int_0^t \beta_2\sqrt{1-\rho^2}dB_{2,s}\right. \\
&\qquad \left. -\tfrac{1}{2}\int_0^t \left(\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho\right) ds\right\} \\
&= \exp\left\{-\int_0^t \begin{bmatrix} -(\beta_1 + \beta_2\rho) \\ -\beta_2\sqrt{1-\rho^2} \end{bmatrix} \cdot dB_s - \tfrac{1}{2}\int_0^t \left\| \begin{bmatrix} -(\beta_1 + \beta_2\rho) \\ -\beta_2\sqrt{1-\rho^2} \end{bmatrix} \right\|^2 ds\right\}.
\end{aligned}$$

Thus $\widetilde{B}_t = dB_t + \begin{bmatrix} -(\beta_1 + \beta_2\rho) \\ -\beta_2\sqrt{1-\rho^2} \end{bmatrix} dt$ by Girsanov's Theorem.

To develop a strategy, we consider the dynamics of $e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)t}\left(\frac{\overline{X}_t}{X_{1,t}X_{2,t}} - \frac{\Delta_{1,t}}{X_{2,t}} - \frac{\Delta_{2,t}}{X_{1,t}}\right)$:

$$d\left[e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)t}\left(\frac{\overline{X}_t}{X_{1,t}X_{2,t}}-\frac{\Delta_{1,t}}{X_{2,t}}-\frac{\Delta_{2,t}}{X_{1,t}}\right)\right]$$
$$=\frac{e^{(\alpha_1+\alpha_2+\beta_1\beta_2\rho-r)t}}{X_{1,t}X_{2,t}}\Big\{\left(\Delta_{1,t}X_{1,t}+\Delta_{2,t}X_{2,t}\right)dt$$
$$+\left(\overline{X}_t-\Delta_{1,t}X_{1,t}-\Delta_{2,t}X_{2,t}\right)\left[d\widetilde{B}_{1,t}\left(-\beta_1-\beta_2\rho\right)+d\widetilde{B}_{2,t}\left(-\beta_2\sqrt{1-\rho^2}\right)\right]\Big\}, \tag{3.14}$$

where $d\widetilde{B}_{1,t}=dW_{1,t}-(\beta_1+\beta_2\rho)dt$, $d\widetilde{B}_{2,t}=\frac{1}{\sqrt{1-\rho^2}}\left(dW_{2,t}-\rho dW_{1,t}-\beta_2(1-\rho^2)dt\right)$ by the above application of Girsanov's Theorem.

While it is enlightening to see the dynamics of the problem from this perspective, it only lets us know that there is a subset of the continuation region by examining where the drift,$(T-t)X_{1,t}+(t-T)X_{2,t}$, is positive. That is, our subset of the continuation region takes the form

$$\begin{aligned}\mathcal{U}&=\{(t,x_1,x_2):(T-t)(X_{1,t}-X_{2,t})>0\}\\&=\{(t,x_1,x_2):X_{2,t}<X_{1,t}\}.\end{aligned} \tag{3.15}$$

We may, however, run numerical simulations using Least-Squares Monte Carlo. As before, we construct the portfolio in two ways. The first via

$$\overline{X}_t=(T-t)(X_{1,t}-X_{2,t})+e^{rt}\int_0^t e^{-rs}(X_{1,s}-X_{2,s})\,ds, \tag{3.16}$$

and the second via

$$dt=1/N$$
$$\overline{X}_0=(T-0)(X_{1,0}-X_{2,0}),$$
$$d\overline{X}_n=(T-n\cdot dt)(X_{1,n+1}-X_{2,n+1}-X_{1,n}+X_{2,n})$$

$$+\left(\overline{X}_n - (T - n \cdot dt)(X_{1,n} - X_{2,n})\right) r \cdot dt,$$

$$\overline{X}_{n+1} = \overline{X}_n + d\overline{X}_n. \tag{3.17}$$

The technique of Least-Squares Monte Carlo requires repeated use of regression techniques to estimate the expected value of continuing. In the single variable case, we followed Longstaff & Schwartz [10] in that we regressed on a degree two polynomial. To do so here required changing to a new independent regression variable for $\overline{X}$: $X_t = X_{1,t} - X_{2,t}$. However, we also chose to examine the possibility of using multilinear regression with $X_{1,t}$ and $X_{2,t}$ as the independent variables. The results between the two techniques, for both portfolio construction approaches, yielded identical results, but the multilinear regression had significantly increased computation times. Both sets of code are provided in the appendix, but the results reported below are from the simulations with X_t.

Table 3.4: Least-Squares Monte Carlo simulations for two variable switching: Both methods and average stopping times for each method.

T	LSM Portfolio 1	Avg. σ_1	LSM Portfolio 2	Avg. σ_2
1	1.02021752002	0.2285	1.01292236985	0.22933333
10	9.54787902156	0.5905	9.54961813294	0.579666667
20	18.1756148417	0.95115	18.1884767641	1.000666667
30	27.5330375199	7.02	27.5217077105	7.02
40	35.2490346596	8.64	35.2284128281	8.64
50	42.5558133353	10.9	42.5182654108	10.7

In Table 3.4 500 paths and 12 time periods per year were simulated for $X_{1,t}$ and $X_{2,t}$ where $\alpha_1 = 0.05$, $\beta_1 = 0.2$, $\alpha_2 = 0.09$, $\beta_2 = 0.2$, $r = 0.1$, $\rho = 0.5$, $X_{1,0} = X_{2,0} = 1$. In addition to the final value Φ for each method, we also report the average value of σ for each method. Note that for some paths, $\sigma = 0$ or $\sigma = T$.

In seeking analytical solutions to the two variable switching problem, we now return our attention to the infinite horizon case and use the CPT method.

3.2 Rewriting the Problem via CPT

Let us consider the infinite horizon problem; that is $T = \infty$:

$$\Phi(x_1, x_2) = \sup_{\sigma} \mathbb{E}_{x_1,x_2} \left\{ \int_0^{\sigma} e^{-rs} X_{1,s}\, ds + \int_{\sigma}^{\infty} e^{-rs} X_{2,s}\, ds \right\}. \tag{3.18}$$

Denoting $C_{i,t} = \int_0^t -e^{-rs} X_{i,s}\, ds$ and $\delta_i(x_i) = \mathbb{E}_{x_i}\, C_{i,\infty}$, we apply the following proposition.

Proposition 3.12. *If $\delta_1(x_1)$ and $\delta_2(x_2)$ are finite on the domain E, then for any $(x_1, x_2) \in E$ Equation (3.18) may be rewritten as*

$$\Phi(x_1, x_2) = \sup_{\sigma \geq 0} \mathbb{E}_{x_1,x_2} \left\{ e^{-r\sigma} \left(\delta_1(X_{1,\sigma}) - \delta_2(X_{2,\sigma}) \right) \right\} - \delta_1(X_{1,0}), \tag{3.19}$$

where σ is any stopping time in the set of all stopping times.

Proof:

Using the identity $C_{i,\infty} = C_{i,\sigma} + e^{-r\sigma} C_{i,\infty} \circ \theta_\sigma$, iterated conditioning and the Strong Markov property we find

$$\begin{aligned}
&\mathbb{E}_{x_1,x_2} \left\{ \mathbb{E} \left\{ -\int_0^{\sigma} -e^{-rs} X_{1,s}\, ds - \int_{\sigma}^{\infty} -e^{-rs} X_{2,s}\, ds \middle| \mathscr{F}_\sigma \right\} \right\} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ \mathbb{E} \left\{ -\int_0^{\sigma} -e^{-rs} X_{1,s}\, ds + \int_0^{\sigma} -e^{-rs} X_{2,s}\, ds \right.\right. \\
&\qquad \left.\left. - \int_0^{\infty} -e^{-rs} X_{2,s}\, ds \middle| \mathscr{F}_\sigma \right\} \right\} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ \mathbb{E} \left\{ -C_{1,\sigma} + (C_{2,\sigma} - C_{2,\infty}) \middle| \mathscr{F}_\sigma \right\} \right\} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ \mathbb{E} \left\{ -C_{1,\infty} + e^{-r\sigma} C_{1,\infty} \circ \theta_\sigma + (-e^{-r\sigma} C_{2,\sigma} \circ \theta_\sigma) \middle| \mathscr{F}_\sigma \right\} \right\} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ e^{-r\sigma} \mathbb{E} \left\{ C_{1,\infty} \circ \theta_\sigma - C_{2,\infty} \circ \theta_\sigma \middle| \mathscr{F}_\sigma \right\} - \mathbb{E}_{X_{1,\sigma}}\, C_{1,\infty} \right\} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ e^{-r\sigma} \mathbb{E}_{X_{1,\sigma}, X_{2,\sigma}} \left\{ C_{1,\infty} - C_{2,\infty} \right\} \right\} - \mathbb{E}_{x_1,x_2}\, C_{1,\infty} \\
&\quad = \mathbb{E}_{x_1,x_2} \left\{ e^{-r\sigma} \left(\delta_1(X_{1,\sigma}) - \delta_2(X_{2,\sigma}) \right) \right\} - \delta_1(X_{1,0}),
\end{aligned}$$

as desired. □

Under the assumption that $\alpha_1, \alpha_2 < r$, both δ_1 and δ_2 are finite and

$$\begin{aligned}\delta_i(x_i) = \mathbb{E}_{x_i}\ C_{i,\infty} &= \int_0^\infty -e^{-rs}\ \mathbb{E}_{x_i}\ X_{i,s}\ ds \\ &= -\frac{x_i}{r-\alpha_i}.\end{aligned} \tag{3.20}$$

And thus we rewrite the problem of Equation (3.18) into

$$\Phi(x_1, x_2) = \sup_\sigma \mathbb{E}_{x_1,x_2}\left\{e^{-r\sigma}\left(\frac{X_{2,\sigma}}{r-\alpha_2} - \frac{X_{1,\sigma}}{r-\alpha_1}\right)\right\} + \frac{X_{1,0}}{r-\alpha_1}. \tag{3.21}$$

For convenience of notation, let us denote $g(x_1, x_2) = \left(\frac{x_2}{r-\alpha_2} - \frac{x_1}{r-\alpha_2}\right)$ and $\hat{\Phi}(x_1, x_2)$ as $\sup_\sigma \mathbb{E}_{x_1,x_2}\{e^{-r\sigma} g(X_{1,\sigma}, X_{2,\sigma})\}$. Then it is sufficient to optimize $\hat{\Phi}$.

3.3 Infinite Horizon: PDE Approach

The infinitesimal generator for this two dimensional problem, with the assumed dynamics, is given by

$$\mathbb{L}f(x_1, x_2) = \tfrac{1}{2}\beta_1^2 x_1^2 \frac{\partial^2 f}{\partial x_1^2} + \beta_1\beta_2\rho x_1 x_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \tfrac{1}{2}\beta_2^2 x_2^2 \frac{\partial^2 f}{\partial x_2^2} + \alpha_1 x_1 \frac{\partial f}{\partial x_1} + \alpha_2 x_2 \frac{\partial f}{\partial x_2} \tag{3.22}$$

for a twice differentiable function f. We seek a solution $\hat{\Phi}$ of the form

$$\phi(x_1, x_2) = \begin{cases} \psi(x_1, x_2) & \text{for } x_2 < \mu x_1 \\ g(x_1, x_2) & \text{for } x_2 \geq \mu x_1 \end{cases}, \tag{3.23}$$

and a continuation region of the form $\mathcal{C} = \{(x_1, x_2) : x_2 < \mu x_1\}$. Our solution $\psi \in C^2(D)$ will satisfy

$$\mathbb{L}\psi(x_1, x_2) = r\psi(x_1, x_2) \quad \text{for } x_2 < \mu x_1,$$

$$\begin{aligned}
\psi(x_1,x_2) &= g(x_1,x_2) \quad \text{for } x_2 = \mu x_1,\\
\nabla\psi(x_1,x_2) &= \nabla g(x_1,x_2) \quad \text{for } x_2 = \mu x_1,\\
\mathbb{L}g(x_1,x_2) &\le r g(x_1,x_2) \quad \text{for } x_2 > \mu x_1,\\
\psi(x_1,x_2) &> g(x_1,x_2) \quad \text{for } x_2 < \mu x_1,
\end{aligned}$$

and we guess that ψ will take the form $\psi(x_1,x_2) = Cx_1^{1-\lambda}x_2^{\lambda}$ for some constants $C, \lambda > 0$. When placed into the equation $\mathbb{L}\psi(x_1,x_2) = r\psi(x_1,x_2)$, we see that

$$\begin{aligned}
\lambda = &\frac{\frac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \frac{1}{2}\beta_2^2 + \alpha_1 - \alpha_2}{\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2}\\
&+ \frac{\sqrt{(\frac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \frac{1}{2}\beta_2^2 + \alpha_1 - \alpha_2)^2 + 4(\frac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \frac{1}{2}\beta_2^2)(r-\alpha_1)}}{\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2}. \quad (3.24)
\end{aligned}$$

Denote $b = \frac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \frac{1}{2}\beta_2^2$ and $a = \alpha_1 - \alpha_2$. Then we have

$$\lambda = \frac{b + a + \sqrt{(b+a)^2 + 4b(r-\alpha_1)}}{2b}.$$

For $\alpha_1, \alpha_2 < r$ and $\rho \in [-1, 1]$, this λ is real.

Examining $\mathbb{L}g(x_1,x_2)$, we find

$$\begin{aligned}
\mathbb{L}g(x_1,x_2) &= \alpha_1 x_1 \frac{-1}{r-\alpha_1} + \alpha_2 x_2 \frac{1}{r-\alpha_2}\\
&< r\left(\frac{x_2}{r-\alpha_2} - \frac{x_1}{r-\alpha_1}\right),
\end{aligned}$$

which holds for $g > 0$ automatically as, by the prior assumptions necessary for $C_{1,\infty}, C_{2,\infty}$ to be finite, we have $\alpha_1, \alpha_2 < r$. The function g is positive for $\frac{x_2}{r-\alpha_2} > \frac{x_1}{r-\alpha_1}$, i.e. $x_2 > x_1\left(\frac{r-\alpha_2}{r-\alpha_1}\right)$, and so we can expect that μ will be proportional to $\frac{r-\alpha_2}{r-\alpha_1}$.

At $x_2 = \mu x_1$ we have

$$\psi(x_1, \mu x_1) = Cx_1^{1-\lambda}(\mu x_1)^{\lambda} = C\mu^{\lambda} x_1$$

$$g(x_1, x_2) = x_1 \left(\frac{\mu}{r - \alpha_2} - \frac{1}{r - \alpha_1} \right)$$
$$\Rightarrow C = \mu^{-\lambda} \left(\frac{\mu}{r - \alpha_2} - \frac{1}{r - \alpha_1} \right),$$

and

$$\nabla\psi(x_1, \mu x_1) = C(1-\lambda)x_1^{-\lambda}(\mu x_1)^{\lambda}\vec{e}_1 + C\lambda x_1^{1-\lambda}(\mu x_1)^{\lambda-1}\vec{e}_2$$
$$= C(1-\lambda)\mu^{\lambda}\vec{e}_1 + C\lambda\mu^{\lambda-1}\vec{e}_2$$
$$\nabla g(x_1, \mu x_1) = \frac{-1}{r-\alpha_1}\vec{e}_1 + \frac{1}{r-\alpha_2}\vec{e}_2$$
$$\Rightarrow \begin{cases} C(1-\lambda)\mu^{\lambda} = \frac{-1}{r-\alpha_1} \\ C\lambda\mu^{\lambda-1} = \frac{1}{r-\alpha_2} \end{cases}$$

Combining this information, we have

$$C = \frac{\mu^{-\lambda}}{\lambda - 1} \left(\frac{1}{r - \alpha_1} \right), \tag{3.25}$$
$$\mu = \frac{\lambda}{\lambda - 1} \left(\frac{r - \alpha_2}{r - \alpha_1} \right). \tag{3.26}$$

So our overall solution to equation (3.18) is

$$\phi(x_1, x_2) = \begin{cases} \frac{1}{\lambda-1} \left(\frac{x_1}{r-\alpha_1} \right) \left(\frac{x_2}{\mu x_1} \right)^{\lambda} + \frac{x_1}{r-\alpha_1} & \text{for } x_2 < \mu x_1, \\ \frac{x_2}{r-\alpha_2} & \text{for } x_2 \geq \mu x_1. \end{cases} \tag{3.27}$$

In Figure 3.1, we see a three dimensional surface plot of $\phi(x_1, x_2)$ as each variable ranges from 0 to 100. The scale of each axes is in units of 10. As is evident from the graph, the surface slopes down sharply as $(X_{1,0}, X_{2,0})$ approach $(0, 0)$. It also slopes up sharply as either $X_{1,0}$ or $X_{2,0}$ approach 100. There is no significant amount of curvature visible in the graph. However, we do find visible evidence of curvature in the next plot.

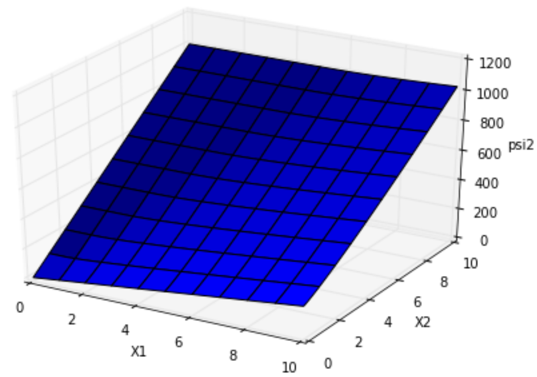

Figure 3.1: A sample plot of $\phi(x_1, x_2)$ in which $\alpha_1 = 0.05$, $\alpha_2 = 0.09$, $\beta_1 = 0.3$, $\beta_2 = 0.1$, and $\rho = 0.1$.

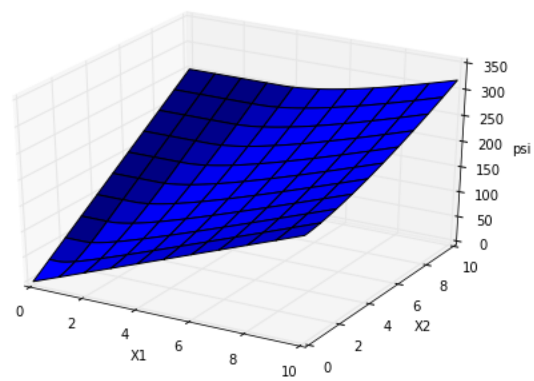

Figure 3.2: A sample plot of $\phi(x_1, x_2)$ in which $\alpha_1 = 0.05$, $\alpha_2 = 0.06$, $\beta_1 = 0.3$, $\beta_2 = 0.4$, and $\rho = 0.5$.

In Figure 3.2 we have another sample plot of $\phi(x_1, x_2)$ over the same range, but the processes $X_{1,t}$ and $X_{2,t}$ are assumed to have dynamics that are similar to one another. While the end behavior slope is similar to the last plot, we see some evidence of curvature as $X_{1,0}$ and $X_{2,0}$ increase together.

3.4 Infinite Horizon: Change of Numeraire Approach

Beginning from the transformed problem, Equation (3.21), we define $Z_t = e^{-\alpha_1 t} X_{1,t}/X_{1,0}$.

As Z_t is a positive martingale starting at one, it satisfies the hypothesis of the Theorem 3.10.

Since the conditions on Z_t of Theorem 3.10 are satisfied, let $\widetilde{\mathbb{P}}$ be as in (3.6). Then

Z_t is the Radon-Nikodým derivative of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. Further, since the process $X_{2,t}$ is nonnegative, we have for $t < T$

$$\begin{aligned}\mathbb{E}_{x_1,x_2}&\left\{\frac{Z_t}{Z_T}e^{-rT}\left(\frac{X_{2,T}}{r-\alpha_2}-\frac{X_{1,T}}{r-\alpha_1}\right)\bigg|\mathscr{F}_t\right\}\\&=\left(\frac{e^{-\alpha_1 t}X_{1,t}}{r-\alpha_2}\right)\widetilde{\mathbb{E}}_{x_1,x_2}\left\{e^{(\alpha_1-r)T}\left(\frac{X_{2,t}}{X_{1,t}}-\frac{r-\alpha_2}{r-\alpha_1}\right)\bigg|\mathscr{F}_t\right\}.\end{aligned}$$

Letting $Y_t = X_{2,t}/X_{1,t}$, $Y_0 = y = x_2/x_1$, we have for $t = 0$

$$\Phi(x_1,x_2) = \left(\frac{x_1}{r-\alpha_2}\right)\sup_{\sigma}\widetilde{\mathbb{E}}_y\left\{e^{(\alpha_1-r)\sigma}\left(Y_\sigma-\frac{r-\alpha_2}{r-\alpha_1}\right)\right\}+\frac{x_1}{r-\alpha_1}. \tag{3.28}$$

Let us denote $\Phi^Y(y)$ as

$$\Phi^Y(y) = \sup_{\sigma}\widetilde{\mathbb{E}}_y\left\{e^{(\alpha_1-r)\sigma}\left(Y_\sigma-\frac{r-\alpha_2}{r-\alpha_1}\right)\right\}. \tag{3.29}$$

The stochastic process Y_t has dynamics

$$dY_t = Y_t\left[(-\alpha_1+\alpha_2+\beta_1^2-\beta_1\beta_2\rho)dt-\beta_1 dW_{1,t}+\beta_1 dW_{2,t}\right], \tag{3.30}$$

which may be verified by either application of Itô's formula to $X_{2,t}/X_{1,t}$ or by direct calculation since $Y_t = y\exp\left\{(\alpha_2-\alpha_1-\frac{1}{2}\beta_2+\frac{1}{2}\beta_1)t-\beta_1 W_{1,t}+\beta_2 W_{2,t}\right\}$.

For the problem at hand, we require independent Brownian motions to proceed so let $W_{1,t} = \sqrt{1-\rho^2}B_{1,t}+\rho B_{2,t}$ and $W_{2,t} = B_{2,t}$. These Brownian motions are independent as $dW_{1,t}dW_{2,t} = dt$. Let $B(t) = (B_{1,t}, B_{2,t})$. Then

$$\begin{aligned}Z_t &= \exp\left\{-\int_0^t -\beta_1 dW_{1,s}-\tfrac{1}{2}\int_0^t \beta_1^2 ds\right\}\\&=\exp\left\{-\int_0^t\begin{bmatrix}-\beta_1\sqrt{1-\rho^2}\\-\beta_1\rho\end{bmatrix}dB(s)-\tfrac{1}{2}\int_0^t\left\|\begin{bmatrix}-\beta_1\sqrt{1-\rho^2}\\-\beta_1\rho\end{bmatrix}\right\|^2 ds\right\},\end{aligned}$$

so we have under after the change of measure

$$d\widetilde{B}_{1,t} = dB_{1,t} - \beta_1\sqrt{1-\rho^2}dt = \frac{dW_{1,t} - \rho dW_{2,t}}{\sqrt{1-\rho^2}} - \beta_1\sqrt{1-\rho^2}dt,$$
$$d\widetilde{B}_{2,t} = dB_{2,t} - \beta_1\rho dt = dW_{2,t} - \beta_1\rho dt.$$

That is,

$$dW_{1,t} = \sqrt{1-\rho^2}d\widetilde{B}_{1,t} + \rho d\widetilde{B}_{2,t} + \beta_1 dt,$$
$$dW_{2,t} = d\widetilde{B}_{2,t} + \beta_1\rho dt.$$

Thus we have for the dynamics of Y_t after the change of measure

$$dY_t = Y_t\left[(-\alpha_1+\alpha_2)dt - \beta_1\sqrt{1-\rho^2}d\widetilde{B}_{1,t} + (-\beta_1\rho+\beta_2)d\widetilde{B}_{2,t}\right].$$

For convenience, we wish to express the Brownian motion terms as a single Brownian motion.

$$\begin{aligned}
(cd\widetilde{B}_{3,t})(cd\widetilde{B}_{3,t}) &= \left(-\beta_1\sqrt{1-\rho^2}d\widetilde{B}_{1,t} + (-\beta_1\rho+\beta_2)d\widetilde{B}_{2,t}\right)^2 \\
&= \left(\beta_1^2(1-\rho^2) + \beta_1^2\rho^2 - 2\beta_1\beta_2\rho + \beta_2^2\right)dt \\
&= \left(\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2\right)dt \\
&\Rightarrow c = \sqrt{\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2}.
\end{aligned}$$

Then $\widetilde{B}_{3,t}$ is a Brownian motion starting at 0 and

$$\widetilde{B}_{3,t} = \frac{-\beta_1\sqrt{1-\rho^2}d\widetilde{B}_{1,t} + (-\beta_1\rho+\beta_2)d\widetilde{B}_{2,t}}{\sqrt{\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2}}. \tag{3.31}$$

We then discover that the dynamics of Y_t take the form

$$dY_t = Y_t \left[(-\alpha_1 + \alpha_2) dt + \sqrt{\beta_1^2 - 2\beta_1\beta_2\rho + \beta_2^2}\, d\widetilde{B}_{3,t} \right], \tag{3.32}$$

which yields the infinitesimal generator

$$\mathbb{L} f(y) = (-\alpha_1 + \alpha_2)\, y \tfrac{d}{dy} f(y) + \left(\tfrac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \tfrac{1}{2}\beta_2^2\right) y^2 \tfrac{d^2}{dy^2} f(y) \tag{3.33}$$

for any twice differentiable function f. As $\Phi(x_1, x_2) = \frac{x_1}{r-\alpha_2}\Phi^Y(y) + \frac{x_1}{r-\alpha_1}$, it suffices to optimize Φ^Y. To this end, we assume a solution of the form $\phi^Y(t, y) = e^{(\alpha_1 - r)t}\psi^Y(y)$ which leads us to examine the ordinary differential equation

$$\begin{cases} (\alpha_1 - r)\,\psi^Y(y) + (-\alpha_1 + \alpha_2)\, y \frac{d}{dy}\psi^Y(y) & \\ \qquad + \left(\frac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \frac{1}{2}\beta_2^2\right) y^2 \frac{d^2}{dy^2}\psi^Y(y) = 0 & \text{for } 0 \le y < y_0, \\ \psi^Y(y) = \left(y - \frac{r-\alpha_2}{r-\alpha_1}\right) & \text{for } y \ge y_0, \end{cases}$$

and we look for a solution of the form Cy^λ. Plugging this function into the above yields precisely the value as in Equation (3.24). Further, since it we seek solutions that are bounded as $y \to 0+$, as before we will only consider the root $\lambda = (b + a + \sqrt{(b+a)^2 + 2b(r - \alpha_1)})/(2b)$. As we want our solution to be continuous at y_0, we have

$$\begin{aligned} Cy_0^\lambda &= y_0 - \frac{r - \alpha_2}{r - \alpha_1}, \\ C &= y_0^{-\lambda}\left(y_0 - \frac{r - \alpha_2}{r - \alpha_1}\right), \\ \Rightarrow \psi^Y(y) &= \left(y_0 - \frac{r - \alpha_2}{r - \alpha_1}\right)\left(\frac{y}{y_0}\right)^\lambda. \end{aligned}$$

And since we seek a solution that is smooth at y_0, we have

$$\begin{aligned}
&\lambda\left(\frac{1}{y_0}\right)\left(y_0-\frac{r-\alpha_2}{r-\alpha_1}\right)\left(\frac{y_0}{y_0}\right)^{\lambda-1}=1,\\
&\lambda\left(y_0-\frac{r-\alpha_2}{r-\alpha_1}\right)=y_0,\\
&y_0\left(\lambda-1\right)=\lambda\left(\frac{r-\alpha_2}{r-\alpha_1}\right),\\
&y_0=\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{r-\alpha_2}{r-\alpha_1}\right).
\end{aligned}$$

Notice that this is precisely the same threshold as in the prior approach, there denoted μ. We now demonstrate that the two approaches, when written out as the full value function ϕ, yield algebraically equivalent functions. For the piecewise domains, and noting that both $x_1, x_2 > 0$,

$$\begin{aligned}
x_2<\mu x_1 &\Rightarrow \frac{x_2}{x_1}<\mu\\
&\Rightarrow 0\le y<\mu=y_0,\\
x_2\ge\mu x_1 &\Rightarrow \frac{x_2}{x_1}\ge\mu\\
&\Rightarrow y\ge\mu=y_0.
\end{aligned}$$

The complete value function for this method, $\phi^Y(x_1,x_2)$, is

$$\begin{aligned}
\phi^Y(x_1,x_2) &= \begin{cases} \left(\frac{x_1}{r-\alpha_2}\right)\left(y_0-\frac{r-\alpha_2}{r-\alpha_1}\right)\left(\frac{y}{y_0}\right)^{\lambda}+\frac{x_1}{r-\alpha_1} & \text{for } 0\le y<y_0\\ \left(\frac{x_1}{r-\alpha_2}\right)\left(y-\frac{r-\alpha_2}{r-\alpha_1}\right)+\frac{x_1}{r-\alpha_1} & \text{for } y\ge y_0 \end{cases},\\
&= \begin{cases} \left(\frac{x_1}{r-\alpha_2}\right)\left(\mu-\frac{r-\alpha_2}{r-\alpha_1}\right)\left(\frac{x_2/x_1}{\mu}\right)^{\lambda}+\frac{x_1}{r-\alpha_1} & \text{for } 0\le\frac{x_2}{x_1}<\mu\\ \left(\frac{x_1}{r-\alpha_2}\right)\left(\frac{x_2}{x_1}-\frac{r-\alpha_2}{r-\alpha_1}\right)+\frac{x_1}{r-\alpha_1} & \text{for } \frac{x_2}{x_1}\ge\mu \end{cases},\\
&= \begin{cases} \frac{1}{\lambda-1}\left(\frac{x_1}{r-\alpha_1}\right)\left(\frac{x_2}{\mu x_1}\right)^{\lambda}+\frac{x_1}{r-\alpha_1} & \text{for} x_2<\mu x_1\\ \left(\frac{x_2}{r-\alpha_2}-\frac{x_1}{r-\alpha_1}\right)+\frac{x_1}{r-\alpha_1} & \text{for } x_2\ge\mu x_1 \end{cases},
\end{aligned}$$

where in the final equality we take advantage of the fact that

$$\begin{aligned}\frac{x_1}{r-\alpha_2}\left(\mu-\frac{r-\alpha_2}{r-\alpha_1}\right) &= \frac{x_1}{r-\alpha_2}\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{r-\alpha_2}{r-\alpha_1}\right)-\frac{x_1}{r-\alpha_1}\\ &= \frac{x_1}{r-\alpha_1}\left(\frac{\lambda}{\lambda-1}-1\right)\\ &= \frac{x_1}{r-\alpha_1}\left(\frac{1}{\lambda-1}\right).\end{aligned}$$

CHAPTER 4: VERIFICATION THEOREMS

In the final chapter we will work through theorems verifying our derived solutions to the single candidate $f(x) = x - K$ case and the two variable switching case.

4.1 Single Candidate, $f(x) = x - K$

Recall that the optimal stopping problem for this case was

$$\Phi_*(x) = \sup_{\sigma} \mathbb{E}_x \left\{ \int_0^{\sigma} e^{-rs} \left(X_s - K \right) \, ds \right\}, \tag{4.1}$$

where $X = (X_t)_{t \geq 0}$ satisfies $dX_t = X_t \left(\alpha dt + \beta dW_t \right)$, and that the infinitessimal generator of X is given by $\mathbb{L}_X f = \frac{1}{2}\beta^2 x^2 f'' + \alpha x f'$.

Theorem 4.13. *The solution to Equation* (4.1) *is given by*

$$\Phi(x) = \begin{cases} \frac{-d^{1-\gamma}}{\gamma(r-\alpha)} x^{\gamma} + \frac{x}{r-\alpha} - \frac{K}{r} & \textit{for } x > d \\ 0 & \textit{for } 0 \leq x \leq d \end{cases}, \tag{4.2}$$

where

$$d = K \frac{\gamma(r-\alpha)}{r(\gamma-1)},$$
$$\gamma = \beta^{-2} \left[\tfrac{1}{2}\beta^2 - \alpha - \sqrt{\left(\tfrac{1}{2}\beta^2 - \alpha\right)^2 + 2r\beta^2} \right].$$

Furthermore, the stopping time

$$\sigma_d = \inf\{t \geq 0 : X_t \leq d\} \tag{4.3}$$

is optimal.

Proof:

It is relatively easy to verify that Φ from Equation (4.2) satisfies the variational inequality

$$\begin{cases} \mathbb{L}_X\phi(x) - r\phi(x) = -(x-K) & \text{for } x > d \\ \phi(x) = 0 & \text{for } 0 \leq x \leq d \\ \phi(d) = 0 & \text{continuity at } d \\ \phi_x(d) = 0 & \text{smoothness at } d \end{cases}.$$

Applying Itô's formula from Peskir & Shiryaev [13], Equation (1.5), we find that

$$\begin{aligned} &\int_0^t e^{-rs}(X_s - K) + e^{-rt}\Phi(X_t) \\ &= \int_0^t e^{-rs}(X_s - K)\ ds + \Phi(x) + \int_0^t e^{-rs}\left[\mathbb{L}_X\Phi - r\Phi\right](X_s)I(X_s \neq d)\ ds \\ &\quad + \int_0^t e^{-rs}\beta X_s\Phi'(X_s)I(X_s \neq d)\ dW_s \\ &\quad + \tfrac{1}{2}\int_0^t e^{-rs}\left(\Phi'(X_s+) - \Phi'(X_s-)\right)I(X_s = d)d\ell_s^c(X) \\ &= \Phi(x) + \int_0^t e^{-rs}(X_s - K)\ ds + \int_0^t e^{-rs}\left[\mathbb{L}_X\Phi - r\Phi\right](X_s)I(X_s \neq d)\ ds \\ &\quad + \int_0^t e^{-rs}\beta X_s\Phi'(X_s)I(X_s \neq d)\ dW_s, \end{aligned}$$

as Φ is smooth at d.

Notice that by Φ being a solution the variational equality, we have that $(\mathbb{L}_X\Phi - r\Phi)(X_t) = -X_t + K$ on the set $\{(t, X_t) : X_t \geq d\}$ and 0 otherwise, and that $X_t - K < 0$ on the set $\{(t, X_t) : 0 \leq X_t \leq d\}$. Therefore, we must have that everywhere on $(0, \infty)$

but d

$$\int_0^t e^{-rs}(X_s - K)\ ds + \int_0^t e^{-rs}\left[\mathbb{L}_X\Phi - r\Phi\right](X_s)I(X_s \neq d)\ ds \leq 0.$$

However, $\mathbb{P}_x(X_s = d) = 0$ for all x and s.

Denote $M = (M_t)_{t\geq 0}$ by $M_t = \int_0^t e^{-rs}\beta X_2\Phi'(X_s)I(X_s \neq d)\ dW_s$. This M_t is a continuous local martingale.

Let $(\sigma_n)_{n\geq 1}$ be a localization sequence of bounded stopping times for M. Then as

$$\int_0^t e^{-rs}(X_s - K)\ ds \leq \int_0^t e^{-rs}(X_s - K)\ ds + e^{-rt}\Phi(X_t) \leq \Phi(x) + M_t,$$

we have for every stopping time σ of X

$$\int_0^{\sigma\wedge\sigma_n} e^{-rs}(X_s - K)\ ds \leq \Phi(x) + M_{\sigma\wedge\sigma_n}.$$

Taking $\mathbb{P}_x$-expectation and using the Optional Sampling Theorem we must conclude that $\mathbb{E}_x M_{\sigma\wedge\sigma_n} = 0$ for all n, and as $n \to \infty$ we find by Fatou's Lemma

$$\mathbb{E}_x\left\{\int_0^{\sigma} e^{-rs}(X_s - K)\ ds\right\} \leq \Phi(x).$$

By taking the supremum over all stopping times σ of X we have $\Phi_*(x) \leq \Phi(x)$. But by the Optional Sampling Theorem we also have

$$\mathbb{E}_x\left\{\int_0^{\sigma_d\wedge\sigma_n} e^{-rs}(X_s - K)\ ds + e^{-r(\sigma_d\wedge\sigma_n)}\Phi(X_{\sigma_d\wedge\sigma_n})\right\} = \Phi(x)$$

for all $n \geq 1$. Letting $n \to \infty$, we notice that $\Phi(X_{\sigma_d}) \equiv 0$ by construction and we have by the Dominated Convergence Theorem that

$$\mathbb{E}_x\left\{\int_0^{\sigma_d} e^{-rs}(X_s - K)\ ds\right\} = \Phi(x),$$

and so σ_d is optimal, and $\Phi = \Phi_*$ for all $x > 0$. □

4.2 Two Variable Switching

For convenience, we shall state the equivalent optimization problem after the CPT transformation. Recall that the optimal stopping problem after the transformation is given by

$$\Phi_*(x_1, x_2) = \sup_{\sigma} \mathbb{E}_{x_1,x_2} \left\{ e^{-r\sigma} \left(\frac{X_{2,\sigma}}{r - \alpha_2} - \frac{X_{1,\sigma}}{r - \alpha_1} \right) \right\}. \tag{4.4}$$

Theorem 4.14. *The solution to Equation* (4.4) *is given by*

$$\Phi(x_1, x_2) = \begin{cases} \frac{1}{\mu^{\lambda}(\lambda-1)(r-\alpha_1)}(x_1)^{1-\lambda}(x_2)^{\lambda} & \text{for } x_2 < \mu x_1 \\ \frac{x_2}{r-\alpha_2} - \frac{x_1}{r-\alpha_1} & \text{for } x_2 \geq \mu x_1 \end{cases}, \tag{4.5}$$

where

$$a = \alpha_1 - \alpha_2,$$
$$b = \tfrac{1}{2}\beta_1^2 - \beta_1\beta_2\rho + \tfrac{1}{2}\beta_2^2,$$
$$\lambda = \frac{b + a + \sqrt{(b+a)^2 + 4b(r - \alpha_1)}}{2b},$$
$$\mu = \frac{\lambda}{\lambda - 1}\left(\frac{r - \alpha_2}{r - \alpha_1}\right),$$

and the switching time

$$\sigma_\mu = \inf\{t \geq 0 : X_{2,t} \geq \mu X_{1,t}\} \tag{4.6}$$

is optimal.

Proof:

Denote the function $g(x_1, x_2) = \frac{x_2}{r-\alpha_2} - \frac{x_1}{r-\alpha_1}$. Any solution $\phi(x_1, x_2)$ to Equation

(4.4) will satisfy the variational inequality

$$\begin{cases} \mathbb{L}_{x_1,x_2}\phi - r\phi = 0 & \text{for } x_2 < \mu x_1 \\ \phi = g & \text{for } x_2 = \mu x_1 \\ \nabla\phi = \nabla g & \text{for } x_2 = \mu x_1 \\ \mathbb{L}_{x_1,x_2} g - rg \leq 0 & \text{for } x_2 > \mu x_1 \\ \phi > g & \text{for } x_2 < \mu x_1 \end{cases}$$

where $\mathbb{L}_{x_1,x_2} f = \frac{1}{2}\beta_1^2 x_1^2 \frac{\partial^2 f}{\partial x_1^2} + \beta_1\beta_2\rho x_1 x_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{1}{2}\beta_2^2 x_2^2 \frac{\partial^2 f}{\partial x_2^2} + \alpha_1 x_1 \frac{\partial f}{\partial x_1} + \alpha_2 x_2 \frac{\partial f}{\partial x_2}$, and $dX_{1,t} = X_{1,t}(\alpha_1 dt + \beta_1 dB_{1,t})$ and $dX_{2,t} = X_{2,t}(\alpha_2 dt + \beta_2\rho dB_{1,t} + \beta_2\sqrt{1-\rho^2}dB_{2,t})$ for $B_{1,t}$ and $B_{2,t}$ independent Brownian Motions.

By Itô's formula we have

$$\begin{aligned} e^{-rt}\Phi(X_{1,t}, X_{2,t}) = \Phi(x_1, x_2) & \\ & + \int_0^t e^{-rs}\left(\mathbb{L}_{x_1,x_2}\Phi - r\Phi\right)(X_{1,s}, X_{2,s}) I(X_{2,s} \neq \mu X_{1,s})\; ds \\ & \int_0^t e^{-rs} \begin{bmatrix} \beta_1 + \beta_2\rho \\ \beta_2\sqrt{1-\rho^2} \end{bmatrix} \cdot \begin{bmatrix} X_{1,s}\frac{\partial\Phi}{\partial x_1} dB_{1,s} \\ X_{2,s}\frac{\partial\Phi}{\partial x_2} dB_{2,s} \end{bmatrix} I(X_{2,s} \neq \mu X_{1,s}), \end{aligned} \tag{4.7}$$

as Φ is smooth at $X_{2,t} = \mu X_{1,t}$. We notice that $\left(\mathbb{L}_{x_1,x_2}\Phi - r\Phi\right)(X_{1,t}, X_{2,t}) \leq 0$ on $\mathbb{R}_+$ by construction.

Denote $M = (M_t)_{t\geq 0}$ by

$$M_t = \int_0^t e^{-rs} \begin{bmatrix} \beta_1 + \beta_2\rho \\ \beta_2\sqrt{1-\rho^2} \end{bmatrix} \cdot \begin{bmatrix} X_{1,s}\frac{\partial\Phi}{\partial x_1}(X_{1,s}, X_{2,s}) dB_{1,s} \\ X_{2,s}\frac{\partial\Phi}{\partial x_2}(X_{1,s}, X_{2,s}) dB_{2,s} \end{bmatrix} I(X_{2,s} \neq \mu X_{1,s}).$$

Since,

$$\frac{\partial\Phi}{\partial x_1}(X_{1,t},X_{2,t}) = \begin{cases} \frac{-1}{r-\alpha_1}\left(\frac{X_{2,t}}{\mu X_{1,t}}\right)^{\lambda} & \text{for } X_{2,t} < \mu X_{1,t} \\ \frac{-1}{r-\alpha_1} & \text{for } X_{2,t} \geq \mu X_{1,t} \end{cases}$$

$$\frac{\partial\Phi}{\partial x_2}(X_{1,t},X_{2,t}) = \begin{cases} \frac{1}{r-\alpha_2}\left(\frac{X_{2,t}}{\mu X_{1,t}}\right)^{\lambda-1} & \text{for } X_{2,t} < \mu X_{1,t} \\ \frac{1}{r-\alpha_2} & \text{for } X_{2,t} \geq \mu X_{1,t} \end{cases}$$

and $\lambda > 0$, these partial derivatives are bounded. M_t is a continuous local martingale. Let $(\sigma_n)_{n\geq 1}$ be a localization sequence of bounded stopping times for M. Then since

$$e^{-rt}\left(\frac{X_{2,t}}{r-\alpha_2} - \frac{X_{1,t}}{r-\alpha_1}\right) \leq e^{-rt}\Phi(X_{1,t},X_{2,t}) \leq \Phi(x_1,x_2) + M_t. \tag{4.8}$$

Then for any switching time σ of $(X_{1,t},X_{2,t})$ we have

$$e^{-r(\sigma\wedge\sigma_n)}\left(\frac{X_{2,\sigma\wedge\sigma_n}}{r-\alpha_2} - \frac{X_{1,\sigma\wedge\sigma_n}}{r-\alpha_1}\right) \leq \Phi(x_1,x_2) + M_{\sigma\wedge\sigma_n} \tag{4.9}$$

for all $n \geq 1$. Taking $\mathbb{P}_{x_1,x_2}$-expectation and using the Optional Sampling Theorem we must conclude that $\mathbb{E}_{x_1,x_2}M_{\sigma\wedge\sigma_n}$ for all n and as $n \to \infty$ we find by Fatou's Lemma

$$\mathbb{E}_{x_1,x_2}\left\{e^{-r\sigma}\left(\frac{X_{2,\sigma}}{r-\alpha_2} - \frac{X_{1,\sigma}}{r-\alpha_1}\right)\right\} \leq \Phi(x_1,x_2). \tag{4.10}$$

Taking the supremum over all switching times σ of $(X_{1,t},X_{2,t})$ we have $\Phi_* \leq \Phi$. But by the Optional Sampling Theorem we have

$$\mathbb{E}_{x_1,x_2}\left\{e^{-r(\sigma_\mu\wedge\sigma_n)}\Phi(X_{1,\sigma_\mu\wedge\sigma_n},X_{2,\sigma_\mu\wedge\sigma_n})\right\} = \Phi(x_1,x_2) \tag{4.11}$$

for all $n \geq 1$. Letting $n \to \infty$ we notice that $\Phi = g$ at the boundary by construction and we have by the Dominated Convergence Theorem that

$$\mathbb{E}_{x_1,x_2}\left\{e^{-r\sigma_\mu}\left(\frac{X_{2,\sigma_\mu}}{r-\alpha_2}-\frac{X_{1,\sigma_\mu}}{r-\alpha_1}\right)\right\}=\Phi(x_1,x_2), \tag{4.12}$$

and so σ_μ is optimal and $\Phi = \Phi_*$ for all $x > 0$. □

REFERENCES

[1] Brekke, Kjell Arne and Bernt Øksendal, 1994: Optimal switching in an economic activity under uncertainty. *SIAM J. Control and Optimization*, **34** no. 4, 1021-1036.

[2] Carmona, René and Nizar Touzi, 2008: Optimal multiple stopping and valuation of swing options. Mathematical Finance, **18** no. 2, 239-268.

[3] Chung, Kai Lai, 2001: A Course in Probability Theory, 3rd Edition. Academic Press.

[4] Cissé, Mamadou, Pierre Patie, and Etienne Tanré, 2012: Optimal stopping problems for some Markov processes. *The Annals of Applied Probability*, **22** no. 3, 1243-1265.

[5] Dahlgren, Eric and Tim Leung, 2015: An optimal multiple stopping approach to infrastructure investment decisions. *Journal of Economic Dynamics & Control*, **53**, 251-267.

[6] Ferguson, Thomas S., 1989: Who Solved the Secretary Problem? *Statistical Science*, **4** no. 3, 282-296.

[7] Freeman, P. R., 1983: The secretary problem and its extensions: a review. *International Statistical Review*, **51** (1983), 189-206.

[8] Haggstrom, Gus W., 1967: Optimal sequential procedures when more than one stop is required. *Annals of Mathematical Statistics*, **38**, no. 6, (1967), 1618-1626.

[9] Karatzas, Ioannis and Steven E. Shreve, 1991: Brownian Motion and Stochastic Calculus, 2nd Edition. Springer-Verlag.

[10] Longstaff, Francis A. and Eduardo S. Schwartz, 2001: Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studites* Spring 2014, **14**, no. 1, pp. 113-147.

[11] Meyer, P.-A. (1966): *Probability and Potentials*. Blaisdell Publishing Co., Ginn and Co., Waltham-Toronto-London.

[12] Øksendal, Bernt, 2003: *Stochastic Differential Equations: An Introduction with Applications*. 6th Edition. Springer, 369 pp.

[13] Peskir, Goran and Albert Shiryaev, 2006: *Optimal Stopping and Free Boundary Problems*. ETH Zurich, 500 pp.

[14] Revuz, Daniel and Marc Yor, 1999: *Continuous Martingales and Brownian Motion*. 3rd Edition. Springer-Verlag.

[15] Shreve, Steven E., 2004: *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer-Verlag, 550 pp.

[16] Večeř, Jan, 2002: Unified Pricing of Asian options. *Risk*, **15**, no. 6, pp. 113-116.

APPENDIX A: SINGLE CANDIDATE CODE

The first appendix will contain all Python code responsible for the single candidate materials, while the second will contain code responsible for the two candidate materials. The Python code used throughout was generated using Python 3.5 through the Anaconda distribution, available at `https://www.continuum.io/downloads`. Please understand that, as Python is a white space language, certain formatting changes were necessary to fit the code within the margins of the document. That is, the scripts will not run as they are presented below. When carriage returns have been inserted there were immediately followed by tabs. It is our hope that this information when combined with some familiarity with the language, or at least an error checking IDE, that any reader will be able to replicate our results with little trouble.

The following is the code that generated the data seen in Table 2.1.

```
# LSM for f(x) = x, single candidate.


import numpy as np

import scipy as spy


def MonteCarlo(M,N,T,S0,A,B):

    dt = 1/N

    S = np.zeros((M,int(T*N)+1))

    S[:,0] = S0

    eps = np.random.normal(0, 1, (M,int(N*T)))

    S[:,1:] = np.exp((A-0.5*B**2)*dt + eps*B*np.sqrt(dt));

    S = np.cumprod(S, axis = 1);

    return S
```

```
def LSM(M,N,T,S,Z,R):
    C = np.zeros((M,int(T*N)+1))
    for m in range(M):
        C[m,T*N] = np.max([S[m,T*N],0]) #K-S[m,T*N],0])
    X = np.zeros((M,int(T*N)))
    Y = np.zeros((M,int(T*N)))
    Exercise = np.zeros((M,int(T*N)))
    Continue = np.zeros((M,int(T*N)))
    for n in range(int(T*N),1,-1):
        x = np.zeros(0)
        y = np.zeros(0)
        for i in range(M):
            if S[i,n-1] > 0: #K-S[i,n-1] > 0:
                X[i,n-1] = Z[i,n-1]
                 # independent variable of regression
                 # should be the underlying Monte
                 # Carlo simulation
                Exercise[i,n-1] = S[i,n-1]
                 # Exercise value if exercise now.
                x = np.append(x,X[i,n-1])
                Y[i,n-1] = C[i,n] #df * C[i,n]
                y = np.append(y,Y[i,n-1])
        if len(x) == 0:
            p = np.array([0,0,0])
        else:
            p = spy.polyfit(x,y,2)
        for i in range(M):
```

```
            if S[i,n-1] > 0: #K-S[i,n-1] > 0:
             # Expected value of continuing,
             # calculated with degreee 2 regression.
                Continue[i,n-1] =
                 p[0]*X[i,n-1]**2 + p[1]*X[i,n-1] + p[2]
        for i in range(M):
         # Exercise now only if expected value
         # of continuing is negative.
            if Continue[i,n-1] < 0:
                C[i,n-1] = Exercise[i,n-1]
                C[i,n:] = 0
    return C

# Parameters for running the model.
M = 500
N = 12
alpha = 0.09
beta = 0.3
r = 0.1
Z0 = 1
K = 0

t = np.linspace(10,100,10)

# Store Output for Table 2.1.
Output = np.zeros((11,3))
T = 1
```

```
# Build the Monte Carlo simulation for X when T=1.
Z = MonteCarlo(M,N,T,Z0,alpha,beta)

# Calculation of integral of e^{-rs}(X_s - K) along each path
# Integration with trapezoidal rule (for non-uniform widths)
dt = 1/N
S = np.zeros((M,T*N+1))
for m in range(M):
    for n in range(T*N):
        S[m,n+1] =
         S[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*(Z[m,n+1]-K)
         +np.exp(-r*dt*n)*(Z[m,n]-K));

# Run LSM for T=1.
C = LSM(M,N,T,S,Z,r)

Output[0,:] =
T, np.mean(C[:,N*T]), Z0/(alpha-r)*(np.exp((alpha-r)*T)-1)

for k in range(len(t)):
    T = int(t[k])
    Z = MonteCarlo(M,N,T,Z0,alpha,beta)
    S = np.zeros((M,T*N+1))
    for m in range(M):
        for n in range(int(T*N)):
            S[m,n+1] =
```

```
            S[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))
            *(Z[m,n+1]-K)+np.exp(-r*dt*n)*(Z[m,n]-K));
    C = LSM(M,N,T,S,Z,r) #, cont, exer
    Output[k+1,:] =
     T, np.mean(C[:,N*T]), Z0/(alpha-r)*(np.exp((alpha-r)*T)-1)
```

The following is the code used in generating the data for Table 2.3.

```
#LSM for x-K, single candidate


import numpy as np
import scipy as spy
import pylab as pl


def MonteCarlo(M,N,T,S0,A,B):
    dt = 1/N
    S = np.zeros((M,int(T*N)+1))
    S[:,0] = S0
    eps = np.random.normal(0, 1, (M,int(N*T)))
    S[:,1:] = np.exp((A-0.5*B**2)*dt + eps*B*np.sqrt(dt));
    S = np.cumprod(S, axis = 1);
    return S


def LSM(M,N,T,S,Z,R):
    C = np.zeros((M,int(T*N)+1))
    for m in range(M):
        C[m,T*N] = np.max([S[m,T*N],0]) #K-S[m,T*N],0])
    X = np.zeros((M,int(T*N)))
```

```
Y = np.zeros((M,int(T*N)))
Exercise = np.zeros((M,int(T*N)))
Continue = np.zeros((M,int(T*N)))
for n in range(int(T*N),1,-1):
    x = np.zeros(0)
    y = np.zeros(0)
    for i in range(M):
        if S[i,n-1] > 0:
            X[i,n-1] = Z[i,n-1]
             # independent variable of regression
             # should be the underlying Monte
             # Carlo simulation.
            Exercise[i,n-1] = S[i,n-1]
             # exercise value if
             # exercise now (current
             # value of integral).
            x = np.append(x,X[i,n-1])
            Y[i,n-1] = C[i,n]
            y = np.append(y,Y[i,n-1])
    if len(x) == 0:
        p = np.array([0,0,0])
    else:
        p = spy.polyfit(x,y,2)
    for i in range(M):
        if S[i,n-1] > 0: #K-S[i,n-1] > 0:
            Continue[i,n-1] = p[0]*X[i,n-1]**2
             + p[1]*X[i,n-1] + p[2]
```

```
                 # Expected value of continuing,
                 # calculated with
                 # degreee 2 regression.
        for i in range(M):
            if Continue[i,n-1] < 0:
             # Exercise now if negative expected
             # value of continuing
             # from the current point.
                C[i,n-1] = Exercise[i,n-1]
                C[i,n:] = 0
    return C, Continue, Exercise

def Sigma(M,N,T,C):
    sigma = np.zeros(M)
    for m in range(M):
        for n in range(int(T*N)):
            if C[m,n+1] != 0:
                sigma[m] = n+1
    return sigma

# Parameters for running the model.
M = 500
T = 10
N = 24
alpha = 0.09
beta = 0.1
r = 0.1
```

```
Z0 = 100
K = 50

# Build the Monte Carlo simulation for X.
Z = MonteCarlo(M,N,T,Z0,alpha,beta)

# Calculation of integral of e^{-rs}(X_s - K) along each path
# Integration with trapezoidal rule (for non-uniform widths)
dt = 1/N
S = np.zeros((M,T*N+1))
for m in range(M):
    for n in range(T*N):
        S[m,n+1] =
         S[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*(Z[m,n+1]-K)
         +np.exp(-r*dt*n)*(Z[m,n]-K));

# Run LSM
C , cont, exer = LSM(M,N,T,S,Z,r)

sigma = np.zeros(M)
#sigma = np.flatnonzero(C[0,:])[0]

for i in range(M):
    if C[i,:].any() != 0:
        sigma[i] = np.flatnonzero(C[i,:])[0]

cvec = np.zeros(M)
```

```
for i in range(M):
    for j in range(N*T+1):
        if C[i,j] != 0:
            cvec[i] = C[i,j]


if K == 0:
    calculated = np.mean(C[:,N*T])
    predicted = Z0/(alpha-r)*(np.exp((alpha-r)*T)-1)
    print('As K=0 was chosen, the following are the calculated
     and predicted values:')
    print('Calculated by averaging final values:', calculated)
    print('Predicted by direct calculation:', predicted)
```

Here we have the code used for the single candidate portfolio data.

```
#LSM for f(x) = x-K, single candidate, portfolio simulations included.


import numpy as np
import scipy as spy
import pylab as pl
from sklearn import linear_model


def f(x,K):
    return x-K


def delta(t,T):
    return T-t
```

```
def phi(Z,barZ,t,T):
    return barZ-delta(t,T)*f(Z,K)


def MonteCarlo(M,N,T,S0,A,B):
    dt = 1/N
    S = np.zeros((M,int(T*N)+1))
    S[:,0] = S0
    eps = np.random.normal(0, 1, (M,int(N*T)))
    S[:,1:] = np.exp((A-0.5*B**2)*dt + eps*B*np.sqrt(dt));
    S = np.cumprod(S, axis = 1);
    return S


def LSM(M,N,T,Phi,Z,R):
    C = np.zeros((M,int(T*N)+1))
    for m in range(M):
        C[m,T*N] = np.max([Phi[m,T*N],0]) #K-S[m,T*N],0])
    X = np.zeros((M,int(T*N)))
    Y = np.zeros((M,int(T*N)))
    Exercise = np.zeros((M,int(T*N)))
    Continue = np.zeros((M,int(T*N)))
    for n in range(int(T*N),1,-1):
        x = np.zeros(0)
        y = np.zeros(0)
        for i in range(M):
            if Phi[i,n-1] > 0:
                X[i,n-1] = Z[i,n-1]
                # x variable of regression should be the underlying
```

```
                # Monte Carlo simulation
                Exercise[i,n-1] = Phi[i,n-1]
                # exercise value if exercise now
                x = np.append(x,X[i,n-1])
                Y[i,n-1] = C[i,n] #df * C[i,n]
                y = np.append(y,Y[i,n-1])
        if len(x) == 0:
            p = np.array([0,0,0])
        else:
            p = spy.polyfit(x,y,2)
        for i in range(M):
            if S[i,n-1] > 0: #K-S[i,n-1] > 0:
                Continue[i,n-1] =
                 p[0]*X[i,n-1]**2 + p[1]*X[i,n-1] + p[2]
                 # Expected value of continuing,
                 # calculated with degreee 2 regression.
        for i in range(M):
            if Continue[i,n-1] < 0:
            # If exercise now value exceeds expected
            # value of continuing.
                C[i,n-1] = Exercise[i,n-1]
                C[i,n:] = 0
    return C, Continue, Exercise

def Sigma(M,N,T,C):
    sigma = np.zeros(M)
    for m in range(M):
```

```
        for n in range(int(T*N)):
            if C[m,n+1] != 0:
                sigma[m] = n+1
    return sigma

# Parameters for running the model.
M = 500
T = 1
N = 24
dt = 1/N
alpha = 0.09
beta = 0.1
r = 0.1
Z0 = 100
K = 50

# Build the Monte Carlo simulation for X.
Z = MonteCarlo(M,N,T,Z0,alpha,beta)

# Calculation of integral of e^{-rs}(X_s - K) along each path
# Integration with trapezoidal rule (for non-uniform widths)
S = np.zeros((M,T*N+1))
for m in range(M):
    for n in range(T*N):
        S[m,n+1] = S[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*f(Z[m,n+1],K)
         +np.exp(-r*dt*n)*f(Z[m,n],K));
```

```
barZ0 = delta(0,T)*f(Z0,K)
barZ = np.zeros((M,T*N+1))
barZ[:,0] = barZ0
for n in range(T*N):
    barZ[:,n+1] = delta((n+1)*dt,T)*f(Z[:,n+1],K)
     +np.exp(r*(n+1)*dt)*S[:,n+1]

barZ2 = np.zeros((M,T*N+1))
barZ2[:,0] = barZ0
for n in range(T*N):
    barZ2[:,n+1] = barZ2[:,n] + delta((n)*dt,T)
     *(f(Z[:,n+1],K)-f(Z[:,n],K))
     +(barZ2[:,n]-delta((n)*dt,T)*f(Z[:,n],K))*r*dt

# Value function with simulated integral.
Phi0 = S

# Value function with portfolio 1
Phi1 = np.zeros((M,T*N+1))
Phi1[:,0] = barZ0-T*f(Z0,K)
for m in range(M):
    for t in range(T*N):
        Phi1[m,t+1] = np.exp(-r*(t+1)*dt)
         *phi(Z[m,t+1],barZ[m,t+1],(t+1)*dt,T)

# Value function with portfolio 2
Phi2 = np.zeros((M,T*N+1))
```

```
Phi2[:,0] = barZ0-T*f(Z0,K)
for m in range(M):
    for t in range(T*N):
        Phi2[m,t+1] = np.exp(-r*(t+1)*dt)
         *phi(Z[m,t+1],barZ2[m,t+1],(t+1)*dt,T)

# Run LSM
C0, con0,exe0 = LSM(M,N,T,Phi0,Z,r)
C1 , con1, exe1 = LSM(M,N,T,Phi1,Z,r)
C2, con2, exe2 = LSM(M,N,T,Phi2,Z,r)

sigma = np.zeros(M)
#sigma = np.flatnonzero(C[0,:])[0]

for i in range(M):
    if C0[i,:].any() != 0:
        sigma[i] = np.flatnonzero(C0[i,:])[0]
    if C1[i,:].any() != 0:
        sigma[i] = np.flatnonzero(C1[i,:])[0]
    if C2[i,:].any() != 0:
        sigma[i] = np.flatnonzero(C2[i,:])[0]

cvec0 = np.zeros(M)
cvec1 = np.zeros(M)
cvec2 = np.zeros(M)
for i in range(M):
    for j in range(N*T+1):
```

```
        if C0[i,j] != 0:
            cvec0[i]= C0[i,j]
        if C1[i,j] != 0:
            cvec1[i] = C1[i,j]
        if C2[i,j] != 0:
            cvec2[i] = C2[i,j]

if K == 0:
    sim0 = np.mean(C0[:,N*T])
    sim1 = np.mean(C1[:,N*T])
    sim2 = np.mean(C2[:,N*T])
    predicted = Z0/(alpha-r)*(np.exp((alpha-r)*T)-1)
    print('As K=0 was chosen, the following
     are the simulated (3 approaches) and predicted values:')
    print('Simulated0 by averaging final values:', sim0)
    print('Simulated1 by averaging final values:', sim1)
    print('Simulated2 by averaging final values:', sim2)
    print('Predicted by direct calculation:', predicted)

if K > 0:
    g1 = beta**(-2)*(0.5*beta**2 - alpha
     + np.sqrt((0.5*beta**2-alpha)**2 +2*r*beta**2))
    g2 = beta**(-2)*(0.5*beta**2 - alpha
     - np.sqrt((0.5*beta**2-alpha)**2 +2*r*beta**2))
    d = K*g2*(r-alpha)/(r*(g2-1))
    C = -d**(-g2+1)/(g2*(r-alpha))
    sim0 , sim1, sim2 = np.mean(cvec0), np.mean(cvec1), np.mean(cvec2)
```

```
print('As K>0 was chosen,
 the following are the simulated values with all 3 approaches:')
print('Simulated0 by averaging final values:', sim0)
print('Simulated1 by averaging final values:', sim1)
print('Simulated2 by averaging final values:', sim2)
print('Infinite horizon problem has parameters C, d, gamma2:'
 , d, g2)
Phiinf = -d/(g2*(r-alpha))*(Z0/d)**(g2) + Z0/(r-alpha) - K/r
print('Infinite horizon problem has solution
 for these parameters:', Phiinf)
```

APPENDIX B: TWO CANDIDATE CODE

Here we provide the Python code used to generate the 3-dimensional plot of $\psi(x_1, x_2)$ seen in Figure 3.1.

```
# 3-D Plot of Two Candidate Switching Solution



import numpy as np
import scipy as spy
import pylab as pl
import mpl_toolkits.mplot3d.axes3d as p3



# Dynamics of first process, X_1,t
alpha1 = 0.05        # drift
beta1 = 0.3          # volatility

# Dynamics of second process, X_2,t
alpha2 = 0.09        # drift
beta2 = 0.1         # volatility

# Discount (interest) rate
r = 0.1

# Correlation
rho = 0.5       # Must be between -1 and 1
```

```
b = (0.5)*(beta1**2 -2*beta1*beta2*rho + beta2**2)
a = alpha1-alpha2

# Calculation of exponent lambda
l = (b+a+np.sqrt((b+a)**2+4*b*(r-alpha1)))/(2*b)

# Calculation of cuttoffs for both methods
mu = l*(r-alpha2)/((l-1)*(r-alpha1))

# Calculation of constant for both methods
C = mu**(-l)/((l-1)*(r-alpha1))

x1 = np.linspace(0,10,101)

x2 = np.linspace(0,10,101)

psi = np.zeros((101,101))

for i in range(len(x2)):
    for j in range(len(x1)):
        if x2[i] < mu*x1[j]:
            psi[i,j] = C*(x1[j]**(1-l))*x2[i]**l + x1[j]/(r-alpha1)
        else:
            psi[i,j] = x2[i]/(r-alpha2)

# Generate grid for 3-D plot.
X, Y = pl.meshgrid(x1,x2)
```

```
# Generate 3-D plot.
fig = pl.figure()
ax = p3.Axes3D(fig)
ax.plot_surface(X,Y,psi)
ax.set_xlabel('X1')
ax.set_ylabel('X2')
ax.set_zlabel('psi')
fig.add_axes(ax)
pl.show()
```

Below we include both sets of code used in generating the portfolio value for the finite horizon switching case. First is the code using multilinear regression.

```
# Two variable switching simulation, f(x)=x.


import numpy as np
import scipy as spy
import pylab as pl
import pandas


# For 3d plots. This import is necessary to have 3D plotting below
from mpl_toolkits.mplot3d import Axes3D


# For statistics. Requires statsmodels 5.0 or more
from statsmodels.formula.api import ols


def MonteCarlo(M,N,T,S0,A,B):
```

```
    dt = 1/N
    S = np.zeros((M,int(T*N)+1))
    S[:,0] = S0
    eps = np.random.normal(0, 1, (M,int(N*T)))
    S[:,1:] = np.exp((A-0.5*B**2)*dt + eps*B*np.sqrt(dt));
    S = np.cumprod(S, axis = 1);
    return S

def phi(xbar,x1,x2,t,T):
    return xbar - (T-t)*(x1-x2)

# Initial parameters

M = 500     # No. of paths per simulation
N = 12      # No. of time periods per year
T = 1       # No. of years

dt = 1/N

rho = 0.5
r = 0.1

a1 = 0.05
b1 = 0.2
a2 = 0.09
b2 = 0.2
```

```
X10 = 1

X20 = 1


X1, X2 = MonteCarlo(M,N,T,X10,a1,b1), MonteCarlo(M,N,T,X20,a2,b2)

#X2 = np.zeros((M,T*N+1))


S1,S2 = np.zeros((M,T*N+1)),np.zeros((M,T*N+1))

for m in range(M):

    for n in range(T*N):

        S1[m,n+1] = S1[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*X1[m,n+1]

         +np.exp(-r*dt*n)*X1[m,n]);

        S2[m,n+1] = S2[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*X2[m,n+1]

         +np.exp(-r*dt*n)*X2[m,n]);


barX0 = T*(X10-X20)

barX = np.zeros((M,T*N+1))

barX[:,0] = barX0

for n in range(T*N):

    barX[:,n+1] = (T-(n+1)*dt)*(X1[:,n+1]-X2[:,n+1])

     + np.exp(r*(n+1)*dt)*(S1[:,n+1]-S2[:,n+1])


Phi = np.zeros((M,T*N+1))

Phi[:,0] = barX0-T*(X10-X20) #+ X20/(a2-r)*(np.exp((a2-r)*T)-1)

for m in range(M):

    for t in range(T*N):

        Phi[m,t+1] = np.exp(-r*(t+1)*dt)

         *phi(barX[m,t+1],X1[m,t+1],X2[m,t+1],(t+1)*dt,T)
```

```
barX2 = np.zeros((M,T*N+1))
barX2[:,0] = barX0
dbarX = np.zeros((M,T*N))
for n in range(T*N):
    dbarX[:,n] = (T-n*dt)*(X1[:,n+1]-X1[:,n])
     + (n*dt-T)*(X2[:,n+1]-X2[:,n])
     + (barX2[:,n]-(T-n*dt)*X1[:,n]-(n*dt-T)*X2[:,n])*r*dt
    barX2[:,n+1] = barX2[:,n] + dbarX[:,n]

Phi2 = np.zeros((M,T*N+1))
Phi2[:,0] = barX0-T*(X10-X20) #+ X20/(a2-r)*(np.exp((a2-r)*T)-1)
for m in range(M):
    for t in range(T*N):
        Phi2[m,t+1] = np.exp(-r*(t+1)*dt)
         *phi(barX2[m,t+1],X1[m,t+1],X2[m,t+1],(t+1)*dt,T)

# LSM
C1, C2 = np.zeros((M,int(T*N)+1)), np.zeros((M,int(T*N)+1))
for m in range(M):
    C1[m,T*N], C2[m,T*N] = np.max([Phi[m,T*N],0]),
      np.max([Phi2[m,T*N],0])
regX1 = np.zeros((M,int(T*N)))
regX2 = np.zeros((M,int(T*N)))
regY1, regY2 = np.zeros((M,int(T*N))), np.zeros((M,int(T*N)))
Exercise1, Exercise2 = np.zeros((M,int(T*N))), np.zeros((M,T*N))
Continue1, Continue2 = np.zeros((M,int(T*N))), np.zeros((M,int(T*N)))
```

```
for n in range(int(T*N),1,-1):
    x1 = np.zeros(0)
    x2 = np.zeros(0)
    y1, y2 = np.zeros(0), np.zeros(0)
    for i in range(M):
        if Phi[i,n-1] > 0:
            regX1[i,n-1] = X1[i,n-1] # x1 Monte Carlo on X1
            regX2[i,n-1] = X2[i,n-1] # x2 Monte Carlo on X2
            Exercise1[i,n-1] = Phi[i,n-1]
            Exercise2[i,n-1] = Phi2[i,n-1]
             # exercise value if exercise now
            x1 = np.append(x1,regX1[i,n-1])
            x2 = np.append(x2,regX2[i,n-1])
            regY1[i,n-1] = C1[i,n] #df * C[i,n]
            y1 = np.append(y1,regY1[i,n-1])
            regY2[i,n-1] = C2[i,n]
            y2 = np.append(y2,regY2[i,n-1])
        X = regX1.flatten()
        Y = regX2.flatten()
        Z1 = regY1.flatten()
        Z2 = regY2.flatten()
        data1 = pandas.DataFrame({'x1': X, 'x2': Y, 'Phi': Z1})
        data2 = pandas.DataFrame({'x1': X, 'x2': Y, 'Phi2': Z2})

        # Fit the model
        model1 = ols("Phi ~ x1 + x2", data1).fit()
        model2 = ols("Phi2 ~ x1 + x2", data2).fit()
```

```
        for i in range(M):
            if Phi[i,n-1] > 0:
                Continue1[i,n-1] = model1._results.params[0]
                 + model1._results.params[1]*X1[i,n-1]
                 + model1._results.params[2]*X2[i,n-1]
                Continue2[i,n-1] = model2._results.params[0]
                 + model2._results.params[1]*X1[i,n-1]
                 + model2._results.params[2]*X2[i,n-1]
                # Expected value of continuing,
                # calculated with multilinear regression.
        for i in range(M):
            if Continue1[i,n-1] < 0:
             # If exercise now value exceeds
             # expected value of continuing.
                C1[i,n-1] = Exercise1[i,n-1]
                C1[i,n:] = 0
            if Continue2[i,n-1] < 0:
                C2[i,n-1] = Exercise2[i,n-1]
                C2[i,n:] = 0

sigma1, sigma2 = np.zeros(M), np.zeros(M)
for i in range(M):
    if C1[i,:].any() != 0:
        sigma1[i] = np.flatnonzero(C1[i,:])[0]
    if C2[i,:].any() != 0:
        sigma2[i] = np.flatnonzero(C2[i,:])[0]
```

```
x1sigma1, x2sigma1 = np.zeros(M), np.zeros(M)
x1sigma2, x2sigma2 = np.zeros(M), np.zeros(M)
for i in range(M):
    x1sigma1[i], x2sigma1[i] = X1[i,sigma1[i]], X2[i,sigma1[i]]
    x1sigma2[i], x2sigma2[i] = X1[i,sigma2[i]], X2[i,sigma2[i]]

cvec1 = np.zeros(M)
cvec2 = np.zeros(M)
for i in range(M):
    for j in range(N*T+1):
        if C1[i,j] != 0:
            cvec1[i] = C1[i,j]
        if C2[i,j] != 0:
            cvec2[i] = C2[i,j]

sim1, sim2 = np.mean(cvec1)+X20/(a2-r)*(np.exp((a2-r)*T)-1)
, np.mean(cvec2)+X20/(a2-r)*(np.exp((a2-r)*T)-1)

print(sim1, sim2)
print('Method 1 average switch time at t=',
np.mean(sigma1)/N)
print('Method 1 average value of x1 and x2 at stop:',
np.mean(x1sigma1), np.mean(x2sigma1))
print('Method 2 average switch time at t=',
np.mean(sigma2)/N)
print('Method 2 average value of x1 and x2 at stop:',
np.mean(x1sigma2), np.mean(x2sigma2))
```

Next we have the Python code for the finite horizon switching problem, but with degree two polynomial regression on $X_t = X_{1,t} - X_{2,t}$.

```
# Two Variable Switch with X=X1-X2


import numpy as np
import scipy as spy
import pylab as pl
import pandas


# For 3d plots. This import is necessary to have 3D plotting below
from mpl_toolkits.mplot3d import Axes3D


# For statistics. Requires statsmodels 5.0 or more
from statsmodels.formula.api import ols


def MonteCarlo(M,N,T,S0,A,B):
    dt = 1/N
    S = np.zeros((M,int(T*N)+1))
    S[:,0] = S0
    eps = np.random.normal(0, 1, (M,int(N*T)))
    S[:,1:] = np.exp((A-0.5*B**2)*dt + eps*B*np.sqrt(dt));
    S = np.cumprod(S, axis = 1);
    return S


def LSM(M,N,T,Phi,Z,R):
    C = np.zeros((M,int(T*N)+1))
    for m in range(M):
```

```
    C[m,T*N] = np.max([Phi[m,T*N],0]) #K-S[m,T*N],0])
X = np.zeros((M,int(T*N)))
Y = np.zeros((M,int(T*N)))
Exercise = np.zeros((M,int(T*N)))
Continue = np.zeros((M,int(T*N)))
for n in range(int(T*N),1,-1):
    x = np.zeros(0)
    y = np.zeros(0)
    for i in range(M):
        if Phi[i,n-1] > 0:
            X[i,n-1] = Z[i,n-1]
            Exercise[i,n-1] = Phi[i,n-1]
            x = np.append(x,X[i,n-1])
            Y[i,n-1] = C[i,n]
            y = np.append(y,Y[i,n-1])
    if len(x) == 0:
        p = np.array([0,0,0])
    else:
        p = spy.polyfit(x,y,2)
    for i in range(M):
        if S[i,n-1] > 0: #K-S[i,n-1] > 0:
            Continue[i,n-1] =
             p[0]*X[i,n-1]**2 + p[1]*X[i,n-1] + p[2]
             # Expected value of continuing,
             # calculated with degreee 2 regression.
    for i in range(M):
        if Continue[i,n-1] < 0:
```

```
                C[i,n-1] = Exercise[i,n-1]
                C[i,n:] = 0
    return C#, Continue, Exercise



def phi(xbar,x,t,T):
    return xbar - (T-t)*(x)

# Initial parameters

M = 500     # No. of paths per simulation
N = 12      # No. of time periods per year
T = 50      # No. of years

dt = 1/N

rho = 0.5
r = 0.1

a1 = 0.05
b1 = 0.2
a2 = 0.09
b2 = 0.2

X10 = 1
X20 = 1
```

```
X1, X2 = MonteCarlo(M,N,T,X10,a1,b1), MonteCarlo(M,N,T,X20,a2,b2)
#X2 = np.zeros((M,T*N+1))

X = X1-X2

#S1,S2 = np.zeros((M,T*N+1)),np.zeros((M,T*N+1))
S = np.zeros((M,T*N+1))
for m in range(M):
    for n in range(T*N):
        S[m,n+1] = S[m,n]+0.5*dt*(np.exp(-r*dt*(n+1))*X[m,n+1]
         +np.exp(-r*dt*n)*X[m,n]);

barX0 = T*(X10-X20)
barX = np.zeros((M,T*N+1))
barX[:,0] = barX0
for n in range(T*N):
    barX[:,n+1] = (T-(n+1)*dt)*(X[:,n+1])
     + np.exp(r*(n+1)*dt)*(S[:,n+1])

Phi1 = np.zeros((M,T*N+1))
Phi1[:,0] = barX0-T*(X10-X20)
for m in range(M):
    for t in range(T*N):
        Phi1[m,t+1] = np.exp(-r*(t+1)*dt)
         *phi(barX[m,t+1],X[m,t+1],(t+1)*dt,T)

barX2 = np.zeros((M,T*N+1))
```

```
barX2[:,0] = barX0
dbarX = np.zeros((M,T*N))
for n in range(T*N):
    dbarX[:,n] = (T-n*dt)*(X[:,n+1]-X[:,n])
     + (barX2[:,n]-(T-n*dt)*X[:,n])*r*dt
    barX2[:,n+1] = barX2[:,n] + dbarX[:,n]

Phi2 = np.zeros((M,T*N+1))
Phi2[:,0] = barX0-T*(X10-X20)
for m in range(M):
    for t in range(T*N):
        Phi2[m,t+1] = np.exp(-r*(t+1)*dt)
         *phi(barX2[m,t+1],X[m,t+1],(t+1)*dt,T)



# LSM
C1, C2 = LSM(M,N,T,Phi1,X,r), LSM(M,N,T,Phi2,X,r)

sigma1, sigma2 = np.zeros(M), np.zeros(M)
for i in range(M):
    if C1[i,:].any() != 0:
        sigma1[i] = np.flatnonzero(C1[i,:])[0]
    if C2[i,:].any() != 0:
        sigma2[i] = np.flatnonzero(C2[i,:])[0]

x1sigma1, x2sigma1 = np.zeros(M), np.zeros(M)
x1sigma2, x2sigma2 = np.zeros(M), np.zeros(M)
```

```
for i in range(M):
    x1sigma1[i], x2sigma1[i] = X1[i,int(sigma1[i])],
      X2[i,int(sigma1[i])]
    x1sigma2[i], x2sigma2[i] = X1[i,int(sigma2[i])],
      X2[i,int(sigma2[i])]

cvec1 = np.zeros(M)
cvec2 = np.zeros(M)
for i in range(M):
    for j in range(N*T+1):
        if C1[i,j] != 0:
            cvec1[i] = C1[i,j]
        if C2[i,j] != 0:
            cvec2[i] = C2[i,j]

sim1, sim2 = np.mean(cvec1) + X20/(a2-r)*(np.exp((a2-r)*T)-1),
np.mean(cvec2) + X20/(a2-r)*(np.exp((a2-r)*T)-1)

b = 0.5*b1**2 - b1*b2*rho + 0.5*b2**2
a = a1-a2
l = (b+a + np.sqrt((b+a)**2 +4*b*(r-a1)))/(2*b)
mu = l/(l-1)*(r-a2)/(r-a1)
stop1 = mu*np.mean(x1sigma1)
stop2 = mu*np.mean(x1sigma2)

print(sim1, sim2)
print('Method 1 average switch time at t=', np.mean(sigma1)/N)
```

```
print('Method 1 average value of x1 and x2 at stop:',
 np.mean(x1sigma1), np.mean(x2sigma1))
print('Method 2 average switch time at t=', np.mean(sigma2)/N)
print('Method 2 average value of x1 and x2 at stop:',
 np.mean(x1sigma2), np.mean(x2sigma2))
print('Predicted stop (infinite horizon) at x2>mu*x1 and avg.
 value of mu*x1 for each method is:', stop1, stop2)
```