We analyze a continuous-time model of dynamic project choices by an agent that is relevant to various important issues in financial economics. An agent controls the drift and volatility of a lognormal output process by dynamically selecting one of $N$ projects. Under broad conditions, the optimal risk-taking policy is characterized by at most $K - 1$ unique switching triggers, where $K$ is the number of spanning projects. As an application, we analyze a continuous-time structural model of capital structure with asset substitution, and obtain novel implications for the effects of systematic risk on capital structure. In contrast with previous studies, there are substantial agency costs of asset substitution, which generate leverage ratios and credit spreads consistent with empirical values.

Key Words: continuous-time model, dynamic optimal control, asset substitution, risk management

JEL Classification: G32, C61, D21, D81
1. Introduction

A number of important problems in financial economics entail the analysis of the dynamic project/risk choices of an agent. For example, a key manifestation of shareholder-debtholder agency conflicts is asset substitution by shareholders that transfers wealth from debtholders to shareholders at the expense of overall firm value. The study of corporate risk management necessitates the analysis of how a firm’s shareholders, or its manager if she controls the firm, respond to their payoff structures by choosing the firm’s risk. An important objective of the regulation of financial institutions is to control the risk-shifting incentives of the managers/shareholders of the institutions that could be detrimental to not only their overall enterprise values, but also their social values more generally because of the possibility of public capital injection in financial distress.

Despite the importance of such risk-taking problems, little is known about the nature of the optimal policies when an agent faces a discrete set of project/risk choices in a dynamic setting. We contribute to the literature by rigorously analyzing a continuous-time model in which an agent controls the evolution of a lognormal output process by dynamically selecting one of a finite, but arbitrary, number of projects, and also chooses the termination time. We show that the agent’s optimal policy depends on the projects’ risk-adjusted drifts that are determined by their drifts and volatilities as well as the curvature/convexity of the agent’s payoff function. The optimal policy only selects projects in the spanning subset, which comprises of the projects whose drifts and variances are extreme points of the upper convex contour of the set defined by the drifts and variances of the original set of projects. Further, if $K$ is the number of projects in the spanning subset, and if their risk-adjusted drifts are consistently ordered—that is, their ordinal ranking remains unaltered for all output values—then the optimal policy is characterized by at most $K - 1$ unique switching triggers.

In the case where the risk-adjusted drifts of the spanning projects need not be consistently ordered, we demonstrate that there could be more than $K - 1$ switching points in general and provide an upper bound on the number of switching points.

As a concrete application of our results, we analyze a continuous-time structural model of capital structure that incorporates agency conflicts arising from asset substitution. We calibrate the model and obtain novel implications for the effects of systematic risk on capital structure, credit
spreads and agency costs. In contrast with the conclusions of previous studies (e.g., Leland (1998)), there are substantial agency costs of asset substitution, which generate leverage ratios and credit spreads consistent with empirical values.

In our continuous-time model, an agent controls the drift and volatility of a lognormal output process by dynamically selecting one of $N$ projects. The agent’s payoff increases with the output. She dynamically chooses the projects and the termination time to maximize her discounted expected payoff stream. We demonstrate that the projects’ risk-adjusted drifts— the drift minus the product of (a half of) the volatility and the local curvature/convexity (which could be positive or negative and vary with the output level) of the agent’s payoff function—play a central role in determining the agent’s dynamic project choices.

We first analyze the case where there are only two projects to illustrate our techniques and the intuition underlying our general results. Consider a low risk project 1 and a high risk project 2. If the projects’ risk-adjusted drifts are consistently ordered, then the agent’s optimal dynamic project choices are characterized by a unique trigger value of the output where the agent switches between the projects. More precisely, if the risk-adjusted drift of project 1 is higher than that of project 2, it is optimal for the agent to select project 1 when the output level is higher than a unique threshold, and project 2 if the output level is below the threshold. If project 2 has a higher risk-adjusted drift than project 1, however, the optimal policy always selects project 2.

The intuition for the structure of the optimal policy and the role of the “consistent ordering” condition is as follows. If one project has a higher risk-adjusted drift for all output realizations, then it has a higher present value (NPV) from the agent’s standpoint than the other project. When the output is high, the probability of termination is low so that it is optimal for the agent to choose the project with the higher risk-adjusted drift because it has a higher NPV. As the output declines, the increase in the termination probability may induce the agent to switch to the other project provided it has higher volatility. In particular, if one project has a higher risk-adjusted drift and volatility than the other, then the agent always selects this project in the optimal policy because it dominates the other project when the output is high owing to its higher risk-adjusted drift as well as when the output is low owing to its higher volatility. If the “consistent ordering” condition is violated, however, the above intuition does not apply because the “relative” dominance of one
project over the other changes as the output varies. Consequently, there could be multiple switching points. (We provide such an example in Section 3.)

Notwithstanding the intuition described above, it is far from obvious that the optimal policy features a single switching trigger even when the consistent ordering condition holds. The rigorous proof of our results hinges on the fact that the Hamilton-Jacobi-Bellman (HJB) equation associated with the agent’s dynamic control problem is an elliptic equation. In our proof, we utilize a key property of elliptic equations—the maximum principle—which ensures that solutions within a region with non-positive boundary conditions remain non-positive inside the region. Given the maximum principle, the consistent ordering of the projects’ risk-adjusted drifts is sufficient to ensure that the agent’s optimal policy is “monotonic” and, therefore, characterized by a single switching trigger.

We then analyze the more general setting where there are an arbitrary number of projects, and provide a sharp characterization of the optimal policy. In particular, we prove that the optimal policy only entails the choices of projects in the spanning subset, that is, projects whose drifts and variances are extreme points of the upper convex contour of the set defined by the drifts and variances of the projects. For a project outside the spanning subset, its risk-adjusted drift is less than or equal to a linear combination of the risk-adjusted drifts of the spanning projects weighted by their variances. Building on the intuition for our results in the two-project case, this condition implies that the project is “dominated” by the spanning projects for high and low output realizations and thus not optimally chosen at all by the agent.

Using extensions of the techniques employed in the two-project case, we then show that if the risk-adjusted drifts of the spanning projects are consistently ordered, then the optimal policy is characterized by at most $K - 1$ switching triggers where $K$ is the number of projects in the spanning subset. Furthermore, we show that the optimal policy selects the spanning project with the highest risk-adjusted drift when output is above a threshold and progressively switches to a spanning project with higher volatility and lower risk-adjusted drift as the output declines (if at some point no such spanning project exists, then the agent no longer switches to another project). This result rigorously formalizes the common intuition about asset substitution (obtained in special, two-period models) that agents optimally switch to higher-risk projects with lower NPVs as output declines. We, however, clearly delineate exactly which projects the agent optimally chooses (among
the set of available projects) as well as the condition (the consistent ordering condition) under which the intuition is valid. Our analysis also leads to analytical characterizations of the optimal switching triggers and the termination threshold via a set of “super-contact” and “smooth-pasting” conditions. We exploit these conditions and our analytical results to outline a general algorithm to derive the optimal risk-taking policy in various applications of our model.

As we discuss in the paper, the consistent ordering condition on the spanning projects holds for a number of applications of interest. Nevertheless, we are able to characterize the optimal policy even when the condition does not hold. In particular, we prove that the number of switching points in the optimal policy is at most $M(K - 1)$ where $M$ is the number of sub-intervals of output values over which the projects are consistently ordered, that is, the ordinal ranking of their risk-adjusted drifts does not change. By the intuition for our results when the spanning projects are consistently ordered, there can be at most $K - 1$ switches within each sub-interval over which the projects are consistently ordered so that the total number of switches is bounded above by $M(K - 1)$. In the process, we also demonstrate that the conventional intuition about asset substitution is incorrect in the absence of the consistent ordering condition.

As a concrete application of our general results, we adapt our model to examine how shareholder-debtholder agency conflicts arising from asset substitution affect the capital structure choices and credit spreads of firms. Specifically, we analyze a structural model of capital structure in which a firm can dynamically switch between two projects. The firm’s earnings evolve with different drifts and volatilities under the projects with the higher volatility project having a higher actual or physical drift (that is, the drift under the physical probability measure). The firm’s capital structure reflects the tradeoff between the tax advantages of debt and financial distress costs that include the agency costs of asset substitution.

Our framework is particularly well suited to investigate how the firm’s optimal capital structure, risk policy, and agency costs of asset substitution change when the systematic portion of the firm’s risk varies. Security values and, therefore, capital structure choices depend on the risk-neutral drifts of the projects. The wedge between the physical and risk-neutral drifts is the risk premium that varies with the firm’s systematic risk. As the firm’s systematic risk increases above a threshold, \textit{ceteris paribus}, the higher-risk project has a lower risk-neutral drift, leading to dramatic changes
in the nature of the optimal risk-taking policy. As the systematic risk increases, the optimal policy changes from one in which only the high risk project is always chosen to one in which the low risk project is chosen in “good” states and the high risk project in “bad” states. Consequently, systematic risk has significant effects on the optimal risk-taking policy and, therefore, the firm’s optimal capital structure.

We calibrate the baseline parameters of the model to data on large U.S. public firms over the period 1962–2009 and quantitatively analyze the effects of systematic risk on the agency costs of asset substitution, capital structure and credit spreads. We find that optimal leverage ratios vary non-monotonically with systematic risk, due to the tradeoffs between the agency costs and tax benefits of debt. The optimal leverage ratio in the baseline model is 33% and varies between 25% and 58%. The optimal leverage ratio achieves its minimum (25%) when agency costs are highest (equal to 6% of firm value). The leverage ratios and credit spreads predicted by our analysis are consistent with those observed in the data. Our findings contrast sharply with those of the comparable study of Leland (1998). Leland’s quantitative analysis generates agency costs that are only about 1% of firm value, much higher leverage ratios, and lower credit spreads than those predicted by our analysis. The stark differences stem from the fact that we model the earnings process that could, in general, have different physical and risk-neutral drifts under the two projects. Systematic risk has significant effects on the agency costs, leverage ratios and credit spreads by affecting the risk premium and thus the risk-neutral drifts of the projects.

To the best of our knowledge, our paper is the first to rigorously analyze the general dynamic model of corporate project/risk choices that arises in the applications that we discussed earlier, and to provide a characterization of the optimal policies. The simple structure of the optimal policies can facilitate quantitative investigations of various issues such as the importance of agency conflicts due to asset substitution, the effects of managerial risk-taking on firm value, etc. Leland (1998) studies the dynamic asset substitution problem of shareholders using a model that is a special case of ours. He assumes that there are only two projects that have the same drifts, but different volatilities. Further, he assumes the form of the optimal policy without rigorously deriving it.

The problem that we examine is related to but very different from the multi-armed bandit
Because the two problems are very different, the well-known “Gittins index” policy (Gittins (1989)) that solves the bandit problem has no analogue in our setting and, therefore, cannot be used to derive the optimal control policy. Our paper is also related to the literature that studies differentiability of value functions in switching problems and the associated super-contact and smooth-pasting conditions satisfied at the switching triggers and termination thresholds. Twice-differentiability of solutions to elliptic equations (such as HJB equations) have been extensively studied in the mathematics literature (see e.g., Evans (1983) and Gilbarg and Trudinger (2001)). Dixit (1991) and Dumas (1991) introduce the very useful super-contact and smooth-pasting conditions to economists.

There is a stream of literature that studies models with risk-taking behavior by investors stemming from the seminal work of Merton (1971) on optimal investment strategies of a rational investor. Karatzas, Lechozky, Sethi and Shreve (1986) and Pliska (1986) obtain solutions to Merton’s portfolio problem for general utility functions (see Karatzas and Shreve (1998) for an exposition). Cadenillas, Cvitanic, and Zapatero (2007) and Ou-Yang (2007) investigate models of delegated portfolio management where an agent makes dynamic investment decisions that affect the wealth process. In these models, the agent invests in any linear combination of a number of securities, while our model corresponds more closely to corporate investment decisions where there are typically only a discrete set of projects available and it is usually impossible to “short” a project.

The results of the application that we study shed light on how aggregate or systematic uncertainty influences firms’ capital structure and risk-taking decisions. These issues are difficult to analyze in traditional contingent claims models of capital structure that carry out their entire analyses in the risk-neutral measure and, therefore, do not examine how the risk premium and systematic risk affect capital structure (e.g., Leland (1994, 1998)). In a related literature on capital structure and economic cycles, which does not consider asset substitution, Almeida and Philippon

---

1In the multi-armed bandit problem, the choice of an arm changes the state of that arm, but leaves the states of all the other arms frozen. The state of the system at any date is, therefore, represented by a vector that specifies the state of each arm. In our setting, however, the state variable is the output process that is controlled by the agent. The choice of a project alters the evolution of the output process. Further, the agent also controls the termination time of the output process. In a recent paper, Strulovici and Szydlowski (2012) show the existence of “classical” solutions (see Fleming and Soner (2005)) of the value function for the optimal control of any time-homogeneous, one-dimensional diffusion. They apply their results to a two-armed bandit problem and characterize the optimal policy.
(2007), Chen (2010), and Bhamra, Kuehn, and Strebulav (2010) argue that the high marginal utility of money in an economic recession leads to higher expected costs of default under the risk-neutral probability, and thus lower firm leverage ratios. Our analysis shows that asset substitution alone can lead to substantial agency costs and low leverage ratios when the effects of systematic risk are incorporated in a capital structure model.

2. Motivation and Description of Class of Problems

We describe and motivate the broad class of stochastic control problems that we analyze in this article. We consider a continuous-time setting with an infinite time horizon $[0, \infty)$. All stochastic processes are defined on an underlying probability space $[\Omega, \mathcal{F}, P]$. A firm generates cash flows $C_t$ that evolve as follows:

$$dC_t = \mu(p_t)C_t dt + \sigma(p_t)C_t dB_t,$$

where $B_t$ is a standard Brownian motion. Let $\{\mathcal{F}_t\}$ denote the complete and augmented filtration generated by the Brownian motion $B$. In (1), $p = \{p_t\}$ is an $\{\mathcal{F}_t\}$-adapted process that takes values in the finite set $\{1, 2, \ldots, N\}$. We interpret $p_t$ as a project choice at date $t$ that affects the distribution of the cash flow process. The drift and volatility of the cash flow process are given by

$$\mu(p_t) = \mu_i, \sigma(p_t) = \sigma_i, \text{if } p_t = i \in \{1, 2, \ldots, N\},$$

where $(\mu_i, \sigma_i)$ are constants for all $1 \leq i \leq N$.

Consider the agent(s) who controls the firm. For example, depending on the specific applications of the framework, the agent could be the firm’s manager or its shareholders. The objective of the agent is to dynamically choose the firm’s projects, $\{p_t\}$, and the time $\tau_B$ at which the firm’s operations are terminated (which is an $\mathcal{F}_t$-stopping time) to solve

$$\sup_{(\{p_t\}, \tau_B)} \left[ \int_0^{\tau_B} e^{-rs} f(C_s) ds \right].$$

(2)
where $f$ is the agent’s payoff function and $r$ is the time discount rate. We assume that

$$
\mu_i < r, \quad 1 \leq i \leq N
$$

so that the integral in (2) is finite.

For example, suppose the firm is controlled by its risk-neutral shareholders. The firm has a capital structure that consists of equity and long-term debt that requires it to make a total debt payment $\theta$ per unit time. In the absence of taxes, the payout flow to shareholders per unit time is

$$
f(C_t) = C_t - \theta.
$$

In the above scenario, the stopping time $\tau_B$ is the time at which the shareholders optimally declare bankruptcy.

In analyses of asset substitution, shareholders could choose risky, negative NPV projects that lead to a reduction in overall firm value. More specifically, consider the scenario where shareholders can dynamically choose between two projects 1 and 2 where

$$
\sigma_1 < \sigma_2, \quad \mu_1 > \mu_2.
$$

The less risky project 1, therefore, has a higher NPV than the more risky project 2. The project that maximizes overall firm value is project 1. The dynamic project choice strategy or policy that maximizes shareholder value, however, could involve switching to the riskier project 2, especially when the firm is in financial distress. Because shareholders effectively hold a call option on the firm’s cash flows, it could be optimal for them to increase risk when their option is sufficiently “out of the money” even though the riskier project has a lower drift. Although this intuition is familiar, the optimality of such a switching policy with a unique switching point is not known in the extant literature. We rigorously establish that this is indeed the case for two projects. It is also a priori unclear how this intuitive result generalizes to the more realistic setting where the number of available projects, $N > 2$. In this scenario, the optimal “asset substitution” policy, which depends on the drifts and volatilities of the projects, need not necessarily involve adopting the next riskier
project in financial distress. One of the implications of our analysis of the general model is the
delineation of conditions under which this is, indeed, the case and conditions under which it is not.

In analyses of risk management, shareholders of a leveraged firm may have lower incentives to
choose safe, positive NPV projects because a large portion of the project’s cash flows may accrue
to debtholders. In the context of our general framework, consider the case where the two projects satisfy

\[ \sigma_1 < \sigma_2, \quad \mu_1 < \mu_2. \]  

(4)

In this scenario, the riskier project has a higher NPV when the firm is in a normal state and, therefore, maximizes total firm value. In the presence of shareholder-debtholder agency conflicts, however, the incentives of shareholders could prevent them from using the less risky project 1 to hedge when the firm is in financial distress. Again, as in the case of the application to shareholder asset substitution, it is a priori unclear how this intuition extends to the scenario where there are more than two available projects, which we address in this paper.

In the above examples, the objective function of shareholders has the form (2). More generally, the firm could have multiple projects with differing drifts and volatilities. In addition, there can be nonlinear costs due to various sources such as taxes, financial distress costs and operating costs that affect shareholders’ payoffs so that the (net) payoff to shareholders \( f(C_t) \) could differ in general from the earnings, \( C_t \). The analysis of the deadweight costs of shareholder-debtholder agency conflicts, therefore, entails solving the general stochastic control problem (2), where there are multiple available projects.

Our framework is also applicable to the analyses of risk-taking by a manager whose objective function differs, in general, from that of shareholders. For example, consider an all-equity firm, where the manager holds a large equity stake so that she is significantly under-diversified. Her utility payoff \( f(C_t) \) is a concave function of the firm’s total cash flow \( C_t \), while the risk-neutral shareholders’ payoff is linear. The manager’s optimal dynamic project choices, therefore, do not maximize shareholders’ value. When the manager is not significantly risk-averse, and her compensation structure is significantly convex (e.g., due to stock option compensation), her payoff \( f(C_t) \) may be a convex function of the firm’s cash flow \( C_t \). In such a scenario, the manager could have
the incentive to engage in asset substitution by choosing a highly risky, lower NPV project. For example, with two available projects that satisfy (3), the convexity of the manager’s payoff structure could induce her to inefficiently choose the riskier project 2 when it is not in the interests of shareholders. In these examples too, the analysis of the inefficiencies due to manager-shareholder agency conflicts requires solving the general stochastic control problem (2).

3. Two Projects

To illustrate the techniques we use to analyze the general stochastic control problem (2), we begin by examining the simpler scenario in which there are two available projects, that is, $N = 2$. First, we set up the problem rigorously. Consider a firm that generates cash flow $C_t$ which is given by the following process

$$dC_t = \mu(p_t)C_t dt + \sigma(p_t)C_t dB_t,$$

where $B_t$ is the standard Brownian motion, and $p_t \in \{1, 2\}$ indicates the project choice of the firm at time $t$. We can write $C_t = C_t(P)$ to explicitly indicate the dependence of the cash flow process on the project choice policy $P = \{p_t\}_{t \geq 0}$. An admissible project choice policy $P$ is $\mathcal{F}_t$-adapted, where $\mathcal{F}_t$ is the complete and augmented filtration generated by the Brownian motion $B$. Let $\mathcal{P}$ denote the space of all admissible project choice processes.

Consider an agent who can dynamically select project 1 or 2 for the firm without additional costs. The agent’s objective function at any date $t$ is

$$S_t(C_t) = E_t \left[ \int_t^{\tau_B} e^{-r(s-t)} f(C_s) ds \right].$$

(6)

where the payoff function $f$ is continuous. The agent chooses the project choice policy and the stopping time to solve

$$\sup_{P \in \mathcal{P}: \tau_B} S_0(C_0).$$

(7)

In most relevant applications of the framework, an agent’s payoff increases with the firm’s cash flow so that we assume that $f$ is strictly increasing. To ensure that the integral in (6) is finite, we
also assume that \( f \) satisfies the following linear growth condition for some constant \( c \),

\[
|f(x)| \leq c(1 + x), \quad \text{for } x > 0. \tag{8}
\]

If \( f \) is concave (not necessarily strictly), the condition (8) is automatically satisfied. Even when \( f \) is convex, if the payoff function is asymptotically linear, then (8) still holds. For example, this is true for the payoff function of a manager endowed with options.\(^2\)

We begin by assuming that \( f \) is twice continuously differentiable, i.e., \( f \in C^2(\mathbb{R}_+) \), and later show that the results can be generalized to the case where \( f \) is not necessarily continuously differentiable. We define the local curvature or degree of convexity of \( f \) at \( x \) as

\[
\gamma_f(x) = -\frac{xf''(x)}{f'(x)}. \tag{9}
\]

Note that the local curvature/convexity could be positive or negative and vary with \( x \) in general. In the case where \( f \) is a utility function, \( \gamma_f(x) \) is the relative risk aversion of \( f \) at \( x \).

**Definition 1.** The risk-adjusted drift of project \( i \) at \( x \) is defined as

\[
\tilde{\mu}_i(x) = \mu_i - \frac{1}{2}\gamma_f(x)\sigma_i^2. \tag{10}
\]

Note that if \( \gamma_f(x) \equiv \gamma \) is a constant, then the risk-adjusted drift \( \tilde{\mu}_i = \mu_i - \frac{1}{2}\gamma\sigma_i^2 \) does not depend on \( x \). In the case where \( f \) is linear, \( \tilde{\mu}_i = \mu_i \). The risk-adjusted drifts play a central role in the characterization of the optimal project choices.

The following theorem shows that, if the projects’ risk-adjusted drifts are monotonically ranked, then the optimal project choice policy is characterized by at most a single trigger value, \( q^* \), of the cash flow process at which the agent switches between the projects. Without loss of generality,

\(^2\)The linear growth condition (8) can be relaxed to \( |f(x)| \leq c(1 + x^k) \), where \( k \) is a constant that satisfies \( 1 \leq k < \min(\gamma_1^+, \ldots, \gamma_N^+) \), \( \gamma_i^+ \) being the positive root of the characteristic equation \( \frac{1}{2}\sigma_i^2x(x-1) + \mu_i x - r = 0 \) (note that \( \gamma_i^+ > 1 \)). For simplicity, we state the linear growth condition throughout the paper.
assume that project 1 has lower volatility, i.e.,

\[ \sigma_1 < \sigma_2. \] (11)

**Theorem 1.** *(Optimal Project Choices with Two Projects)*

i) If \( \tilde{\mu}_1(x) > \tilde{\mu}_2(x) \) for all \( x > 0 \), then there exists a threshold level, \( C_B^* \geq 0 \), and a trigger \( \infty \geq q^* > C_B^* \) such that the agent’s optimal project choice policy \( P^* = \{p_t^*\} \) is as follows:

\[
p_t^* = \begin{cases} 
1, & \text{if } C_t \geq q^*, \\
2, & \text{if } q^* > C_t \geq C_B^*.
\end{cases}
\]

The agent optimally terminates the firm at the first time \( \tau_B^* \) that the firm’s cash flow hits the level \( C_B^* \).

ii) If \( \tilde{\mu}_1(x) \leq \tilde{\mu}_2(x) \) for all \( x > 0 \), then the agent always selects project 2. There exists a threshold level \( C_B^* \geq 0 \) such that the agent optimally terminates the firm at the first time \( \tau_B^* \) that the firm’s cash flow hits the level \( C_B^* \geq 0 \).

As suggested by Definition 1, the risk-adjusted drifts of the projects incorporate their volatilities and the curvature of the agent’s payoff function. In the hypothetical absence of the possibility of termination, the agent always prefers the project with the higher risk-adjusted drift because the present value of its payoffs to the agent are higher. Consequently, when the firm’s cash flows are sufficiently high so that the probability of termination is low, the agent optimally chooses the project with the higher risk-adjusted drift. In case i), when the firm’s cash flows are below an endogenous threshold, the higher probability of termination could cause the agent to switch to the riskier project even though it has lower risk-adjusted drift. In case ii), however, the high-risk project 2 also has a higher risk-adjusted drift so that it is never optimal for the agent to choose project 1. In other words, project 2 completely dominates project 1 because the agent prefers it when cash flows are high owing to its higher risk-adjusted drift as well as when cash flows are low.
owing to its higher volatility.

It is worth emphasizing that it is far from obvious that the optimal project choice policy is characterized by a single switching trigger even when the projects’ risk-adjusted drifts are “consistently ordered” (that is, project 1 has a higher risk-adjusted drift for all earnings realizations) as in case i) of the Theorem, especially when the agent’s payoff function $f$ is more general. Theorem 1 rigorously establishes that this is, indeed, the case. The proof of Theorem 1 makes use of the fact that the Hamilton-Jacobi-Bellman equations of the stochastic control problem (7) is an elliptic differential equation. Using the maximum principle for elliptic equations, which implies that solutions within a region with non-positive boundary conditions remain non-positive inside the region, we show that the switching trigger is unique. In our discussion below, we show that there could be multiple switching points when the risk-adjusted drifts are not consistently ordered.

In the proposition below, we show that, if $f$ is asymptotic to a CRRA function with risk aversion $0 \leq \gamma \leq 1$, then the switching point $q^*$ in Theorem 1 is finite.

**Proposition 1.** [Finiteness of Switching Trigger] Suppose the assumptions of Theorem 1 hold. Assume that $0 \leq \gamma \leq 1$ and

$$f(x) = c_1 + c_2 \frac{1}{1-\gamma} x^{1-\gamma} + o(x^{1-\gamma})$$

for $x$ sufficiently large, with $c_2 > 0$ (when $\gamma = 1$, assume $c_1 > 0$ and $c_2 = 0$), then the optimal switching trigger $q^*$ in Theorem 1 is finite.

Note that, even if $f$ is convex, the conditions of Proposition 1 can be satisfied if $f$ is asymptotic to a linear function (the case $\gamma = 0$). In many of the illustrative applications discussed in Section 2, the function $f$ is either concave or convex. The following corollary adapts the results in Theorem 1 to the case where $f$ is a general concave or convex function that is not necessarily continuously differentiable.

**Corollary 1.** [Optimal Project Choices for Concave or Convex Payoff Functions]

i) If $\mu_1 \geq \mu_2$ and $f$ is concave (not necessarily strict), then the optimal policy is characterized
by a unique switching trigger as in case i) of Theorem 1.

ii) If $\mu_1 \leq \mu_2$ and $f$ is convex (not necessarily strict), then the agent always selects the high-risk project 2 as in case ii) of Theorem 1.

The intuition for the result of part (i) hinges on the fact that a concave function has a non-negative curvature that can be unbounded if the function is not continuously differentiable. Consequently, if the low-volatility project 1 has a higher drift, then it also has a higher risk-adjusted drift. In the proof, we show that Theorem 1 case i) “applies” even though the payoff function may not be smooth. The intuition for part (ii) when the payoff function is convex is analogous.

In general, if the risk-adjusted drifts cannot be consistently ordered for all realizations of the firm’s earnings, there can be multiple switching points in the agent’s optimal policy. By the continuity of the risk-adjusted drifts as functions of the firm’s earnings, we can nevertheless decompose the positive real axis into regions such that the risk-adjusted drifts are monotonically ranked within each region. The following theorem shows that the number of switching points in the optimal policy is bounded by the number of such regions.

**Theorem 2.** Assume that the payoff function $f \in C^2$ and that there are $0 = x_0 < x_1 < x_2 < \ldots < x_{M-1} < x_M = \infty$ such that for any $1 \leq i \leq M$, $\tilde{\mu}_1(x) - \tilde{\mu}_2(x)$ does not change sign on $x \in (x_{i-1}, x_i)$. Then the total number of switching points in the optimal policy is bounded above by $M$.

The intuition for Theorem 2 builds on that of Theorem 1. Within an interval of values of the firm’s earnings, if the projects’ risk-adjusted drifts are consistently ordered—that is, their difference does not change sign—then it is optimal for the agent to choose the project with the higher risk-adjusted drift for higher earnings values within the interval and possibly switch to the riskier project for lower values of the earnings so that there is at most one switching point within the interval. The total number of switching points is, therefore, bounded above by the number of regions over which the risk-adjusted drifts are consistently ordered.
Panel A: Optimal policies

Panel B: Optimal value functions

Figure 1: Optimal policy and value functions when $\mu_2$ changes. In this case, the payoff function is linear, $f(x) = x - 1$. In Panel B, the value functions are computed at an initial cash flow level $C_0 = 1$.

In general, if the payoff function is not twice continuously differentiable, we can, in many cases of interest, find a limiting sequence of smooth payoff functions that converges to it. The optimal policy for the nonsmooth payoff function can then be characterized by the limits of the sequence of optimal policies corresponding to the approximating sequence of smooth payoff functions. As long as there is a uniform bound on the number of consistently ordered regions for the sequence of smooth payoff functions, the number of switching points in the limiting optimal policy is also subject to the same bound. Thus our results can also be applied to the general case with nonsmooth payoff functions given the existence of such a limiting sequence of smooth functions.

Illustrative Examples

Theorem 1, Corollary 1 and Theorem 2 together provide a characterization of the optimal switching policy in the case of two projects. We now illustrate the results with some numerical examples. We derive the optimal policies in the examples using the analytical characterization and computational algorithm for the general switching problem with an arbitrary number of projects that we outline in 5.

Example 1. Optimal policies with one switching point.
We fix \( r = 0.04 \) in the following examples. Consider a CRRA payoff function \( f(x) = x^{1-\gamma} - 1 \), where \( 0 \leq \gamma < 1 \). Consider two projects with \( \mu_1 = 0.02, \sigma_1 = 0.1, \sigma_2 = 0.2 \). In the following, we examine the optimal policies for different values of \( \mu_2 \). First, consider the case \( \gamma = 0 \), i.e., the payoff function is linear. If \( \mu_1 \geq \mu_2 \), then by Theorem 1 case (i), there exists a unique switching point \( q^* \in (0, \infty) \) and an optimal bankruptcy level \( C_B^* \) such that the optimal project choice takes the form

\[
p_t^* = \begin{cases} 
1, & \text{if } C_t \geq q^*, \\
2, & \text{if } q^* > C_t \geq C_B^*.
\end{cases}
\]

If \( \mu_1 < \mu_2 \), then by Theorem 1 case (ii), there exists an optimal bankruptcy level \( C_B^* \) such that the optimal project choice \( p_t^* \equiv 2 \) for \( C_t \geq C_B^* \).

Since \( f \) is a CRRA function, the value function can be written in closed form (as piecewise sums of power functions) and the switching and the bankruptcy points can also be expressed in closed form (see Section 5). The closed-form representation allows easy calculation of the optimal policies for different parameters in the model.

Panel A of Figure 1 plots the optimal switching points and bankruptcy thresholds for different values of \( \mu_2 \). We see that as \( \mu_2 \nearrow \mu_1 = 0.02 \), the optimal switching threshold \( q^* \) increases and approaches infinity and the optimal policy approaches the policy of always choosing project 2. In the figure, the line given by the optimal switching trigger \( q^* \) divides the area above bankruptcy into two regions, which correspond to the parameter values for which one of the two projects is optimal.

Panel B of Figure 1 compares the agent’s utility function under the optimal policy and under the policies of always choosing one project. When \( \mu_2 < \mu_1 = 0.02 \), the optimal policy improves the agent’s utility over the policy of always choosing either project 1 or 2 alone. For example, when \( \mu_2 = 0.019 \), the optimal policy yields a utility value of 26.1, a 4.1% (6.0%) improvement compared to 25.1 (24.6) for the policy of choosing only project 1 (2). As noted above, when \( \mu_2 \geq \mu_1 \), in this case, the optimal policy coincides with the policy of choosing project 2 alone.

Next we turn to the case \( \gamma > 0 \). In this case, by Theorem 1, the analysis is similar to above, except that the form of the optimal policies now change at the point \( \bar{\mu}_2 = \mu_2 - \frac{1}{2} \gamma \sigma_2^2 = \mu_1 - \frac{1}{2} \gamma \sigma_1^2 = 0.04 \).
Figure 2: **Optimal strategy and value functions when \( \mu_2 \) changes.** In this case, the payoff function is CRRA with \( \gamma = 0.5 \), \( f(x) = \sqrt{x} - 1 \). In Panel B, the value functions are computed at an initial cash flow level \( C_0 = 1 \).

\[ \tilde{\mu}_1 \]. For example, when \( \gamma = 0.5 \), this implies that the policies change form at \( \mu_2 = 0.0275 \). Figure 2 illustrates the optimal policies and value functions in the case \( \gamma = 0.5 \) and shows that the optimal policies are indeed similar to those in the case \( \gamma = 0 \), but with the critical point being \( \mu_2 = 0.0275 \).

**Example 2. Optimal policies with multiple switching points.**

We now present an example which shows that, if the conditions of Theorem 1 are violated, the optimal policy may be characterized by multiple switching points. Consider the payoff function given by the following:

\[
 f(x) = \begin{cases} 
 0.2x + 0.3, & \text{if } x \geq 1, \\
 x - 0.5, & \text{if } 0 \leq x < 1.
\end{cases}
\]

The function \( f \) is concave over \([0, \infty)\) but not continuously differentiable (\( f'(x) \) is not continuous at \( x = 1 \)). Therefore, Theorem 1 does not apply. Consider two projects with \((\mu_1, \sigma_1) = (-0.01, 0.1), (\mu_2, \sigma_2) = (0, 0.2)\). Since \( \sigma_1 < \sigma_2 \) and \( \mu_1 < \mu_2 \), Corollary 1 also does not apply in this case. Although we cannot directly apply Theorem 2 (because \( f \) is not \( C^2 \)), we can consider a sequence of smooth payoff functions that converge to \( f \) and show by a limiting argument that there can be at most two switching points in the optimal policy. In fact, the optimal policy in this case involves two switching triggers \((q_1^* = 1.345, q_2^* = 0.921)\) and the bankruptcy threshold \( C_B = 0.280 \):
\[ p_t^* = \begin{cases} 
2, & \text{if } C_t \geq 1.345, \\
1, & \text{if } 1.345 > C_t \geq 0.921, \\
2, & \text{if } 0.921 > C_t \geq 0.280.
\end{cases} \]

The intuition for the structure of the optimal policy is that because \( \mu_2 > \mu_1 \), it is optimal to choose project 2 when the cash flow is sufficiently high. Because \( \sigma_1 < \sigma_2 \) and the payoff function is linear when \( x < 1 \), it is also optimal to select project 2 when cash flow is low. However, when the cash flow is close to 1, the concavity of the payoff function implies that it is optimal for the agent to hedge and choose project 1. For simplicity, we construct this example with a piecewise linear payoff function \( f \). It is easy to alter \( f \) so that it is a smooth function without altering the structure of the optimal policy that is characterized by multiple switching triggers.

If we restrict the agent to switch projects at most once, then the restricted optimal policy is to always select project 2. The agent’s utility under the globally optimal policy (with two switching points as above) is 7.882, a 4.8% improvement compared to the value 7.520 for the restricted optimal strategy of choosing only project 2.

The above counterexample illustrates the importance of the conditions in Theorem 1 for policies with a single switching trigger to be optimal. It also shows the potential complexity of optimal risk-taking policies when the projects’ risk-adjusted drifts are not consistently ordered.

4. Multiple Projects

In this section, we consider the general problem where the agent dynamically selects among \( N \) projects, \( i = 1, 2, \ldots, N \). Without loss of generality, we assume that the volatilities of the projects satisfy

\[ \sigma_1 < \sigma_2 < \ldots < \sigma_N. \] (12)

We first show that the optimal policy only selects projects in the spanning subset that is defined as follows.
**Definition 2.** A subset \( \{i_j\}_{1 \leq j \leq K} \) where \( 1 = i_1 < i_2 < \ldots < i_K = N \) is called a spanning subset of the projects if

i) The drifts of projects in the subset are strictly concave in their variances,

\[
\mu_{ik} > \frac{(\sigma^2_i - \sigma^2_{ik})\mu_{ij} + (\sigma^2_{ik} - \sigma^2_{ij})\mu_{il}}{\sigma^2_i - \sigma^2_{ij}}, \quad \text{for } 1 \leq j < k < l \leq K.
\] (13)

ii) For any \( i_k < j < i_{k+1} \), the drifts of the projects \( \{i_k, j, i_{k+1}\} \) are weakly convex in their variances,

\[
\mu_j \leq \frac{(\sigma^2_{ik+1} - \sigma^2_j)\mu_{ik} + (\sigma^2_{j} - \sigma^2_{ik})\mu_{ik+1}}{\sigma^2_{ik+1} - \sigma^2_{ik}}.
\] (14)

Note that the spanning subset only depends on the drifts and volatilities of the projects and not on the agent’s payoff function. In the following lemma, we show that there always exists a unique spanning subset. In fact, conditions (13) and (14) imply that the spanning subset of projects correspond to the extreme points of the upper convex contour of the set of points \( \{(\mu_i, \sigma^2_i)\}_{i=1,\ldots,N} \) in \( \mathbb{R}^2 \).

**Lemma 1.** There exists a unique spanning subset for any given set of projects.

The following proposition shows that the agent’s optimal project choice policy only selects projects in the spanning subset. In this proposition, we assume that the agent’s optimal policy can be characterized by a finite number of switching thresholds of output levels at which the agent switches from one project to another. We later show that this is indeed the case.

**Proposition 2.** Assume that \( f \in C^2(\mathbb{R}_+) \) and the agent’s optimal policy is characterized by a finite number of switching thresholds.

i) The agent only selects projects in the spanning subset \( \{i_j\}_{1 \leq j \leq K} \) in the optimal policy.

ii) If the agent selects project \( i_j \) when the output is just above a switching threshold \( q \) and switches
to a project $i_k$ when the output level is just below $q$, then either $k = j - 1$ or $k = j + 1$. In other words, the agent only switches between adjacent projects in the spanning subset.

This proposition greatly facilitates the derivation of the optimal policy by substantially restricting not only the set of projects that may appear in the optimal policy, but also the possible switching scenarios. The spanning condition implies that any project $(\mu_i, \sigma_i^2)$ that is a convex combination of spanning projects is dominated by the spanning projects and is, therefore, not optimally chosen. To obtain some intuition for this result, it might be useful to compare the result to the well-known result that a convex function defined over a compact, convex set attains its maximum at an extreme point. As mentioned earlier, the spanning subset corresponds to the extreme points of the upper convex contour of the set of points defined by the drifts and variances of the projects. Building on the intuition for Theorem 1, the optimal policy trades off the risk-adjusted drifts and volatilities of the projects and selects the projects with higher risk-adjusted drifts at higher earnings levels and the projects with higher volatilities at lower earnings levels. Extending the intuition for case ii) of Theorem 1, when a project’s drift and variance lie in the (relative) interior of the convex contour of the set of drifts and variances of the projects, then it is dominated by spanning projects for all earnings levels so that it is not selected at all in the optimal policy.

Analogous to the two-project case, we consider the following “consistent ordering” condition for the risk-adjusted drifts of the spanning projects.

**Definition 3.** The risk-adjusted drifts of the spanning subset of projects are **consistently ordered** on an interval $(a, b)$ if there exists a permutation \( \{j_1, \ldots, j_K\} \) of the indices \( \{1, \ldots, K\} \) such that

\[
\tilde{\mu}_{ij_1}(x) \geq \tilde{\mu}_{ij_2}(x) \geq \cdots \geq \tilde{\mu}_{ij_K}(x), \quad \text{for all } a < x < b.
\]

The “consistent ordering” condition essentially implies that the ordinal ranking of the projects’ risk-adjusted drifts remains unaltered over the interval $(a, b)$. Note that the condition trivially holds
when the risk-adjusted drifts are constants as in the case where the payoff function $f$ has a power form, that is, $f(x) = x^c$, where $c$ is a constant.

We have the following characterization of the optimal policy of the agent if the risk-adjusted drifts of the spanning projects can be consistently ordered.

**Theorem 3.** Assume that $f \in C^2(\mathbb{R}_+)$ and the risk-adjusted drifts of the spanning subset of projects are consistently ordered on $(0, \infty)$. Let project $i_h$ be the spanning project with the highest risk-adjusted drift (if there are multiple such projects, let $i_h$ be the one with the highest index, that is, the highest volatility). There exist a threshold level $C^*_B \geq 0$ and switching points $C^*_B = q^*_K < q^*_{K-1} < \ldots < q^*_{m'} < q^*_{m'-1} = \infty$ with $m' \geq h$ such that the agent’s optimal project choices $\{p^*_t\}$ are as follows:

$$p^*_t = i_k, \text{ if } q^*_k \leq C_t < q^*_{k-1}, \text{ for } m' \leq k \leq K.$$ 

The agent optimally terminates the firm at the first time $\tau^*_B$ that the firm’s cash flow hits the level $C^*_B$.

The above theorem generalizes Theorem 1 to the case of multiple projects. In the optimal policy, the agent selects only the spanning project with the highest risk-adjusted drift and the ones with higher volatilities (project $i_h, i_{h+1}, \ldots, i_K$). Furthermore, the agent always switches from a spanning project to the next one with higher volatility as the firm’s earnings decline across the switching thresholds. The intuition for the nature of the switching policy can be understood from the results of Theorem 1. When the firm’s earnings are high, the probability of termination is very low so that it is optimal for the agent to select the project $i_h$ with the highest risk-adjusted drift because it has the highest present value. As the earnings decline, the termination probability increases so that it may be optimal for the agent to increase risk. Building on the intuition for part ii) of Theorem 1, it is sub-optimal for the agent to choose a project with a lower risk-adjusted drift and lower volatility than the project $i_h$ because it is dominated by project $i_h$.

As in the case of Theorem 1, the consistent ordering of the projects’ risk-adjusted drifts is the sufficient condition which ensures that the optimal policy is characterized by at most $K - 1$
switching triggers. We deal with the case where the risk-adjusted drifts are not consistently ordered shortly. As mentioned earlier, if the payoff function has a power form, then the risk-adjusted drifts of projects are constants and thus consistently ordered. Therefore, the optimal policy for an agent with such payoff functions and an arbitrary number of projects involves at most $K - 1$ switching points, where $K$ is the number of spanning projects.

The theorem is broadly applicable because we only need the projects in the spanning subset to be consistently ordered to characterize the optimal switching policy with multiple projects. In many applications of interest, the spanning subset may have much fewer projects than the original set of projects. For example, in our illustrative example 2 below, the spanning subset may contain only two projects.

In the proposition below, we show that, if $f$ is asymptotic to a CRRA function with risk aversion $0 \leq \gamma \leq 1$, then it is always optimal for the agent to adopt project $i_1$ when $C_t$ is sufficiently large. The intuition is that project $i_1$ has the highest risk-adjusted drift and thus the highest NPV when the output level is sufficiently high. We omit the proof because it is essentially identical to that of Proposition 1.

**Proposition 3.** [Finiteness of Switching Points] Assume that the conditions of Theorem 3 hold. Furthermore, for some $0 \leq \gamma \leq 1$,

$$f(x) = c_1 + c_2 \frac{1}{1 - \gamma} x^{1-\gamma} + o(x^{1-\gamma})$$

for $x$ sufficiently large, with $c_2 > 0$ (when $\gamma = 1$, assume $c_1 > 0$ and $c_2 = 0$), then $m' = 1$ in Theorem 3.

In the following theorem, we consider the general case where the risk-adjusted drifts of spanning projects are not globally consistently ordered and characterize the optimal policy.

**Theorem 4.** Assume that the payoff function $f \in C^2$ and that there are $0 = x_0 < x_1 < x_2 < \ldots < x_{M-1} < x_M = \infty$ such that for any $1 \leq i \leq M$, the risk-adjusted drifts of the spanning subset
of projects are consistently ranked on \((x_{i-1}, x_i)\). Then the total number of switching points in the optimal policy is bounded above by \(M(K - 1)\).

The proof of this theorem and the intuition underlying it are extensions of that of Theorem 2 in the two-project case. We, therefore, omit the proof for brevity.

Illustrative Examples

Example 1. A Case with Three Projects

Consider three projects with the following characteristics: \(\sigma_1 = 0.1, \sigma_2 = 0.2, \sigma_3 = 0.3, \mu_1 = 0.02, \mu_3 = 0\). We let \(\mu_2\) vary over a range of possible values and examine how the optimal project choices change. For concreteness, we fix the discount rate \(r = 0.04\) and assume that the payoff function is linear and \(f(x) = x - 1\), but the results do not depend on the discount rate or the curvature of the payoff function. Since the curvature of \(f\) is zero, the risk-adjusted drifts of the three projects are monotone in their volatilities when \(0 = \mu_3 < \mu_2 < \mu_1 = 0.02\). A critical value for the optimal policy is

\[
\tilde{\mu}_2 = \frac{(\sigma_3^2 - \sigma_2^2)\mu_1 + (\sigma_2^2 - \sigma_1^2)\mu_3}{\sigma_3^2 - \sigma_1^2} = 0.0125.
\]

When \(\mu_2 > \tilde{\mu}_2\), the projects satisfy the strict concavity condition (B.3). By Theorem 3, there exists switching thresholds \(q_1^* > q_2^*\) and bankruptcy boundary \(C_B^*\) such that the optimal project choices are

\[
\tilde{p}_t^* = \begin{cases} 
1, & \text{if } C_t \geq q_1^*, \\
2, & \text{if } q_2^* \leq C_t < q_1^*, \\
3, & \text{if } C_B^* \leq C_t < q_2^*.
\end{cases}
\] (15)

If \(\mu_2 \leq \tilde{\mu}_2\), project 2 is “dominated” by projects 1 and 3 in the sense that \((1, 3)\) constitute a spanning subset of the projects. Therefore, by Proposition 2 and Theorem 3, only projects 1 and
Figure 3: Example with three projects: optimal strategy when $\mu_2$ changes.

3 are selected in the optimal policy. There exists a single switching threshold $q_1^*$ and bankruptcy boundary $C_B^*$ such that the optimal project choices are

$$p_t^* = \begin{cases} 
1, & \text{if } C_t \geq q_1^*, \\
3, & \text{if } C_B^* \leq C_t < q_1^*.
\end{cases}$$

(16)

Figure 3 plots the optimal switching and bankruptcy triggers for different values of $\mu_2$. When $\mu_2 > \bar{\mu}_2$, the optimal policy is consistent with (15), with two different switching points $q_1^* > q_2^*$. When $\mu_2 \leq \bar{\mu}_2$, the figure shows that the two triggers coincide, i.e., $q_1^* = q_2^*$. The agent finds it optimal to switch directly from project 3 to 1 when output level crosses the trigger $q_1^*$, consistent with the one-trigger policy given in (16). In the figure, the lines given by the optimal triggers divide the area above bankruptcy into three regions, which correspond to the parameter values for which one of the three projects is optimal.

Example 2. CRRA payoff function and projects with drifts linear in volatilities

We now apply our general results to the case where the drifts of $N$ projects are linear in their volatilities and the agent’s payoff function is a CRRA function with relative risk aversion $\gamma$. We

---

3In this example, if $\mu_2 > 0.02$, then the optimal policy is a one-trigger policy involving only projects 2 and 3, because in that case project 2 has the highest drift and thus by Theorem 3 project 1 does not appear in the optimal policy.
will see that in this case the optimal risk-taking policy has a particularly simple form. Assume that

\[ \mu_i = a + b \sigma_i, \quad 1 \leq i \leq N. \]

Then

\[ \tilde{\mu}_i = a + b \sigma_i - \frac{1}{2} \gamma \sigma_i^2 = f(\sigma_i^2), \]

where

\[ f(x) = a + b \sqrt{x} - \frac{1}{2} \gamma x. \]

Case i). \( b > 0 \), i.e., the projects’ drifts increase in volatilities. In this case, \( f(x) \) is a strictly concave function of \( x \) and therefore the entire set of projects constitutes the spanning subset. Let project \( i_0 \) be the one that minimizes \( |\sigma_i - \frac{b}{\gamma}| \), then \( i_0 \) has the highest risk-adjusted drift. By Theorem 3, the optimal policy is given by at most \( N - i_0 \) unique switching thresholds \( C_{B}^* \leq q_{N-1}^* \leq \ldots \leq q_{i_0}^* \leq \infty \). The agent selects project \( i \) when \( q_i^* \leq C_t < q_{i-1}^* \).

Case ii). \( b \leq 0 \), i.e., the projects’ drifts decrease in volatilities. In this case, \( f(x) \) is a weakly convex function of \( x \) and it is easy to see that projects 1 and \( N \) form a spanning subset of all the projects. Therefore, only projects 1 and \( N \) will be selected in the optimal policy. Since the risk-adjusted drift of project 1 is higher than that of project \( N \), there exists a unique switching trigger \( q^* \geq C_{B}^* \) such that project 1 is selected when \( C_t \geq q^* \) and project \( N \) is selected when \( C_t < q^* \).

5. Analytical Characterization and Computational Algorithm

In this section, we describe the algorithm for the computation of the optimal risk-taking policy for an arbitrary set of projects, and provide the associated analytical expressions. As in the previous section, there are \( N \) projects with drifts and volatilities \( \{\mu_i, \sigma_i\}_{i=1,\ldots,N} \) with \( \sigma_1 < \sigma_2 < \ldots < \sigma_N \). The agent’s payoff function is \( f(x) \) so that her instantaneous cash flow is given by \( f(C_t)dt \) when the output is \( C_t \). There are four steps in the computational algorithm.

**Step 1. Finding the spanning subset of projects**

We first determine the unique spanning subset of the set of projects \( \{1, \ldots, N\} \). As discussed
earlier, this is equivalent to finding the extreme points of the upper convex contour of the points \( \{(\mu_i, \sigma_i^2)\}_{i=1,...,N} \) in \( \mathbb{R}^2 \) that correspond to projects in the spanning subset. (The proof of Lemma 1 provides an explicit algorithm of finding the spanning subset.) Suppose that \( \{i_j\}_{j=1,...,K} \) is the unique spanning subset of projects with \( i_1 < i_2 < \ldots < i_K \). We note that in the case of \( N = 2 \) projects, this step can be omitted as the set of the two projects spans itself.

**Step 2. Checking the consistent ordering condition**

In this step, we determine whether the risk-adjusted drifts of the spanning subset, \( \{\bar{\mu}_{ij}(x)\}_{j=1,...,K} \), are consistently ordered for \( x > 0 \). In applications, the agent’s payoff function and thus the risk-adjusted drifts usually have an explicit functional form (such as piecewise polynomials) in which case the consistent ordering condition is straightforward to verify.\(^4\)

**Step 3. Closed-Form Representation of the Agent’s Value Function**

The proof of Theorem 3 provides us with analytical characterizations of the agent’s optimal value function given a set of switching thresholds and the bankruptcy trigger. Let \( i_h \) be the spanning project with the highest risk-adjusted drift. We consider the generic case where the switching thresholds and the bankruptcy trigger satisfy \( C_B = q_i < q_{i_K} < q_{i_{K-1}} < \ldots < q_i < q_{i_{h-1}} = \infty \), i.e., the agent’s optimal policy consists of choosing project \( i_j \) when the output \( q_{i_{j+1}} \leq C_t < q_i \) for \( j = h, \ldots, K \) and declaring bankruptcy when \( C_t = C_B \). The other cases can be analyzed similarly.

We first obtain a special solution \( g(x) \in C^1([C_B, \infty)) \) of the following piecewise linear second-order ODE:

\[
L_{ij}g(x) + f(x) = 0, \quad q_{i_{j+1}} \leq C < q_{i_j}, \quad j = h, \ldots, K,
\]

where \( L_k g = \frac{1}{2} \sigma_k^2 \frac{d^2 g}{dx^2} + \mu_k \frac{dg}{dx} - rg \). In applications, the payoff function \( f(x) \) usually has a closed-form representation (such as piecewise polynomial functions) that allows us to obtain \( g(x) \) in closed form.

\(^4\)If no consistent order exists, then by Theorem 4, we can still analytically characterize the optimal policy, though with more switching points. For the sake of space, we omit the details here. The results are available from the authors upon request.
Next, the agent’s value function can be written in the following closed form,

\[ S(C) = \begin{cases} 
    A_1 C^{\gamma_1^h} + g(C), & \text{if } C > q_h \\
    A_2 C^{\gamma_{h+1}^+} + A_3 C^{\gamma_h^-} + g(C), & \text{if } q_{h+1} \leq C < q_h \\
    \vdots \\
    A_{2K-2h} C^{\gamma^+_K} + A_{2K-2h+1} C^{\gamma^-_K} + g(C), & \text{if } C_B \leq C < q_{K-1}
\end{cases} \tag{17} \]

where \( A_1, \ldots, A_{2K-1} \) are constants and \( \gamma_{ij}^\pm \) are the positive and negative roots of the characteristic equations

\[ \frac{\sigma_{ij}^2}{2} \gamma^2 + (\mu_{ij} - \frac{\sigma_{ij}^2}{2}) \gamma - r = 0, \quad j = h, \ldots, K. \tag{18} \]

The coefficients \( \{A_j\}_{j=1, \ldots, 2K-2h+1} \) are determined by the condition that the value function \( S \) is continuously differentiable at each switching threshold \( q_j \), i.e., it satisfies the following conditions,

\[ S|_{x=q_j^+} = S|_{x=q_j^-}, \quad \left. \frac{\partial S}{\partial x} \right|_{x=q_j^+} = \left. \frac{\partial S}{\partial x} \right|_{x=q_j^-}, \quad j = h, \ldots, K - 1, \tag{19} \]

and the boundary condition at the bankruptcy trigger,

\[ S(C_B) = 0. \tag{20} \]

Since the closed-form representation of \( S(C) \) in (17) is linear in the coefficients \( \{A_j\} \), the equations (19) and (20) form a system of \( 2K - 2h + 1 \) linear equations. We can thus obtain the unique values of \( \{A_j\}_{j=1, \ldots, 2K-2h+1} \) by solving this linear system.

**Step 4. Finding the optimal switching thresholds and bankruptcy trigger.**

From the proof of Theorem 3, the switching thresholds are determined by the following \( K - h \) "super-contact" conditions,

\[ \left. \frac{\partial^2 S}{\partial x^2} \right|_{x=q_j^+} = \left. \frac{\partial^2 S}{\partial x^2} \right|_{x=q_j^-}, \quad j = h, \ldots, K - 1 \tag{21} \]
and the bankruptcy trigger $C_B$ is determined by the “smooth-pasting” condition,

$$S'(C_B) = 0. \quad (22)$$

Since we have obtained the closed-form representation of $S(C)$ from the previous step, equations (21) and (22) provide $K - h + 1$ non-linear equations that we can solve and determine the $K - h$ switching thresholds and the bankruptcy trigger.


In this section, we present a concrete application of our general results to the investigation of the optimal capital structure choices of firms in the presence of shareholder-debtholder agency conflicts arising from risk-taking/asset substitution by shareholders. Specifically, we analyze a structural model of capital structure in which a firm can dynamically switch between projects with different drifts and volatilities. The firm’s capital structure reflects the tradeoff between the tax advantages of debt and financial distress costs that include the agency costs of asset substitution (Leland (1998)).

As we elaborate in more detail below, our framework is particularly well suited to investigate how the firm’s optimal capital structure, risk policy, and agency costs of asset substitution change when its systematic risk varies. The implications of our analysis shed light on how aggregate or systematic uncertainty influences firms’ capital structure and risk-taking decisions. As our subsequent discussion will make clear, these questions are difficult to fully analyze in traditional “contingent claims” models of capital structure.

6.1. Model Setup

The earnings before interest and taxes (EBIT) of a firm evolve as follows:

$$\frac{dC_t}{C_t} = \bar{\mu} p_t dt + \sigma p_t C_t d\hat{B}_t, \quad (23)$$

where $\hat{B}_t$ is a standard Brownian motion under the physical or “real world” probability measure.
The firm may dynamically choose from a menu of two projects, \( p_t \in \{1, 2\} \). The drifts and volatilities of the projects satisfy
\[
\hat{\mu}_1 < \hat{\mu}_2 \text{ and } \sigma_1 < \sigma_2,
\]
which captures the intuitive notion that the drift/expected growth of the firm’s earnings is higher under the higher risk project.

By the theory of contingent claims valuation, security prices are given by their expected discounted payoff streams, where the expectation is under the risk-neutral or pricing measure and the discount rate is the risk-free rate (see Duffie (2001)). For simplicity, we assume that the market price of risk or Sharpe Ratio, \( \lambda \), is the same for the two projects\(^5\). The firm’s EBIT process under the risk-neutral measure is given by
\[
\frac{dC_t}{C_t} = \mu_{p_t} dt + \sigma_{p_t} dB_t
\]
where \( B_t \) is a standard Brownian motion under the risk-neutral measure and the risk neutral drifts are
\[
\mu_i = \hat{\mu}_i - \lambda \sigma_i, \quad i = 1, 2.
\]

In (26), \( \lambda \sigma_i \), which is the difference between the physical and risk-neutral drifts, is the risk premium of the project that increases with its systematic risk, and the market price of risk/Sharpe Ratio \( \lambda \) increases with the proportion of the project’s total risk that is systematic.\(^6\)

\(^5\)Allowing different market prices of risk for the two projects will introduce an additional parameter in the model and will not change the results qualitatively.
\(^6\)To see this, rewrite the firm’s earnings under the physical measure in (23) as follows.
\[
\frac{dC_t}{C_t} = \mu_{p_t} dt + \sigma_{p_t}^{\text{systematic}} d\tilde{B}_t^{\text{systematic}} + \sigma_{p_t}^{\text{idiosyncratic}} d\tilde{B}_t^{\text{idiosyncratic}},
\]
where \( \tilde{B}_t^{\text{systematic}} \) and \( \tilde{B}_t^{\text{idiosyncratic}} \) are orthogonal Brownian motions representing the systematic and idiosyncratic components of the project’s risk, and \( (\sigma_{p_t}^{\text{systematic}})^2 + (\sigma_{p_t}^{\text{idiosyncratic}})^2 = (\sigma_{p_t})^2 \). If \( \lambda^{S} \) is the market price of the project’s systematic risk (the market price of idiosyncratic risk is zero), then
\[
\mu_i = \hat{\mu}_i - \lambda^{S} \sigma_i^{\text{systematic}}.
\]
Comparing (26) with (28), we see that
\[
\lambda = \lambda^{S} \frac{\sigma_i^{\text{systematic}}}{\sigma_i}.
\]
Consequently, \( \lambda \) increases with the ratio of systematic risk to total risk.
Remark 1. By (26), it is easy to see that, if \( \lambda \) is above a threshold, the relative ranking of the projects’ risk-neutral drifts differs from that of the physical drifts, that is, the higher risk project has a lower risk-neutral drift. As our subsequent analysis makes clear, this observation has important implications for the effects of systematic risk on the firm’s risk choices and capital structure. Previous “contingent claims” models of capital structure (e.g., Leland (1994, 1998), Goldstein, Ju and Leland (2001), Strebulaev (2007)) directly work under the risk-neutral measure, which makes it difficult to disentangle the effects of the physical drifts of projects’ earnings from those of the market price of risk and systematic risk. In fact, in the study that is most directly comparable with our present analysis, Leland (1998) models the unlevered asset value instead of the EBIT process. Apart from some important conceptual issues associated with modeling the unlevered assets instead of the EBIT process (see Goldstein et al. (2001) for a discussion), the risk-neutral drifts of the unlevered asset value process under the two possible projects in Leland’s (1998) analysis are equal, which eliminates the possibility of the aforementioned effects of systematic risk on risk-taking and capital structure.

At time \( t = 0 \), the firm issues long-term (infinite horizon and completely amortized) debt that pays a perpetual coupon \( \theta \) per unit time. Coupon payments are shielded from corporate taxes. The corporate tax rate is \( \tau > 0 \). Thus the instantaneous cash flow to shareholders at time \( t \) is \( (1 - \tau)(C_t - \theta)dt \) and the cash flow to debt holders is \( \theta dt \). As in Leland (1998), shareholders dynamically select the firm’s projects and endogenously declare bankruptcy at the bankruptcy (stopping) time \( \tau_B \) when it is no longer optimal for them to continue servicing debt. The original shareholders of the unlevered firm choose the initial leverage at time \( t = 0 \) to maximize total firm value that is given by

\[
\text{Firm Value} = E_0 \left[ \int_0^{\tau_B} \left\{ \text{total after-tax cash flow to firm} \right\} dt + \text{bankruptcy payoff} \right], \tag{30}
\]

where \( \tau_B \) is the bankruptcy stopping time. As in Leland (1998), the firm incurs bankruptcy costs that are a proportion \( \alpha \in (0, 1) \) of the unlevered firm value at bankruptcy so that the bankruptcy payoff \( V_B \) is a proportion \((1 - \alpha)\) of the unlevered firm value at the bankruptcy time, \( \tau_B \).
Given the initial leverage choice that is determined by $\theta$, the shareholders’ problem is a special case of Theorem 1, where the payoff function is given by $f(x) = (1 - \tau)(x - \theta)$. Applying the theorem, we have the following results:

i) If $\mu_1 < \mu_2$, there exists a bankruptcy threshold $C_B > 0$ such that shareholders always select project 2 when $C_B \leq C < \infty$;

ii) if $\mu_1 \geq \mu_2$, there exists a unique switching threshold $C_S$ and bankruptcy threshold $C_B > 0$ such that the shareholders choose project 1 when $C > C_S$ and project 2 when $C_B \leq C \leq C_S$.

In the two cases above, we use the same variables to denote the switching and bankruptcy thresholds to avoid cluttering the notation, but it should be borne in mind that they differ in the two cases. From the above, we see that, in case (i), shareholders optimally choose the “high risk” project always. In case (ii), however, they choose the high risk project in financial distress and the “low risk” project when earnings are above a threshold. We, therefore, see that, as alluded to in Remark 1, the optimal asset substitution policy differs dramatically depending on the ranking of the projects’ risk-neutral drifts relative to the physical drifts.

By (24), there exist unique constants $\bar{\mu}$ and $\beta > 0$ such that

$$\hat{\mu}_i = \bar{\mu} + \beta \sigma_i, \quad i = 1, 2. \quad (31)$$

From (26),

$$\mu_i = \bar{\mu} + (\beta - \lambda)\sigma_i, \quad i = 1, 2 \quad (32)$$

Equation (32) implies that the critical point where the ranking of the projects’ risk-neutral drifts changes is $\lambda = \beta$. Consequently, as the market price of risk or Sharpe Ratio increases above this threshold, the optimal asset substitution policy changes from choosing the “high risk” project always to one that chooses the “high risk” project in bad states and the “low risk” project in good ones. As we see shortly, this behavior has significant consequences for the firm’s capital structure and the agency costs of asset substitution.
6.2. Calibration and Quantitative Analysis

To obtain quantitative implications for the optimal capital structure, risk-taking policy, and agency costs, we calibrate the baseline parameters of our model as follows. We set the effective corporate tax rate $\tau$ to 0.2 following Leland (1994, 1998), which reflects a corporate tax rate of 0.35 and the difference between personal tax rates on interest payments and dividends or capital gains. The bankruptcy cost $\alpha = 0.20$ falls within the range of such costs estimated by Andrade and Kaplan (1998) and Davydenko, Strebulaev, and Zhao (2011). The risk-free rate $r = 0.053$ is the average short-term T-bill rate from 1962 to 2009.

We use corporate data from Compustat to calibrate the drifts and volatilities of the projects. We choose the sample of large U.S. public nonfinancial and nonutility firms (book assets $\geq$ $100 million) over the period from 1962 to 2009. We also require these firms to have monthly stock returns from CRSP. We use the formula $\sigma_A = \sqrt{(1 - L)^2 \sigma_E^2 + L^2 \sigma_D^2}$ to estimate the asset volatility of a firm. We compute the equity volatility $\sigma_E$ from lagged 36 monthly stock returns, $\sigma_D = 0.081$ is the average 10-year T-bond return volatility, and $L$ is the market leverage of the firm.\(^7\) Within each industry (Fama and French 48 industry classification), we separate the firms into two subsamples of high and low volatility using the median volatility in the industry. Intuitively, we can regard these two samples as the high risk project and the low risk project in the model. We compute the mean asset volatility for these two samples and then average across industries to obtain the estimates $\sigma_1 = 0.16$ and $\sigma_2 = 0.33$. The risk-neutral drifts can be estimated via the approximate formula

$$V = \frac{(1 - \tau)C}{r - \mu} + \frac{\tau \theta}{r},$$

or equivalently,

$$\mu = r - \frac{(1 - \tau)C}{1 - \tau \theta},$$

(33)

where $\frac{C}{r}$ and $\frac{\theta}{r}$ can be estimated by the ratios of earnings and interest payments to total assets. We use (33) to compute the risk-neutral drifts for the high and low volatility samples and then average across industries and obtain $\mu_1 = -0.030$ and $\mu_2 = -0.041$. Note here that the risk-neutral

\(^7\)This is a simplified version of the more elaborate estimation of asset volatility in Schaefer and Strebulaev (2008).
drifts can be negative since they are adjusted for risk.

To estimate the actual or physical drifts, we first estimate the average Sharpe ratio \( \lambda = 0.45 \) using monthly value-weighted market returns over the period 1962 to 2009 obtained from CRSP. We then use (26) to obtain the physical drifts \( \hat{\mu}_1 = 0.042 \) and \( \hat{\mu}_2 = 0.109 \). The following table summarizes the baseline parameters of the model:

<table>
<thead>
<tr>
<th>Description</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk-free rate</td>
<td>( r )</td>
<td>0.053</td>
</tr>
<tr>
<td>Market Price of Risk/Sharpe Ratio</td>
<td>( \lambda )</td>
<td>0.45</td>
</tr>
<tr>
<td>Physical Drifts</td>
<td>( \hat{\mu}_1 )</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>( \hat{\mu}_2 )</td>
<td>0.109</td>
</tr>
<tr>
<td>Volatilities</td>
<td>( \sigma_1 )</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>( \sigma_2 )</td>
<td>0.33</td>
</tr>
<tr>
<td>Effective Tax Rate</td>
<td>( \tau )</td>
<td>0.20</td>
</tr>
<tr>
<td>Bankruptcy cost</td>
<td>( \alpha )</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Next, we compute the optimal risk-taking policy, the optimal leverage, the firm’s total value (equity value + debt value), and the credit spread as the market price of risk, \( \lambda \), varies keeping the other parameters fixed at their baseline values. In the numerical procedure, we first determine the optimal switching threshold \( C^*_S \) and the bankruptcy trigger \( C^*_B \) for a given initial debt level \( \theta \). We then determine the optimal debt level \( \theta^* \) by maximizing the initial firm value.\(^8\)

Figure 4 presents the results. We first observe that (31) and (32) imply that, at the critical point \( \lambda = \beta = 0.39 \), the risk-neutral drifts of the two projects are equal. Panel A of Figure 4 shows that when \( \lambda \leq 0.39 \), the shareholders always choose the high volatility project 2 which also has a higher risk-neutral drift, consistent with the result in Theorem 1 case ii). When \( \lambda > 0.39 \), shareholders prefer project 1 at high cash flow levels and project 2 at low cash flow levels. The optimal switching threshold \( C^*_S \) is decreasing in \( \lambda \) and approaches \( \infty \) as \( \lambda \searrow 0.39 \). This is consistent with our analysis in Section 6.1. Panel B shows that the firm value is a strictly decreasing function

\(^8\)Since the model is homogeneous of degree one in the pair \( (C, \theta) \), the optimal policies depend linearly on the debt level \( \theta \). This scaling-invariant feature simplifies the numerical procedure.
of the market price of risk $\lambda$, because the risk-neutral drifts are negatively related to $\lambda$.

Interestingly, we find that the optimal leverage is non-monotonic in $\lambda$ (Panel C of Figure 4). The optimal leverage ratio is relatively stable when $\lambda \leq 0.39$ with values close to 60%, but is sharply decreasing when $0.39 \leq \lambda \leq 0.41$ and reaches a minimum of 25% when $\lambda = 0.41$. The leverage increases with $\lambda$ when $\lambda > 0.41$ but does not increase to levels above 40%. The non-monotonicity of the optimal leverage ratios comes from the tradeoff between asset substitution costs and tax benefits. Intuitively, higher leverage provides higher tax benefits, but also lead to
higher asset substitution costs because shareholders switch to the high-risk project sooner. When \( \lambda \) is close to 0.41, asset substitution is a more significant concern so that the optimal leverage is low. At the baseline value of \( \lambda = 0.45 \), the average leverage ratio is 33%, very close to the average market leverage ratio 31% in the empirical sample. It is worth noting that our model can generate leverage ratios that are significantly lower than in models such as Leland (1994, 1998) even though we don’t have dynamic debt restructuring. Our model can, therefore, help to explain the fact that non-financial firms’ leverage ratios are much lower than the levels obtained by traditional trade-off models. Finally, Panel D of Figure 4 shows that, as with the leverage ratio, the credit spread also varies non-monotonically with \( \lambda \). We, therefore, obtain an interesting link between the market price of risk and credit spreads. It is worth noting that the credit spreads are higher than those predicted by Leland (1998) that are around 100 basis points despite the significantly lower leverage ratios in our model.

To quantify the agency costs of asset substitution, we follow Leland (1998) by considering an alternative “benchmark” model in which the firm’s capital structure and project choices maximize the total firm value. As in (30), the instantaneous cash flow to the firm is the sum of the cash flows to shareholders and debtholders and is equal to \( [(1 - \tau)C_t + \tau \theta] dt \).

We compute the optimal risk-taking policies in this benchmark problem as well as the the associated firm values and leverage ratios. We compute the agency costs as the difference between the firm values in the benchmark model and in the actual model where project choices maximize shareholder value instead of total firm value.

Panel E of Figure 4 shows that the leverage ratios in the benchmark model are also non-monotonic, but are relatively stable in the range [54%, 68%]. This is in stark contrast with the actual model, in which the leverage can reach a minimum of 25% (for \( \lambda = 0.41 \)). We also note that the low leverage phenomenon in the actual model occurs only when the drifts of the two projects are sufficiently different, which explains why low leverage ratios cannot be generated in the Leland (1998) model where the projects’ risk-neutral drifts are equal.

Panel F shows the agency costs of asset substitution as percentages of firm values. The agency costs (as percentage of firm values) are 4.4% at the baseline value \( \lambda = 0.45 \) and achieve a maximum of 6.1% at \( \lambda = 0.41 \). Interestingly, this maximum point is also the point where the leverage ratio
in the actual model is lowest. These results suggest that the asset substitution behavior can have substantial agency costs and lead to significantly lower leverage ratios relative to the benchmark model. It is also worth noting that the agency costs of asset substitution predicted by our model are significantly higher than those predicted by Leland (1998), which are around 1%. Our analysis shows that the agency costs of asset substitution could, in fact, be substantial and significantly influence leverage levels and credit spreads.

7. Conclusions

We analyze a class of continuous-time models in which an agent controls the evolution of an output process that affects her payoff stream by dynamically selecting among a set of arbitrary, but finite, number of projects. The log-normal output process evolves with differing drifts and volatilities under the feasible projects. The models we analyze are broadly applicable to the study of a number of important problems in financial economics. To the best of our knowledge, ours is the first paper to rigorously analyze this class of models and provide analytical characterizations of the optimal policies. We show that the optimal policy is determined by the risk-adjusted drifts of the projects, which incorporate the projects’ drifts and volatilities and the curvature of the agent’s payoff function. The optimal policy only selects projects in the spanning subset, which comprises of projects that correspond to the extreme points of the upper convex contour of the set of points defined by the drifts and variances of the original set of projects. If the number of spanning projects is \( K \), and their risk-adjusted drifts can be consistently ordered, then the optimal policy is characterized by at most \( K - 1 \) unique switching triggers at which the agent switches from projects with higher risk-adjusted drifts to lower risk-adjusted drifts. In other words, the consistent ordering of the projects’ risk-adjusted drifts implies the “monotonicity” of the optimal policy. We generalize our results to the case where the projects may not be consistently ordered and provide bounds on the number of switching points in the optimal policy. Our results have applications to the general class of “corporate risk-taking” problems such as dynamic asset substitution by shareholders of a leverage firm, dynamic risk-taking/risk management by firm managers, and the
prudential regulation of financial institutions (see e.g., Subramanian and Yang (2012)).

We present a concrete application of our general results to the investigation of the capital structure choices of firms in the presence of shareholder-debtholder agency conflicts due to risk-taking/asset substitution. We obtain novel implications for the effects of systematic risk on capital structure, credit spreads and agency costs. In particular, our framework can generate low leverage ratios and high credit spreads consistent with values observed in the data. Further, in sharp contrast with previous literature that has examined the problem in more special settings, our results suggest that the agency costs of asset substitution are, in fact, substantial and are important determinants of capital structure and credit spreads.

Appendix A. Proofs in the Case of Two Projects

We begin by stating the relevant dynamic programming verification theorem for our analysis.

**Proposition 4.** [Dynamic Programming Verification Theorem] Let $S_q(C)$ denote the value of the switching policy when the state variable takes the value $C$ where project 1 is chosen when the state variable $C$ exceeds $q$ and project 2 is chosen when the state variable is below $q$. Suppose that $S_q$ satisfies the following Hamilton-Jacobi-Bellman (HJB) equation:

$$
\max_{i \in \{1,2\}} L_i S_q + f(C) = 0,
$$

$$
S_q(C_B) = S'_q(C_B) = 0,
$$

$$
S_q(C) \leq c(1 + |C|), \quad \text{for } C > 0
$$

where $c$ is a constant and

$$
L_i S_q = \frac{1}{2} \sigma_i^2 C^2 \frac{S''_q}{dC^2} + \mu_i C \frac{dS_q}{dC} - r S_q, \quad i \in \{1,2\}
$$

Then $S_q$ is the optimal value function that solves problem (7), $q$ is the optimal switching trigger and $C_B$ is the optimal termination boundary.

Since the above follows from the general verification theorem for dynamic programming, we omit its proof for brevity and refer the reader to Fleming and Soner (2005) (Theorem IV.3.1).

We begin the proof by proving a series of lemmas. The following lemma is useful in establishing bounds of derivatives on functions we consider.

**Lemma 2.** [Bounds for Derivatives] Let $L = \frac{1}{2} \sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{dx}{dx} - r$, where $\sigma > 0, \mu, r$ are constants. If $g \in C^2(R_+)$ satisfies

$$
|Lg(x)|, |g(x)| \leq c_1(1 + x), \quad x \geq a,
$$

(A.3)
for some \( a > 0 \) and constant \( c_1 > 0 \), then

\[
|x^2 g''(x)| + |xg'(x)| \leq c(1 + x), \quad x \geq a,
\]

for some constant \( c \).

**Proof.** Consider the substitution of variables \( y = \log(x) \) and \( h(y) = g(x) = g(e^y) \). Then \( h'(y) = e^y g'(e^y) = xg'(x) \) and \( h''(y) = x^2 g''(x) + xg'(x) \), and \( Lg(x) = \frac{1}{2} \sigma^2 h''(y) + (\mu - \frac{1}{2} \sigma^2)h'(y) - rh(y) \equiv \tilde{L}h(y) \). The bounds in (A.3) translate into

\[
|\tilde{L}h(y)|, |h(y)| \leq c_1(1 + e^y) \quad y \geq b = \log(a).
\]

To show (A.4), it suffices to show that \( |h''(y)| + |h'(y)| \leq c(1 + |e^y|) \), \( y \geq b \), for some constant \( c > 0 \). Because \( \tilde{L}h \) is a linear combination of \( h'', h' \) and \( h \), with a non-zero coefficient on \( h'' \), it suffices to show that

\[
|h'(y)| \leq c_2 e^y, \quad y \geq b,
\]

for some constant \( c_2 \). Assume to the contrary that, for any constant \( A > 0 \), there exists a \( y_0 > b \) such that

\[
|h'(y_0)| > A(1 + e^{y_0}).
\]

Without loss of generality, assume \( h'(y_0) > 0 \). Let \( \lambda = \frac{\mu - \frac{1}{2} \sigma^2}{2 \sigma^2} \), then

\[
\tilde{L}h(y) = \frac{1}{2} \sigma^2 [h''(y) + \lambda h'(y)] - rh(y).
\]

By (A.5),

\[
\left| \left( e^{\lambda y} h'(y) \right)' \right| = \left| e^{\lambda y} (h''(y) + \lambda h'(y)) \right| = \left| e^{\lambda y} \frac{2}{\sigma^2} (\tilde{L}h(y) + rh(y)) \right|
\]

\[
\leq c_3 e^{(1 + \lambda) y}.
\]

Therefore, for \( 0 \leq s \leq 1 \),

\[
\left| e^{\lambda (y_0 + s)} h'(y_0 + s) - e^{\lambda y_0} h'(y_0) \right| = \left| \int_{y_0}^{y_0 + s} \left( e^{\lambda y} h'(y) \right)' dy \right|
\]

\[
\leq c_3 \int_{y_0}^{y_0 + s} e^{(1 + \lambda) y} dy \leq c_4 e^{(1 + \lambda) y_0},
\]

(B.8)

By (A.8) and (A.7),

\[
h'(y_0 + s) \geq e^{-\lambda s} h'(y_0) - c_4 e^{y_0} > \frac{A}{2} e^{y_0},
\]

(A.9)

if \( A > 2 \max(1, e^\lambda) c_4 \). Now

\[
h(y_0 + 1) - h(y_0) = \int_{y_0}^{y_0 + 1} h'(y_0 + s) ds \geq \frac{A}{2} e^{y_0}.
\]

(A.10)
However, by (A.5),
\[ h(y_0 + 1) - h(y_0) \leq 2c_1(1 + e)^y_0. \]  
(A.11)
If \( A > 4c_1(1 + e) \), (A.10) and (A.11) lead to a contradiction. Therefore, (A.6) holds for \( c_2 \) sufficiently large and the lemma is proved.

The following lemma is a converse of the well-known Dynkin’s lemma. It states that if a function satisfies certain linear bounds and solves a class of differential equations associated with diffusion processes, it can be presented via an integration formula as in the Dynkin’s lemma.

**Lemma 3.** [Converse of Dynkin’s Lemma] Let \( L = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{dx}{dx} - r \), where \( \sigma > 0, \mu < r \) are constants. Assume \( g \in C^2(R_+) \) satisfies
\[ |Lg(x)|, |g(x)| \leq c(1 + |x|), \quad \text{for} \quad x \geq a. \]

Let \( x_0 > a \) and \( \{y_t\}_{t \geq 0} \) be the diffusion process given by \( dy_t = \mu y_t dt + \sigma y_t dB_t \), \( y_0 = x_0 \), and let \( \tau_a \) be the first time \( y_t = a \). Then
\[ g(x_0) = g(a) - E \left[ \int_{0}^{\tau_a} e^{-rs}(Lg)(y_s) ds \right]. \]  
(A.12)

**Proof.** Define \( h_t = g(a) - E_t \left[ \int_{0}^{\tau_a} e^{-r(s-t)}(Lg)(y_s) ds \right] \). Then \( h_t = h(y_t) \) because the process \( y_t \) is a Markov process. Also, \( |h(x)| \leq c(1 + |x|) \), due to the bound on \( Lg \) and the integration presentation of \( h \) (note that \( \int_{0}^{\tau_a} e^{-rs}(1 + C_s) ds \leq \frac{1}{r}(1 + C_0) \)). By Ito’s lemma, \( Lh(x) = Lg(x) \) for \( x \geq a \). Therefore, \( L(h-g) = 0 \), for \( x \geq a \), and thus \( h(x) - g(x) = a_1 x^{\gamma^-} + a_2 x^{\gamma^+} \), where \( \gamma^\pm \) are the positive and negative roots of the characteristic equation of operator \( L \), \( \frac{1}{2}\sigma^2 \gamma(\gamma - 1) + \mu \gamma - r = 0 \). It is easy to see from \( \mu < r \) that \( \gamma^+ > 1 \). Since \( g \) and \( h \) are bounded above by linear functions, we have \( a_2 = 0 \). Now \( h(a) = g(a) \) implies that \( a_1 = 0 \). Therefore, \( h = g \). Now \( h_0 = h(x_0) = g(x_0) \) and thus (A.12) holds.

The following lemma allows the comparison of two linearly bounded functions via comparison of the operator \( L \) applied to them.

**Lemma 4.** [Comparison Theorem] Let \( L = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{dx}{dx} - r \), where \( \sigma > 0, \mu < r \) are constants. Assume \( g, h \in C^2(R_+) \) satisfy that
\[ g(a) = h(a), \quad \text{for some} \quad a > 0 \]
\[ Lg(x) \geq Lh(x), \quad \text{for} \quad x > a \]
\[ |g(x)|, |h(x)| \leq c(1 + |x|), \quad \text{for} \quad x \geq a \]

Then
\[ g(x) \leq h(x), \quad \text{for} \quad x \geq a. \]
Proof. Consider \( x > a \). Let \( \{ y_t \}_{t \geq 0} \) be the diffusion process given by \( d y_t = \mu y_t dt + \sigma y_t dB_t \), \( y_0 = x \), and let \( \tau_a \) be the first time \( y_t = a \). By Lemma 1, we have
\[
\begin{align*}
    h(x) &= h(a) - \int_0^{\tau_a} e^{-r(s-t)} (Lh)(y_s) ds, \\
    g(x) &= g(a) - \int_0^{\tau_a} e^{-r(s-t)} (Lg)(y_s) ds.
\end{align*}
\]
Therefore, \( h(x) - g(x) = -\int_0^{\tau_a} e^{-r(s-t)} (Lh - Lg)(y_s) ds \geq 0 \).

The following lemma is the maximum principle on finite intervals. Essentially, it means that for a function \( g \) with \( Lg \geq 0 \), an interior positive maximum cannot be achieved.

Lemma 5. [Maximum Principle] Let \( L = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{d}{dx} - r \), where \( \sigma, \mu \) and \( r \) are constants and \( r > 0 \). \( g \in C^2([a, b]) \). If \( Lg(x) \geq 0 \), for \( a < x < b \), and \( g(a), g(b) \leq 0 \), then \( g(x) \leq 0 \) for \( a \leq x \leq b \).

Proof. Assume that to the contrary, \( g(x) \) achieves a positive maximum at \( x_0 \in (a, b) \), then \( g''(x_0) \leq 0, g'(x_0) = 0, g(x_0) > 0 \). This implies \( Lg(x_0) = \frac{1}{2}\sigma^2 x_0^2 g''(x_0) + \mu x g'(x_0) - rg(x_0) < 0 \), a contradiction.

The following lemma is a version of the maximum principle for functions \( g \) with \( Lg \geq 0 \) on a semi-finite interval \([a, \infty)\). Note that we only require that the function value of \( g \) at the finite end of the interval to be negative (or nonpositive) and that \( g \) has linear growth.

Lemma 6. [Maximum Principle at Infinity] Let \( L = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{d}{dx} - r \), where \( \sigma > 0, \mu < r \) are constants. Assume \( g \in C^2(R+) \) satisfies \( g(x) \leq c(1+|x|) \) for some constant \( c > 0 \).

i) If \( g(a) \leq 0 \), and \( Lg(x) \geq 0 \) for \( x > a \), then \( g(x) \leq 0 \) for \( x \geq a \).

ii) If \( g(a) < 0 \), and \( Lg(x) \geq 0 \) for \( x > a \), then \( g(x) < 0 \) for \( x \geq a \).

Proof. Define \( h(x) = g(a) \frac{x-a}{a} \) for \( x \geq a \). Then \( h(a) = g(a) \) and \( Lh(x) = 0 \leq Lg(x) \), for \( x \geq a \).

Since \( |g(x)|, |h(x)| \leq c(1+|x|) \), applying Lemma 4 to \( g, h \), we obtain \( g(x) \leq h(x) \) for all \( x \geq a \). If \( g(a) = 0 \), then \( h(x) \equiv 0 \) and thus \( g(x) \leq 0 \) for \( x \geq a \). If \( g(a) < 0 \), then \( h(x) < 0 \) for \( x \geq a \), and thus \( g(x) < 0 \) for \( x \geq a \).

The following lemma shows that a \( C^1 \) function \( g \) that satisfies conditions analogous to \( Lg \geq 0 \) can only cross zero "once".

Lemma 7. [Single Crossing Property] Let \( \bar{L}_i = \frac{1}{2}\sigma_i^2 x^2 \frac{d^2}{dx^2} + \bar{\mu}_i x \frac{d}{dx} - r, \ i = 1, 2 \), where \( \bar{\sigma}_1, \bar{\sigma}_2 > 0, \bar{\mu}_1, \bar{\mu}_2 < r \) are constants.

i) Let \( g \in C^2([a, b]) \cap C^2([b, \infty)) \) and \( g \) is \( C^1 \) on \([a, \infty), 0 < a < b \), and \( |g(x)| \leq c(1+x) \) for \( x \geq a \). Assume that
\[
\bar{L}_1 g(x) \geq 0, \ a \leq x < b, \quad \bar{L}_1 g(b-) > 0, \quad \bar{L}_2 g(x) \geq 0, \ x > b.
\]
If \( g(b) = 0 \), then

\[
\begin{align*}
g(x) &\geq 0, \quad a \leq x < b, \\
g(x) &\leq 0, \quad x > b.
\end{align*}
\]

ii) Let \( g \in C^2([a, b]) \cap C^2([b, d]) \) and \( g \) is \( C^1 \) on \([a, d], 0 < a < b < d\). Assume that

\[
\begin{align*}
\tilde{L}_1g(x) &\geq 0, \quad a \leq x < b, \quad \tilde{L}_1g(b-) > 0, \\
\tilde{L}_2g(x) &\geq 0, \quad b < x \leq d.
\end{align*}
\]

If \( g(b) = 0 \) and \( g(d) \leq 0 \), then

\[
\begin{align*}
g(x) &\geq 0, \quad a \leq x < b, \\
g(x) &\leq 0, \quad b < x \leq d.
\end{align*}
\]

**Proof.** i) By Lemma 6 i), \( g(x) \leq 0 \) for \( x \geq b \) and thus \( g'(b+) \leq 0 \). Now assume \( g(x_0) < 0 \) for some \( a \leq x_0 < b \). Since \( g(b) = 0 \) and \( \tilde{L}_1g \geq 0 \) on \([a, b] \), by Lemma 5, \( g(x) \leq 0 \) for \( x_0 \leq x \leq b \). This implies that the function \( g \) restricted to \([a, b] \) achieves maximum at \( b \). Thus \( g'(b-) \geq 0 \) and \( g''(b-) \leq 0 \). If \( g'(b-) = 0 \), then \( \tilde{L}_1g(b-) = \frac{1}{2}\sigma^2 x_0^2 g''(b-) + \mu x g'(b-) - rg(b-) \leq 0 \), a contradiction with the assumption \( \tilde{L}_1g(b-) > 0 \). If \( g'(b-) > 0 \), then \( g'(b-) > g'(b+) \), contradicting with the fact that \( g \) is \( C^1 \) on \([a, \infty) \). Therefore, \( g(x) \geq 0 \) for \( a \leq x < b \).

ii) By Lemma 5, we have \( g(x) \leq 0 \) for \( x \in (b, d] \) and thus \( g'(b+) \leq 0 \). The rest of the proof proceeds the same as in i). \( \square \)

The following lemma is useful to prove the optimality of the switching policy described in the statement of Theorem 1.

**Lemma 8.** Assume that \( f \in C^2(R_+) \) is strictly increasing and linearly bounded at infinity. Assume that the risk-adjusted drifts satisfy

\[
\tilde{\mu}_1(x) > \tilde{\mu}_2(x) \quad \text{for all } x > 0.
\]

Let \( S_q \) denote the value function of the switching policy that selects project 1 if \( C_t \geq q \) and project 2 if \( C_t < q \).

i) If \( (L_2S_q(x) + f(x))|_{q+} \leq 0 \), then \( (L_2S_q)(x) + f(x) \leq 0 \) for \( x \geq q \).

ii) If \( (L_2S_q)(x) + f(x)|_{q+} = 0 \), then \( (L_1S_q)(x) + f(x) \leq 0 \) for \( q > x \geq C_B \).

**Proof.** i) First, we note that the operators \( L_1 \) and \( L_2 \) commute, \( L_1L_2 = L_2L_1 \), because \( x^2 \frac{d^2}{dx^2}(x \frac{d}{dx}) = \)
\[ x^3 \frac{d^3}{dx^3} + 2x^2 \frac{d^2}{dx^2} = x \frac{d}{dx} \left( x^2 \frac{d}{dx} \right). \] For \( x > q \), we have \( L_1 S_q(x) + f(x) = 0 \). Therefore,

\[ L_1(L_2 S_q + f) = L_2(L_1 S_q) + L_1 f = (L_1 - L_2)f \]
\[ = \frac{1}{2}(\sigma_1^2 - \sigma_2^2)x^2 f'' + (\mu_1 - \mu_2)x f' = f'(\bar{\mu}_1(x) - \bar{\mu}_2(x)) > 0, \quad x > q \quad (A.13) \]

where we use the facts that \( (\sigma_1^2 - \sigma_2^2)f'' \geq 0 \) and \( (\mu_1 - \mu_2)f' > 0 \). Because \( S_q \) has the integration representation (6) and \( f \) is linearly bounded at infinity, \(|f|, |S_q(x)|, |L_1 S_q(x)| \leq c(1 + x)\) for \( x \) large. By Lemma 2,

\[ |L_2 S_q(x) + f(x)| \leq c(1 + x) \quad (A.14) \]

for some constant \( c \). Using (A.13) and (A.14) and \( L_2 S_q + f|_{q+} \leq 0 \), we can apply Lemma 6 to \( g = L_2 S_q + f \) and \( a = q \) and obtain \( (L_2 S_q)(x) + f(x) \leq 0 \) for \( x > q \).

ii) It is easy to see that \( (L_2 S_q)(x) + f(x)|_{q+} = 0 \) implies that \( S_q(x) \) is twice differentiable at \( x = q \). Note that for \( C_B \leq x < q, L_2 S_q(x) + f(x) = 0 \). Therefore,

\[ L_2(L_1 S_q + f) = L_1(L_2 S_q) + L_2 f = (L_2 - L_1)f = \frac{1}{2}(\sigma_2^2 - \sigma_1^2)x^2 f'' + (\mu_2 - \mu_1)x f' < 0, \quad C_B \leq x < q. \quad (A.15) \]

Define the function \( w \) as follows:

\[ w(x) = \begin{cases} -\sigma_1^2(L_1 S_q + f), & \text{if } q > x > C_B, \\ \sigma_2^2(L_2 S_q + f), & \text{if } x \geq q. \end{cases} \]

From (A.13) and (A.15),

\[ L_2 w(x) > 0, \quad \text{if } q > x > C_B \quad (A.16) \]
\[ L_2 w|_{q-} > 0, \quad (A.17) \]
\[ L_1 w(x) > 0, \quad \text{if } x \geq q. \quad (A.18) \]

Furthermore,

\[ w(x) = \sigma_2^2(L_2 S_q + f) = (\sigma_2^2 L_2 S_q - \sigma_1^2 L_1 S_q) + (\sigma_2^2 - \sigma_1^2)f \]
\[ = (\sigma_2^2 \mu_2 - \sigma_1^2 \mu_1) S'_q + (\sigma_2^2 - \sigma_1^2)(-\bar{r} S_q + f), \quad x \geq q \]
\[ w(x) = -\sigma_1^2(L_1 S_q + f) = (\sigma_2^2 L_2 S_q - \sigma_1^2 L_1 S_q) + (\sigma_2^2 - \sigma_1^2)f \]
\[ = (\sigma_2^2 \mu_2 - \sigma_1^2 \mu_1) S'_q + (\sigma_2^2 - \sigma_1^2)(-\bar{r} S_q + f), \quad q > x > C_B \]

Therefore, from the continuous twice differentiability of \( S_q, w(x) \) is \( C^1 \) on \((C_B, \infty)\), and \(|w(x)| \leq c(1 + x)\) (from (A.14)). \( L_2 w(q-) = -\sigma_1^2(L_2 - L_1)f(q-) > 0 \). By Lemma 7 i), \( w(x) \geq 0 \) for \( x \leq q \). Therefore, \( L_1 S_q + f = -\frac{1}{\sigma_1} w(x) \leq 0 \) for \( q > x > C_B \).
Below we proceed with the proof of Theorem 1.

**Proof of Theorem 1.** We first prove that the optimal risk-taking policy involves only one switch in both cases i) and ii). For this purpose we only need to consider the case i). The proof of the other case is identical and will be omitted. First, we consider the case where the strict inequality holds for the risk-adjusted drifts, i.e., where \( \mu_1(x) > \mu_2(x) \) for all \( x > 0 \).

**Case 1.** Let \( S_0 \) be the utility function for the policy of always choosing project 1. Assume \( L_2S_0(x) + f(x)|_{C_B^+} \leq 0 \). Then according to Lemma 8, \( L_2S_0(x) \leq 0 \) for all \( x \geq C_B \). The verification theorem then implies that \( S_0 \) is optimal.

**Case 2.** Assume that \( L_2S_0(x) + f(x)|_{C_B^+} > 0 \) and \( L_2S_q(x) + f(x)|_{q^+} \leq 0 \) for some \( q > 0 \). Then there exists \( q^* \in (C_B, q) \) such that \( L_2S_{q^*}(x) + f(x)|_{q^+} = 0 \). This implies that \( S_{q^*} \) is twice differentiable at \( q^* \) and \( L_1S_{q^*}(x) + f(x)|_{q^+} = 0 \). Lemma 3 then implies that \( L_1S_{q^*}(x) + f(x) \leq 0 \) for \( x < q \) and \( L_1S_{q^*} + f(x) \leq 0 \) for \( x > q \). The verification theorem then implies that \( S_{q^*} \) is optimal.

**Case 3.** If \( L_2S_q(x) + f(x)|_{q^+} > 0 \) for all \( q > 0 \), then \( L_1S_q(x) + f(x)|_{q^+} = -\frac{\sigma_2^2}{\sigma_1^2}(L_2S_q(x) + f(x)) \) for all \( q > 0 \). Let \( S_\infty \) be the the utility function for strategy of always choosing the control value of 2. Then \( L_1S_\infty(x) + f(x) < 0 \) for all \( x > 0 \) and \( S_\infty \) is thus optimal.

Next, we deal with the case where the risk-adjusted drifts satisfy \( \tilde{\mu}_1(x) \geq \tilde{\mu}_2(x) \), for all \( x > 0 \). We will apply a limiting argument to the results in the case of strict inequality that we show above.

Consider a strictly increasing sequence \( \{\mu_{2,n}\}_{n\geq 1} \) such that \( \lim_{n \to \infty} \mu_{2,n} = \mu_2 \). For the projects \((\mu_1, \sigma_1)\) and \((\mu_{2,n}, \sigma_2)\), the risk-adjusted drifts satisfy the strict inequality \( \tilde{\mu}_1(x) > \tilde{\mu}_{2,n}(x) \), for all \( x > 0 \). Therefore, as we have shown above, there exist switching triggers \( q_n \), which gives the optimal policy for the agent selecting between the two projects \((\mu_1, \sigma_1)\) and \((\mu_{2,n}, \sigma_2)\). Let \( S_{q_n} \) be the corresponding value function. There exists a subsequence \( n_k \) such that \( q_{n_k} \) converges as \( k \to \infty \) (the limit may be infinite). Without loss of generality, assume that \( q_n \to q^* \) as \( n \to \infty \). We shall prove that the strategy \( S_{q^*} \) is optimal.

By the integration representation (6), the value functions \( S_{q_n} \) converge to the value function \( S_{q^*} \). Since \( L_1S_{q_n} = -f \) for \( x > q_n \), and \( L_2S_{q_n} = -f \) for \( C_B < x < q_n \), where \( (L_{2,n}g)(x) = \frac{1}{2}\sigma_2^2 \frac{d^2g}{dx^2}(x) + \frac{\mu_2}{2\sigma_1^2} \frac{dg}{dx}(x) - rg(x) \), it follows from Lemma 2 that the convergence of \( S_{q_n} \) to \( S_{q^*} \) is \( C^2 \) on any compact intervals bounded away from \( q^* \). Therefore, the optimality conditions \( L_2nS_{q_n} + f \leq 0 \) for \( x > q_n \) and imply that \( L_2S_q + f \leq 0 \) for \( x > q^* \). Similarly, the optimality conditions \( L_1S_{q_n} + f \leq 0 \) for \( C_B < x < q_n \) imply \( L_1S_q + f \leq 0 \) for \( C_B < x < q^* \). Thus the strategy \( S_{q^*} \) is indeed optimal.

Next, we shall prove that in case i), the optimal switching point \( q^* > C_B \) and that in case ii), \( q^* = C_B \). We first consider case i). From Case 1 in the proof above, we see that the strategy of always choosing project 1 is optimal if and only

\[
L_2S_0(x) + f(x)|_{C_B^+} \leq 0. \tag{A.19}
\]
From the verification theorem (Proposition 4), \( S_0^*(C_B) = S_0(C_B) = 0 \). Therefore,

\[
0 = L_1 S_0(x) + f(x)|_{C_B^+} = \frac{1}{2} \sigma_1^2 S_0''(C_B) + f(C_B) \quad \text{(A.20)}
\]

It is clear that \( f(C_B) < 0 \) because otherwise the agent’s utility value can be increased by lowering the bankruptcy threshold slightly. By (A.20), \( S_0''(C_B) > 0 \). This implies

\[
L_2 S_0(x) + f(x)|_{C_B^+} = \frac{1}{2} \sigma_2^2 S_0''(C_B) + f(C_B) > \frac{1}{2} \sigma_1^2 S_0''(C_B) + f(C_B) = 0.
\]

Therefore, the strategy of choosing only project 1 cannot be optimal and \( q^* > C_B \).

In case ii), let \( S_0^b(x) \) be the strategy of always choosing project 2. By the same argument as above, we can show that

\[
L_1 S_0^b(x) + f(x)|_{C_B^+} < 0.
\]

By Lemma 8, \( L_1 S_0^b(x) + f(x) \leq 0 \) for \( x \geq C_B^+ \) and hence \( S_0^b \) is optimal, i.e., \( q^* = C_B \).

**Proof of Proposition 1.** Consider the strategy \( \infty \) where project 2 is always adopted. Then from the equation \( L_2 S_\infty + f = 0 \), we obtain

\[
S_\infty(x) = -\frac{c_1}{r} + \frac{c_2}{1 - \gamma} \left( \frac{x^{1-\gamma}}{(1 - \gamma - \gamma_1)(1 - \gamma - \gamma_2^+)} \right) + a_1 x^{\gamma_2^-} + o(x^{1-\gamma}).
\]

Therefore,

\[
L_1 S_\infty(x) + f(x) = -\frac{c_2}{1 - \gamma} \left[ (1 - \gamma - \gamma_1)(1 - \gamma - \gamma_2^+) - 1 \right] x^{1-\gamma} + a_1(\gamma_2^- - \gamma_1^-)(\gamma_2^- - \gamma_1^+) x^{\gamma_2^-} + o(x^{1-\gamma}).
\]

\[
= -\frac{c_2}{1 - \gamma} \left[ \frac{1}{2}(\sigma_1^2 - \sigma_2^2)(1 - \gamma)(1 - \gamma - \gamma_2^+) + (\mu_1 - \mu_2)(1 - \gamma) \right] x^{1-\gamma} + a_1(\gamma_2^- - \gamma_1^-)(\gamma_2^- - \gamma_1^+) x^{\gamma_2^-} + o(x^{1-\gamma}).
\]

\[
= -\left[ \frac{\mu_1 - \frac{1}{2}\sigma_1^2 \gamma - (\mu_2 - \frac{1}{2}\sigma_2^2 \gamma)}{(1 - \gamma - \gamma_2^-)(1 - \gamma - \gamma_2^+)} \right] x^{1-\gamma} + a_1(\gamma_2^- - \gamma_1^-)(\gamma_2^- - \gamma_1^+) x^{\gamma_2^-} + o(x^{1-\gamma}) > 0, \quad x \to \infty.
\]

In the last step above, we used the fact that \( 1 - \gamma - \gamma_2^- \geq -\gamma_2^- > 0, 1 - \gamma - \gamma_2^+ \leq 1 - \gamma_2^+ < 0, \mu_1 - \frac{1}{2}\sigma_1^2 \gamma - (\mu_2 - \frac{1}{2}\sigma_2^2 \gamma) > 0 \). Therefore, the strategy \( q \) that adopts project 1 when \( x \) is sufficiently large will generate a higher utility value. This implies that the optimal switching point \( q^* < \infty \).

**Proof of Corollary 1.** The general case where \( f \) is a continuous function can be obtained as the limiting case of smooth functions. Assume \( f \) is concave (the convex case is similar). Construct a decreasing sequence of smooth concave functions \( f_1 > f_2 > \ldots \) such that \( f_n \) converges uniformly
to $f$ on any compact intervals. By the theorem, there exists unique triggers $q_n^*$ such that the value function $S_{q_n^*}^n$ is the optimal value function with payoff function $f_n$. There exists a subsequence $n_i$ such that $q_{n_i}^* \to q^*$, $i \to \infty$. For simplicity, assume $q_n^* \to q^*$ as $n \to \infty$. We shall prove that the strategy $S_{q^*}$ is optimal.

By the integration representation (6) and the convergence of $f_n$, the value functions $S_{q_n^*}^n$ converge to the value function $S_{q^*}$. Since $L_1S_{q_n^*}^n = -f_n$ for $x > q_n^*$, and $L_2S_{q_n^*}^n = -f_n$ for $C_B < x < q_n^*$, it follows from Lemma 2 that the convergence of $S_{q_n^*}^n$ to $S_{q^*}$ is $C^2$ on any compact intervals bounded away from $q^*$. Therefore, the optimality conditions $L_2S_{q_n^*}^n + f_n \leq 0$ for $x > q_n^*$ and imply that $L_2S_q + f \leq 0$ for $x > q^*$. Similarly, the optimality conditions $L_1S_{q_n^*}^n + f_n \leq 0$ for $C_B < x < q_n^*$ imply $L_1S_{q^*} + f \leq 0$ for $C_B < x < q^*$. Thus the strategy $S_{q^*}$ is indeed optimal.

In the proof of Theorem 2, we need the following two lemmas. The following lemma states that if a project is dominated by another on an interval, there cannot be a sub-interval bounded by switching points on which the dominated project is selected.

**Lemma 9.** i) If $\bar{\mu}_1(x) - \bar{\mu}_2(x) > 0$ on $x \in (a, b) \subset [C_B, \infty)$, then there do not exist two adjacent switching points $a \leq q_1 < q_2 \leq b$ such that project 2 is selected for $x \in (q_1, q_2)$ in the optimal policy.

ii) If $\bar{\mu}_1(x) - \bar{\mu}_2(x) < 0$ on $x \in (a, b) \subset [C_B, \infty)$, then there do not exist two adjacent switching points $a \leq q_1 < q_2 \leq b$ such that project 1 is selected for $x \in (q_1, q_2)$ in the optimal policy.

**Proof.** We only need to prove part i) since part ii) then follows by symmetry. Assume to the contrary that there exist switching points $a \leq q_1 < q_2 \leq b$ such that in the optimal policy the agent selects project 2 when $x \in (q_1, q_2)$. This implies that $L_2S^+ = 0$ for $x \in (q_1, q_2)$, $L_1S^+|_{q_1}^b = 0$, and $L_1S^+|_{q_2}^b = 0$. Using (A.15), we have $L_2(L_1S^+ + f) < 0$ for $x \in (q_1, q_2)$. Therefore, by the maximum principle (Lemma 5), $L_1S^+ + f \geq 0$ for $x \in (q_1, q_2)$ and strict inequality holds for some $y \in (q_1, q_2)$ (otherwise $L_2(L_1S^+ + f) \equiv 0$ on $(q_1, q_2)$). This is in contradiction with the verification theorem, which requires $L_1S^+ + f \leq 0$ for $x \in (q_1, q_2)$.

The following lemma provides the bounds on the number of switching points when the project choice at the right end of the interval is given.

**Lemma 10.** Assume that $\{x_i\}_{i=0,\ldots,L}$ satisfies $C_B = x_0 < x_1 < \ldots < x_L \leq \infty$ and $\bar{\mu}_1(x) - \bar{\mu}_2(x)$ does not change sign on $x \in (x_{i-1}, x_i)$ for all $1 \leq i \leq L$ and changes signs at each point $x_i$, for $1 \leq i \leq L$. Also assume that project $p_L \in \{1, 2\}$ is chosen for $x \in (x_L - \varepsilon, x_L]$ for some $\varepsilon > 0$. Then

i) If $\bar{\mu}_1(x) - \bar{\mu}_2(x) > 0$ for $x \in (x_{L-1}, x_L)$ and $p_L = 1$, or $\bar{\mu}_1(x) - \bar{\mu}_2(x) < 0$ for $x \in (x_{L-1}, x_L)$ and $p_L = 2$, then the number of switching points in the interval $[x_0, x_L]$ is bounded above by $L$.

ii) If $\bar{\mu}_1(x) - \bar{\mu}_2(x) < 0$ for $x \in (x_{L-1}, x_L)$ and $p_L = 1$, or $\bar{\mu}_1(x) - \bar{\mu}_2(x) > 0$ for $x \in (x_{L-1}, x_L)$ and $p_L = 2$, then the number of switching points in the interval $[x_0, x_L]$ is bounded above by $L+1$. 45
Proof. We prove the lemma by induction. Assume that the lemma holds for the case of $L = L' - 1$. First, we consider case i) for $L = L'$. By symmetry, we only need to consider the case \( \hat{\mu}_1(x) - \hat{\mu}_2(x) > 0 \) for $x \in (x_{L'-1}, x_{L'})$ and $p_{L'} = 1$. There cannot be more than one switching points in \([x_{L'-1}, x_{L'})\) because otherwise the region bounded by the two greatest switching points would give rise to the sub-interval described in Lemma 9 part i). If there are no switching points in \([x_{L'-1}, x_{L'})\), then the agent still selects project 1 at $x = x_{L'-1}$, i.e., $p_{L'-1} = p_{L'} = 1$. Case ii) for $L = L' - 1$ implies that there are at most $L'$ switching points on $[x_0, x_{L'})$. If there is one switching point in \([x_{L'-1}, x_{L'})\), then $p_{L'-1} = 2$ and case i) for $L = L' - 1$ implies that there at most $L' - 1 + 1 = L'$ switching points on $[x_0, x_{L'})$.

Next, we consider case ii) for $L = L'$. By symmetry, we only need to consider the case \( \hat{\mu}_1(x) - \hat{\mu}_2(x) < 0 \) for $x \in (x_{L'-1}, x_{L'})$ and $p_{L'} = 1$. Lemma 9 part ii) implies that there can be no more than two switching points in \([x_{L'-1}, x_{L'})\). Using the results for $L = L' - 1$ it is then easy to verify that there are at most $L' + 1$ switching points on $[x_0, x_{L'})$.

Proof of Theorem 2. Without loss of generality, we assume that $\hat{\mu}_1(x) > \hat{\mu}_2(x)$ for $x_{M-1} < x < x_M = \infty$. The proof of the main theorem shows that when $x$ is sufficiently large, the agent selects project 1 in the optimal policy. Therefore, by Lemma 10 Case i), the total number of switching points of the optimal policy on $[C_B, \infty)$ is bounded above by $M$.

Appendix B. Proofs in the Case of Multiple Projects

We first establish the existence of a unique spanning subset of projects.

Proof of Lemma 1. We use induction to prove the existence of a spanning subset. When the number of projects $N = 1, 2$, the spanning subset is the entire set of projects. Assume that spanning subset exists for any $N - 1$ projects. We now consider the case of $N$ projects. We first select the spanning subset \( \{j_1, j_2, \ldots, j_L\} \) of projects \( \{1, 2, \ldots, N - 1\} \). Let $l$ be the index that minimizes the slope from \((\mu_{j_l}, \sigma^2_{j_l})\) to \((\mu_N, \sigma^2_N)\),

\[
 l = \arg \min_{1 \leq l \leq L} \frac{\mu_{j_l} - \mu_N}{\sigma^2_{j_l} - \sigma^2_N}. \tag{B.1}
\]

If there are multiple minimum points for (B.1), then we choose $l$ to be the smallest such index. For any project $j_k$ with $L \geq k > l$, (B.1) implies that

\[
 \frac{\mu_{j_l} - \mu_N}{\sigma^2_{j_l} - \sigma^2_N} \leq \frac{\mu_{j_k} - \mu_N}{\sigma^2_{j_k} - \sigma^2_N}
\]

which is equivalent to the condition that $j_k$ is dominated by $j_l$ and $N$ (as defined in (14)). It is then easy to verify that \( \{j_1, j_2, \ldots, j_l, N\} \) form a spanning subset of all the projects \( \{1, \ldots, N\} \).

To show uniqueness, consider any spanning subset \( \{i_1, \ldots, i_K\} \). We know that $i_K = N$ and we will show that
\[ i_{K-1} = j^* = \min \left( \arg \min_{1 \leq j \leq N} \frac{\mu_j - \mu_N}{\sigma_j^2 - \sigma_N^2} \right). \quad (B.2) \]

If this is true, then by induction we can easily show that the spanning subset must be unique. Now if \( i_{K-1} \neq j^* \), there are three cases.

Case i): \( i_{K-1} < j^* \). Then \( j^* \) is dominated by \( i_{K-1} \) and \( i_K = N \), by property (ii) of spanning subsets. But this implies the slope from \((\mu_{i_{K-1}}, \sigma_{i_{K-1}}^2)\) to \((\mu_N, \sigma_N^2)\) is smaller than or equal to that from \((\mu_{j^*}, \sigma_{j^*}^2)\) to \((\mu_N, \sigma_N^2)\), a contradiction with (B.2).

Case ii): \( i_{K-1} > j^* = i_1 = 1 \). Then \( i_{K-1} \) is dominated by projects \( i_1 = j^* \) and \( i_K = N \), a contradiction to the properties of spanning subsets.

Case iii): \( i_{K-1} > j^* > i_1 \). Then there exists \( s \) such that \( i_s > j^* > i_{s-1} \) (if \( j^* \) is equal to some \( i_s \), it is easy to show a contradiction). Therefore \( j^* \) is dominated by \( i_{s-1} \) and \( i_s \). However, this means that \( j^* \) cannot satisfy the minimum slope condition in (B.2).

The following lemma is an implication of the strict concavity condition of project drifts and variances and will be useful in proving Proposition 2.

**Lemma 11.** Assume \( 1 \leq i < j < k \leq N \), \( S \in \mathbb{C}^2 \) is the agent’s utility function for a control policy, and the drifts and variances of projects \( \{i, j, k\} \) satisfy the strict concavity condition

\[ \mu_j > \frac{(\sigma_k^2 - \sigma_i^2)\mu_i + (\sigma_j^2 - \sigma_i^2)\mu_k}{\sigma_k^2 - \sigma_i^2}, \quad (B.3) \]

then the following is true:

i) There exists \( 0 < \alpha < 1 \) and \( \delta > 0 \) such that

\[ L_jS(x) + f(x) = \alpha(L_iS(x) + f(x)) + (1 - \alpha)(L_kS(x) + f(x)) + \delta xS'(x). \quad (B.4) \]

ii) If \( L_iS(x) + f(x) = 0 \) and \( L_jS(x) + f(x) \leq 0 \) for some \( x \), then \( L_kS(x) + f(x) < 0 \).

iii) If \( L_jS(x) + f(x) \leq 0 \) and \( L_kS(x) + f(x) = 0 \) for some \( x \), then \( L_iS(x) + f(x) < 0 \).

iv) If \( L_iS(x) + f(x) = 0 \) and \( L_kS(x) + f(x) = 0 \) for some \( x \), then \( L_jS(x) + f(x) > 0 \).

**Proof.** Let \( \alpha = \frac{\sigma_k^2 - \sigma_i^2}{\sigma_k^2 - \sigma_i^2} \). Then \( \sigma_j^2 = \alpha\sigma_i^2 + (1 - \alpha)\sigma_k^2 \). Because \( \sigma_i^2 < \sigma_j^2 < \sigma_k^2 \), we have \( 0 < \alpha < 1 \). The concavity condition (B.3) implies that \( \mu_j > \alpha \mu_i + (1 - \alpha)\mu_k \). Therefore,

\[ L_jS(x) + f(x) = \frac{1}{2} \sigma_j^2 x^2S''(x) + \mu_jx^2S'(x) - rS(x) + f(x) \]

\[ = \frac{1}{2} \left( \alpha \sigma_i^2 + (1 - \alpha)\sigma_k^2 \right) x^2S''(x) \]

\[ + [\alpha \mu_i + (1 - \alpha)\mu_k + \mu_j - (\alpha \mu_i + (1 - \alpha)\mu_k)]xS'(x) - rS(x) + f(x) \]

\[ = \alpha(L_iS(x) + f(x)) + (1 - \alpha)(L_kS(x) + f(x)) + \mu_j - (\alpha \mu_i + (1 - \alpha)\mu_k)]xS'(x) \]

\[ > \alpha(L_iS(x) + f(x)) + (1 - \alpha)(L_kS(x) + f(x)). \quad (B.5) \]
where we have used the fact that $S'(x) > 0$ (because $f'(x) > 0$). It is straightforward to see that i) to iv) follow from (B.5).

**Proof of Proposition 2. Part i):** Consider the optimal policy $P^*$ in the constrained problem where the agent only select projects from the spanning subset of projects $\{i_k\}_{k=1,\ldots,K}$ (the existence of such an optimal policy will be shown in Theorem 3). We shall prove that $P^*$ is also optimal among all policies in which the agent can select from the full set of projects. Let $S_{P^*}$ be the value function that corresponds to the strategy $P^*$. By the verification theorem (Proposition 4),

$$L_{i_k}S_{P^*}(x) + f(x) \leq 0, \quad x \geq C_B, 1 \leq k \leq K. \quad \text{(B.6)}$$

Assume $j \notin I$ is not in the spanning subset, then there exists $k$ such that $i_k < j < i_{k+1}$. Let $\alpha = \frac{\sigma_{i_{k+1}}^2 - \sigma_{i_k}^2}{\sigma_{i_{k+1}}^2 - \sigma_{i_k}^2}$, then $0 < \alpha < 1$. Condition (14) then implies that

$$L_jS_{P^*}(x) + f(x) \leq \alpha(L_{i_k}S_{P^*}(x) + f(x)) + (1 - \alpha)(L_{i_{k+1}}S_{P^*}(x) + f(x)) \leq 0, \quad x \geq C_B \quad \text{(B.7)}$$

By the verification theorem (Proposition 4), (B.6) and (B.7), the control policy given by $P^*$ is globally optimal.

**Part ii):** Again, let $P^*$ be the optimal switching policy. Assume that under this policy the agent switches from project $i_j$ to project $i_k$ as cash flow crosses a switching threshold $q$ and $k > j + 1$. The fact that $P^*$ is optimal among all control policies implies that

$$L_{i_k}S_{P^*} + f|_{q^+} = 0. \quad \text{(B.8)}$$

Since $L_{i_j}S_{P^*} + f = 0$ for $x \in (q, q + \varepsilon)$,

$$L_{i_j}S_{P^*} + f|_{q^+} = 0. \quad \text{(B.9)}$$

(B.8), (B.9), and Lemma 11 (iv) imply that

$$L_{i_{j+1}}S_{P^*} + f|_{q^+} > 0. \quad \text{(B.10)}$$

This is contradictory to the optimality of $P^*$ and the verification theorem.

We next set up some notations in the proof of Theorem 3. Consider project choice policies that are characterized by a decreasing sequence of $K - 1$ switching triggers $\infty \geq q_1 \geq q_2 \geq \ldots \geq q_{K-1} \geq C_B$. Note that we allow the possibility that $q_i = \infty$. For ease of notation, we let $q_0 = \infty$ and $q_K = C_B$. Define a policy $P_Q$ as follows:

$$P_Q(C_t) = i_k, \quad \text{if } q_k < C_t \leq q_{k-1}, 1 \leq k \leq K. \quad \text{(B.11)}$$
Under the policy \( PQ \), the agent selects only projects in the spanning subset. Let \( S_Q \) be the agent’s utility function associated with the strategy \( PQ \). For ease of notation, we define \( q_0 = \infty \), and \( q_K = C_B \). (5), (6), and Ito’s lemma imply that

\[
L_{ik} S_Q(x) + f(x) = 0, \quad q_k < x < q_{k-1}, 1 \leq k \leq K.
\]

The following lemma provides a set of “local” equations that imply global optimality of such a switching policy, provided that the spanning projects have consistently ordered risk-adjusted drifts.

**Lemma 12.** [Global Optimality of Switching Strategies] Assume that \( f \in C^2(\mathbb{R}_+) \) has linear growth and the risk-adjusted drifts of the spanning projects are consistently ordered on \( \mathbb{R}_+ \) and \( i_h \) is the project with the highest risk-adjusted drift. If there exists an \((K-1)\)-tuple \( Q = (q_1, \ldots, q_{K-1}) \) with \( \infty = \ldots = q_{m'-1} > q_{m'} > \ldots > q_m > q_{m+1} = \ldots = C_B \) and \( m' \geq h \) such that

\[
L_{i_{m'-1}} S_Q(x_n) + f(x_n) \leq 0, \quad \text{for some } x_n \to \infty, \quad (B.12)
\]

\[
L_{ik+1} S_Q + f|_{q_k} = 0, \quad m' \leq k \leq m, \quad (B.13)
\]

\[
L_{ik} S_Q + f|_{q_k} = 0, \quad m' \leq k \leq m, \quad (B.14)
\]

\[
L_{i_{m+2}} S_Q + f|_{C_B} \leq 0. \quad (B.15)
\]

Then the strategy \( PQ \) is globally optimal among all admissible strategies.

**Proof.** The strict concavity condition \(((B.3))\) and the assumption that \( i_h \) has the highest risk-adjusted drift imply that

\[
\bar{\mu}_{ih}(x) \geq \bar{\mu}_{ih+1}(x) \geq \ldots \geq \bar{\mu}_{iK}(x), \quad x > 0. \quad (B.16)
\]

By Lemma (11), we only need to show that \( PQ \) is optimal among all strategies selecting projects in the spanning subset. First we note that \( L_{ij} \) and \( L_{ik} \) commute, i.e., \( L_{ij} L_{ik} = L_{ik} L_{ij} \), because

\[
x^2 \frac{d^2}{dx^2}(x \frac{d}{dx}) = x^3 \frac{d^3}{dx^3} + 2x^2 \frac{d^2}{dx^2} = x \frac{d}{dx} \left( x^2 \frac{d^2}{dx^2} \right). \]

Therefore, if \( m' \leq j < k \leq m + 1 \), then for \( q_j < x < q_{j-1} \), because \( L_{ij} S_Q(x) + f(x) = 0 \),

\[
L_{ij}(L_{ik} S_Q + f)(x) = L_{ik} L_{ij} S_Q(x) + L_{ij} f(x) = L_{ij} f(x) - L_{ik} f(x) \]

\[
= \frac{1}{2} (\sigma_{ij}^2 - \sigma_{ik}^2) f''(x) + (\mu_{ij} - \mu_{ik}) f'(x) \]

\[
= f'(x) \left( \bar{\mu}_{ij}(x) - \bar{\mu}_{ik}(x) \right) \geq 0. \quad (B.17)
\]
where the last inequality follows from (B.16). Similarly, for \( q_k < x < q_{k-1} \),

\[
L_{ik}(L_{ij}S_Q + f)(x) \leq 0. \tag{B.18}
\]

To show that \( PQ \) is globally optimal, by the verification theorem (Proposition 4), it suffices to show that

\[
L_{ik}S_Q(x) + f(x) \leq 0, \quad 1 \leq k \leq K, \quad x \geq C_B. \tag{B.19}
\]

**Step 1.** We will first show that

\[
L_{ik}S_Q + f \leq 0, \quad \text{for } x \in [q_{k+1}, q_{k-2}], \quad m' - 1 \leq k \leq m + 2,
\]

where we let \( q_0 = \infty \) and \( q_K = C_B \) to simplify notation.

**Case a.** \( k = m' - 1 \). Assume that \( L_{im' - 1}S_Q + f > 0 \) for some \( x_0 \in [q_{m'}, \infty) \). By (B.18), \( L_{im'}(L_{im' - 1}S_Q + f) < 0 \) for \( x \in [q_{m'}, \infty) \). By Lemma 6 (applied to \( -L_{im' - 1}S_Q + f \)), \( L_{im' - 1}S_Q + f > 0 \) for \( x \geq x_0 \). But this contradicts (B.12). Therefore, \( L_{im' - 1}S_Q + f \leq 0 \) for all \( x \in [q_{m'}, \infty) \).

**Case b.** \( m' \leq k \leq m + 1 \). Define the function \( w(x) \) as follows:

\[
w(x) = \begin{cases} 
-\sigma^2_{ik}(L_{ik}S_Q + f), & \text{if } q_{k+1} \leq x < q_k, \\
\sigma^2_{ik+1}(L_{ik+1}S_Q + f), & \text{if } q_k \leq x < q_{k-1}.
\end{cases}
\]

Similar to the proof of Lemma 8, since \( L_{ik}S_Q + f = 0 \) for \( q_k \leq x < q_{k-1} \) and \( L_{ik+1}S_Q + f = 0 \) for \( q_{k+1} \leq x < q_k \), \( w(x) \) is continuously differentiable at \( q_k \). From (B.17) and (B.18),

\[
L_{ik+1}w(x) > 0, \quad \text{if } q_{k+1} \leq x < q_k, \tag{B.21}
\]

\[
L_{ik+1}w|_{q_k} > 0, \tag{B.22}
\]

\[
L_{ik}w(x) > 0, \quad \text{if } q_k \leq x \leq q_{k-1}. \tag{B.23}
\]

Note \( w(q_k) = 0 \) by (B.14). The facts that \( L_{ik}S_Q + f|_{q_k-1} = L_{ik-1}S_Q + f|_{q_k-1} = 0 \) and Lemma 11 imply that \( L_{ik+1}S_Q + f|_{q_k-1} \leq 0 \), i.e., \( w(q_k-1) \leq 0 \). Now Lemma 7 applies to \( w \) on the interval. Therefore, \( w(x) \geq 0 \) for \( q_{k+1} \leq x < q_k \), i.e.,

\[
L_{ik}S_Q(x) + f(x) \leq 0, \quad \text{for } x \in [q_{k+1}, q_k].
\]

In addition, \( w(x) \leq 0 \) for \( q_k \leq x \leq q_{k-1} \), i.e., \( L_{ik+1}S_Q(x) + f(x) \leq 0 \), for \( x \in [q_k, q_{k-1}] \).

The same argument as above applied to \( L_{ik-1} \) and \( L_{ik} \) on the interval \( [q_k, q_{k-2}] \) implies that

\[
L_{ik}S_Q(x) + f(x) \leq 0, \quad \text{for } x \in [q_{k-1}, q_{k-2}].
\]

Since \( L_{ik}S_Q(x) + f(x) = 0 \) for \( x \in [q_k, q_{k-1}] \), we have shown that (B.20) holds for index \( k \).
Case e. $k = m + 2$. We have $L_{i_{m+2}}S_Q + f|_{C_B^+} \leq 0$ from (B.19). We also have $L_{i_m}S_Q + f|_{q_m^+} = L_{i_{m+1}}S_Q + f|_{q_m^+} = 0$, which implies $L_{i_{m+2}}S_Q + f|_{q_m^+} < 0$ from Lemma 11. Since $L_{i_{m+1}}(L_{i_{m+2}}S_Q + f) \geq 0$ for $x \in [C_B, q_m]$, the maximum principle (Lemma 5) implies that $L_{i_{m+2}}S_Q + f \leq 0$ for $x \in [C_B, q_m]$. Thus, (B.20) also holds for $k = m + 2$.

Step 2. We will now show that $L_{i_k}S_Q + f \leq 0$ for all $1 \leq k \leq K$ and on each interval $[q_j, q_{j-1}]$, $m' \leq j \leq m + 1$. These are equivalent to the global optimality conditions (B.19).

The proof follows easily from (B.20) and the concavity property in Lemma 11. Consider $x \in [q_j, q_{j-1}]$ for $m' \leq j \leq m + 1$. From (B.20), $L_{i_{j+1}}S_Q(x) + f(x) \leq 0$ and $L_{i_{j-1}}S_Q(x) + f(x) \leq 0$. Since $L_{i_j}S_Q(x) + f(x) = 0$, Lemma 11 implies $L_{i_k}S_Q(x) + f(x) < 0$, for $k < j - 1$ and $k > j + 1$, and thus $L_{i_k}S_Q(x) + f(x) \leq 0$ for all $1 \leq k \leq N$.

Next we proceed to prove Theorem 3.

**Proof of Theorem 3.** To prove the theorem, consider all decreasing $(K - 1)$-tuples $Q = (q_1, \ldots, q_{K-1}) : C_B \leq q_{K-1} \leq \ldots \leq q_1 \leq \infty$ and the corresponding value functions $S_Q$. Fix $x_0 > C_B$ and define

$$S_0^* = \sup_Q S_Q(x_0)$$

There exists a sequence $Q^n = (q^n_1, \ldots, q^n_{N-1})$, $n = 1, 2, \ldots$, such that $\lim_{n \to \infty} S_{Q^n}(x_0) = S_0^*$ and $\lim_{n \to \infty} q^n_k = q_k^*$. Now by continuity, $Q^* = (q_1^*, \ldots, q_{K-1}^*)$ is a decreasing $(K - 1)$-tuple, and $S_{Q^*}(x_0)$ achieves the optimum $S_0^*$. Lemma 11 implies that there exists $m' \leq m$ such that

$$\infty = \ldots = q_{m'-1} > q_{m'} > \ldots > q_m > q_{m+1} = \ldots = C_B.$$

We next show that $m' \geq h$. Assume to the contrary that $m' < h$. First, if $m + 1 = m'$, i.e., only project $m'$ is selected, then by the conditions $S'_{Q^*}(C_B) = S_{Q^*}(C_B) = 0$, we have

$$L_{i_{m'+1}}S_{Q^*} + f|_{C_B^+} = \frac{1}{2} \sigma_{i_{m'+1}}^2 S''_{Q^*}(C_B) + f(C_B) + \frac{1}{2} \sigma_{i_{m'+1}}^2 S''_{Q^*}(C_B) + f(C_B) = 0,$$

in contradiction with the optimality of $P_{Q^*}$ among policies given by decreasing $(K - 1)$-tuples. Therefore, $m + 1 > m'$ and

$$L_{i_{m'+1}}S_{Q^*} + f|_{q_{m'}^+} = 0. \quad (B.24)$$

Since $m' < h$, by the strict concavity condition ((B.3)) and the assumption that $i_h$ has the highest risk-adjusted drift,

$$\tilde{\mu}_{i_h}(x) < \tilde{\mu}_{i_{m'+1}}(x). \quad (B.25)$$

Hence,

$$L_{i_{m'}}(L_{i_{m'+1}}S_{Q^*} + f) = 0, \quad q_{m'} < x < \infty. \quad (B.26)$$

By (B.24) and (B.26), the maximum principle at infinity (Lemma 6) applies to the function
\[ -L_{i_{m'}+1}S_Q^* + f \] and thus
\[ L_{i_{m'}+1}S_Q^* + f \geq 0, \quad q_{m'} \leq x < \infty. \] (B.27)

The inequality in (B.27) has to hold everywhere except at \( q_{m'} \) because of the strict inequality (B.26). But then it is clear by Ito’s lemma that selecting project \( i_{m'}+1 \) for \( x > q_{m'} \) is actually better than selecting project \( i_{m'} \). This is again in contradiction with the optimality of \( P_Q^* \) among policies given by decreasing \((K-1)\)-tuples. Therefore, we have shown by contradiction that \( m' \geq h \).

It is easy to show by optimality of \( S_Q^* \) and Ito’s lemma that
\[ L_{i_{m'-1}}S_Q^*(x_n) + f(x_n) \leq 0, \text{ for some } x_n \to \infty, \]
\[ L_{i_{k+1}}S_Q^* + f|_{q_k+} = 0, \quad m' \leq k \leq m, \]
\[ L_{i_k}S_Q^* + f|_{q_k-} = 0, \quad m' \leq k \leq m, \]
\[ L_{i_{m+2}}S_Q^* + f|_{C_B+} \leq 0. \]

We can now apply Lemma 12 and it follows that the strategy \( P_Q^* \) determined by \( Q^* \) is indeed globally optimal.

Finally, by the same argument as in Theorem 1, we can show that \( m = K - 1 \), i.e., the agent always switches to the spanning project \( i_K \) with the highest volatility when the firm’s output is sufficiently low.

References


