2014

Essays on dynamic information economics

Wong, Tak-Yuen

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESSAYS ON DYNAMIC INFORMATION ECONOMICS

by

TAK-YUEN WONG

B. Economics and Finance, University of Hong Kong, 2006
M.Phil., Economics, Chinese University of Hong Kong, 2008

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2014

Approved by

First Reader

Jianjun Miao, Ph.D.
Professor of Economics

Second Reader

Juan Ortner, Ph.D.
Assistant Professor of Economics

Third Reader

Dilip Mookherjee, Ph.D.
Professor of Economics

This dissertation is dedicated to my fiancée Crystal,
to my parents Ching-Kit and Wai-Hung,
and to my brother Hong-Yuen.

Acknowledgments

I would like to express my deepest gratitude to my advisers Jianjun Miao, Juan Ortner and Dilip Mookherjee for their guidance, patience, support and encouragement. It is a wonderful experience to learn from them. I would like to thank Larry Epstein, Bart Lipman, Hsueh-Ling Huynh, Albert Ma, Michael Manove, Sambuddha Ghosh for generously sharing their knowledge in economics. It is also my pleasure to study and work with my classmates Jacopo Bizzoto, Atanu Bandyopadhyay, Seong Yeon Chang, Hyungseok Joo, Sudipto Karmakar, Alejandro Rivera, Eyno Rots, Kaz Sakamoto, Julie Shi and Qiang Wang. I learn a lot of economics, finance, social issues, politics, mathematics and physical science from them. Lastly, I thank Julian Tsz-Kin Chan for his invaluable help and comments.

ESSAYS ON DYNAMIC INFORMATION ECONOMICS

(Order No.)

TAK-YUEN WONG

Boston University, Graduate School of Arts and Sciences, 2014

Major Professor: Jianjun Miao, Professor of Economics

ABSTRACT

This dissertation studies moral hazard problems and an information acquisition problem in dynamic economic environments. In chapter 1, I study a continuous-time principal-agent model in which a risk-neutral agent protected by limited liability exerts costly efforts to manage a project for her principal. Unobserved risk-taking by the agent is value-reducing in the sense that it increases the chance of large losses, even though it raises short-term profits. In the optimal contract, severe punishment that follows a large loss prevents the agent from taking hidden risks. However, after some histories, punishment can no longer be used because of limited liability. The principal allows the agent to take hidden risk when the firm is close to liquidation. In addition, I explore the roles of standard securities in implementing the optimal contract. The implementation shows that driven by the agency conflicts, incomplete hedging against Poisson risk provides incentives for the agent to take the safe project. Moreover, I study the optimality of "high-water mark" contract widely used in the hedge fund industry and find that "distance-to-threshold" is important in understanding the risk-shifting problem in a dynamic context.

In chapter 2, I study a continuous-time moral hazard model in which the principal hires a team of agents to run the business. The firm consists of multiple divisions and agents exert costly efforts to improve the divisional cash flows. The firm size evolves stochastically based on the aggregate cash flows.The model delivers a negative relationship between firm sizes and pay-for-divisional incentives, and I characterize conditions under which joint/relative performance evaluation will be used. I also explore the implications

of team production on the firm's optimal capital structure and financial policy.

In chapter 3, I study a multi-armed bandits problem with ambiguity. Decision-maker views the probabilities underlying each arm as imprecise and his preference is represented by recursive multiple-priors. I show that the classical "Gittins Index" generalizes to a "Multiple-Priors Gittins Index". In the setting with one safe arm and one ambiguous arm, the decision-maker plays the ambiguous arm if its "Multiple-Priors Gittins Index" is higher than the return delivered by the safe arm. In the multi-armed environment, I obtain the "Multiple-Priors Index Theorem" which states that the optimal strategy for the decision-maker is to play the ambiguous arm with the highest Multiple-Priors Index.

Contents

1 Dynamic Agency and Endogenous Risk-Taking **1**

1.1 Introduction 1

1.1.1 Motivation 1

1.1.2 Literature Review 5

1.2 The Model 8

1.3 Model Solution 12

1.3.1 Incentive compatibility. 12

1.3.2 Derivation of the Optimal Contract. 14

1.3.3 Discussion. 20

1.4 Implication and Analysis 24

1.4.1 Quantitative Analysis. 25

1.4.2 Capital structure implementation. 28

1.4.3 Compensation in Hedge Fund Industry. 34

1.5 Extensions 37

1.5.1 Replacing the agent. 37

1.5.2 Multiple Projects. 39

1.5.3 Random Jump Size. 41

1.6 Conculsion 42

2 Dynamic Team Incentives **44**

2.1 Introduction 44

2.2 The Model 46

2.2.1 The Contracting Problem. 48
2.2.2 First-Best Solution. 50
2.3 Model Solution 51
2.3.1 The Dynamics of Continuation Utility. 51
2.3.2 HJB-equation. 54
2.3.3 Model Implication 56
2.4 Capital Structure 59
2.4.1 Optimal Contracting in an Unlevered Firm 60
2.4.2 Optimal Contracting in a Levered Firm. 61
2.4.3 Discussion. 64
2.5 Conclusion 67

3 Multi-armed Bandit Problems with Ambiguity **69**
3.1 Introduction 69
3.2 The Model 72
3.3 Main Results 74
3.3.1 The Multiple-Priors Gittins Index. 75
3.3.2 The Multiple-Priors Gittins Index Theorem 78
3.4 Conclusion 79

Appendix A **80**

Bibliography **101**

Curriculum Vitae **105**

List of Tables

1.1 **Comparative Statics of** α, γ, L**.** 26
1.2 **Comparative Statics of** μ, l, σ, B**.** 27

2.1 **Capital Structure and default policy with different number of agents.** 66

List of Figures

1.4.1 **The Principal's Value Function.** 25

1.5.1 **Replacing the agent**. 38

2.4.1 **Effort policy in unlevered firms.** 64

2.4.2 **Effort Policy in levered firms**. 65

Chapter 1

Dynamic Agency and Endogenous Risk-Taking

1.1 Introduction

1.1.1 Motivation

Hidden risk-taking is a pervasive fact of life. Consider an engineer working in an oil refinery. His daily task is to help the firm earn enough profits and at the same time ensure the refinery does not explode. However, he could tune up the production rate secretly and expose the facilities to risk in order to generate more profits in the short run. In 2007-2009 financial crisis, we discover that many financial institutions excessively engaged in risky transactions. For instance, financial institutions purchased a lot of "toxic" assets or entered into risky credit default swaps with certain counterparties. While these assets are certainly profitable during normal times, infrequent defaults lead to severe losses. These financial institutions have incentives to engage in such risky investment because they reap a share of the investment profits. Very often, investment strategies and portfolio choices are complex and costly to communicate to the investors.

In hedge funds industry, the fund managers are usually compensated by the high-water mark contract. The contract delivers cash payment to the manager when the managed asset crosses the high-water mark threshold. Notably, recent empirical evidence drawn from time-series data shows that hedge funds exhibit higher return's volatility when their asset shrinks and the managers' performance deteriorates.[1] This implies that fund managers choose to hold more risky assets in their portfolio and invest in a more aggressive way when the asset size lies far below the payment threshold. The asset-substitution argument

[1] see Shelef (2013) and Kolokolova and Mattes (2013) and discussion in section 4.

suggests that fund returns volatility should be high all the times and independent of the size of the asset. Therefore, the standard argument fails to accommodate the new empirics. Why is this the case? How do we rationalize managers' behaviors? Are the existing compensation contracts to the hedge fund managers efficient? Does this type of contract induce risk-taking? The goal of this paper is to study dynamic contracting with hidden risk-taking and propose an agency-based explanation for this newly-documented phenomenon.

In my model, investors hire a manager who enjoys limited liability to run the firm and both parties are risk-neutral. Moral hazard is *two dimensional*: the agent chooses his effort level and risk strategy at any moment of time in order to operate the business. One can interpret the risk dimension of moral hazard as (i) choice of risky or safe projects; or (ii) choice of risky or safe financial assets. Firm's cash flows follow a jump-diffusion. High effort increases the drift of the diffusion component without affecting the volatility of the cash flows process. Risk-taking[2] enhances short term profits by raising the cash flow drift, but at the same time increases the intensity of the jump part, which models rare losses. I assume that downside risk-taking destroys values: it reduces the firm's net present value (NPV). To solve for the optimal contract in the continuous-time setting, I follow the martingale method developed by DeMarzo and Sannikov (2006), Sannikov (2008), and Williams (2009,2011). The method is a powerful tool that provides clean and elegant description of the agent's continuation value, which serves as a state variable in the principal's stochastic control problem. By controlling how the continuation value responds to the observed cash flows, the principal will be able to provide appropriate incentives. As the agent has two choices, there are two incentive constraints restricting her behavior. Specifically, incentive loadings on the Brownian shock control the agent's effort, and the exposure to Poisson risk affects the agent's risk-taking behavior. These two constraints are related by the fact that risk-taking simultaneously affects the drift and intensity.

The optimal contract between the investors and the agent specifies (i) the agent's cash compensation and her promised utility; (ii) the firm's liquidation decision; and (iii) the

[2] I use the terms "risk-taking", "risk-shifting", and "gambling" interchangeably.

principal's recommended risk strategy, as a function of the entire past cash flows history. In the optimal contract, to incentivize the agent to work hard, continuation value has to increase following a high cash flow realization. When this incentive is too strong, it is tempting for the agent to take risks. In order to prevent risk-taking, an upper bound on the incentive loadings to the Brownian risk is imposed by introducing sufficiently large punishment, that is a downward jump on the agent's continuation value, following loss events. However, when the continuation value is low, punishment is in conflict with the limited liability constraint in the sense that the agent will be left with a negative payoff in the game after being punished severely.

In addition, I find that the following trade-off is the key to understand the optimal risk-taking policy. On one hand, allowing the agent to gamble leads to lower NPV of the firm. On the other hand, implementing the safe action to prevent large risks is costly to the principal in the following sense: punishing the agent after a loss means the agent is exposed to Poisson risks, which brings in additional volatility to the agent's continuation value process. This in turn implies that inefficient liquidation would occur more often. This cost is referred as the *agency cost of preventing risk-taking*. Intuitively, this agency cost is decreasing in the agent's continuation value because when the firm is close to the liquidation boundary, exposing the agent to risks is more costly. To determine the optimal action, the principal balances the expected monetary loss and the agency cost. Along the equilibrium path, the principal implements the safe action when the agent's continuation value is high and allows the agent to secretly gamble otherwise. Therefore, the model generates rich dynamics on project or asset selections.

The optimal contract can be implemented using standard securities. In the capital structure that is considered in section 4, the firm accumulates cash in order to fund its operation and cover any short term financing needs. From the incentive perspective, the cash balance keeps track of the agent's performance. The agent is compensated by holding a fraction of firm's common stocks. Inside stake provides the right incentive for the agent to work hard. Outside investors hold long-term bonds and common stocks. The firm would also purchase an insurance contract, or any derivative contract. The insurance contract, provided by a competitive insurance company, gives the firm an opportunity to hedge against the Poisson risk by paying a periodic contractual premium to the insurance company.

The security design exercise features that to motivate the agent to choose the safe project and not to take any hidden risk, incomplete hedging is required: If hedging were complete, then the agent's payoff was completely insured against downside risks and share no loss from any undesirable events, and thus would have incentive to secretly gamble. The results shed light on how risk-management policy and project selection are linked up by agency conflicts and provides an agency-based explanation for the reasons why sometimes we see firms only hedge partially. Therefore my model generates a dynamic capital structure for a firm that has flexibility to choose over projects that carry different degree of large risks. And the model predicts how cash flows volatility, expected returns, hedging, and credit rating are correlated.

On top of the contractual implications on capital structure, the optimal contract in the present setting can be interpreted as the "high-water mark" contract. The high-water mark contract is a predominant form of compensation contract for hedge fund managers. The manager is delegated by the investors certain amount of capital to invest in different financial assets. When the fund performs well and hits a threshold, bonus payment is delivered to the manager, and the threshold moves up–hence the term "high-water mark". The next time the manager is paid by bonus, the fund's cumulative performance has to reach this new threshold, and any previous loss has to be recouped before she get paid. Because the contract exhibits a convex payoff structure: when the fund's performance is below the threshold, the manager gets no bonus, and is compensated once the performance crosses the threshold.

A critical issue is whether the contract actually induces the manager to take excessive hidden tail risk. My results show that the "high-water mark" contract is an optimal contract, and whether the agent will secretly gamble depends critically on the current fund performance. In particular, when the performance is close to the bonus threshold, an impatient manager will not take hidden risks because more frequent losses "drag" the fund away from the threshold. However, when the fund performance is poor and drift away from the water mark, risk-taking problem becomes more severe. As the manager is protected by limited liability, risk-taking helps pushing the fund to the threshold, and she is not required to bear all the losses. Therefore, my model predicts that in a dynamic environment, the distance of the fund to its high-water mark threshold is positively correlated with the volatility of the fund's returns and negatively with its expected performance. This provides an agency-based explanation for currently documented "distance-to-threshold" phenomenon and highlights that "distance-to-threshold" is the key to understand the risk-shifting problem.

1.1.2 Literature Review

This paper belongs to a fast-growing literature on continuous-time dynamic contract theory based on techniques developed by Sannikov (2008) and Williams (2006). Using a martingale method, DeMarzo and Sannikov (2006) (DS hereafter) studies an agency model in which cash flows follow an arithmetic Brownian motion. They found that the optimal contract can be implemented using a combination of credit line, debt and equity. Enriching their setting to jump-diffusions, my model provides an agency-based explanation for "gambling for resurrection" and develops additional implications on capital structure. Other recent works on dynamic agency include Biais, Mariotti, Rochet, and Villeneuve (2010) (BMRV thereafter), DeMarzo *et al.* (2012), He (2008), Hoffmann and Pfeil (2010), Piskoroski and Tchistyi (2010), Miao and Rivera (2013), Williams (2009, 2011), Wong (2013), Zhang (2009) and Zhu (2012).[3] However, agency issue is one dimensional in all these works. This paper attempts to introduce an additional dimension of moral hazard in continuous-time dynamic contract theory.

[3] Cvitanic and Zhang (2012) summarize some recent developments and techniques used in continuous-time contract theory. See also Sannikov (2012) for applications in corporate finance.

In an independent work, DeMarzo, Livdan, and Tchistyi (2011) is closely related to my paper. They consider an agency model with three-outcome space: high cash flows, low cash flows, and a "disaster" state, which corresponds to a large loss in my case. The agent in their model, like mine, can choose to divert cash flows and at the same time choose the riskiness of the project. In particular, risk-shifting leads to a higher probability of a disaster. They show that to prevent risk-taking, it would be much less costly if contract can be written directly on the state of nature. After I finished the first draft of my paper, I found that DeMarzo, Livdan and Tchistyi have developed in a continuous-time extension of their model. The cash flows we both consider is a jump-diffusion. The key difference is that in their work, the downward jump is a "disaster" that leads to the termination of firm immediately and it only occurs when the agent takes gambles while I allow for multiple jumps. Similar to my result, they show that in the optimal contract, the principal would allow the agent to engage in sub-optimal risk-taking when the agent performs poorly. Their optimal contract also calls for randomization over some interval of the continuation utility, which seems to make capital structure implementation less appealing. However, if one focuses on the effect of agency issue on risk-taking, our results are complementary to each other.

Biais, Mariotti, Rochet, and Villeneuve (2010) develops a continuous-time model in which the agent exerts costly effort to prevent large loss. The cash flows in their model is a pure jump process and their focus is on the firm's capital dynamics. They show that when incentive-compatibility is in conflict with limited liability, the principal will downsize, or partially liquidate, the firm's capital. This happens on equilibrium when the firm hits by a sequence of bad shocks. My model differ from their in that I do not consider contractible capital adjustment and allow for multi-tasking. Providing incentive for the agent to take an inefficient action is second-best in my model. Thus my analysis highlights that partial liquidation and risk-taking could potentially be two alternative ways to provide incentives.

Another related paper is Szydlowski (2012). He introduces multi-tasking in the q-theory model by DeMarzo *et al.* (2012). There are n independent tasks in his model and the output of each task is driven by an arithmetic Brownian motion. The agent exerts costly effort to control the drift of each Brownian motion. Similar to my model, even if a task generates positive NPV, if the incentive cost of motivating the agent to work hard in a task is too high, the principal prefers to let the agent shirk in that task. However, the total cash flows volatility remains the same regardless of the number of tasks the agent is putting effort in, as the agent only controls the drift, not the volatility. Unlike my model, the agent could affect the Poisson intensity and thus the variance of the cash flows. Hence his model cannot capture the relationship between project choice dynamics and asset riskiness.

In the static context, Biais and Casamtta (1999) study an agency model with risk-taking. The project generates three outcomes: high, medium and low. Risk-taking is hidden, destroys value and alters the returns distribution in the sense of second-order stochastic dominance. The incentive constraints they obtain are similar to mine. They further show that the optimal contract can be implemented by a mixture of debt and equity and when moral hazard issue on effort is severe enough, stock option helps to provide incentives. Palomino and Prat (2003) study risk-taking in a delegated portfolio choice model in which the underlying technology displays a high-risk-high-return relation. Their optimal contract also has the bonus feature. While I focus more on the downsize risk, my model is more suitable for discussing the high-water mark contract because this contract is dynamic in nature: the bonus threshold change over time depending on the fund performance.

In continuous-time, Sung (1995) studies linear contract and project selection. In his model, an agent with exponential utility selects project at time 0 by affecting the volatility of the cash flows process. To obtain a meaningful moral hazard problem, the principal is assumed to observe the investment return only at the terminal time. Since there is no interim information, project is chosen at the beginning of time and there is no project selection dynamics. Cadenillas et al. (2007) also consider project selection in continuous-time. However, their focus is on the first-best optimal risk-sharing rule, that is, they do not consider any form of agency issue.

The rest of the paper is structured as follows. Section 1.2 develops the formal model. Section 1.3 describes the model's solution and characterizes the optimal contract. Capital structure implementations and empirical implications are discussed in section 1.4. Section 1.5 presents extensions. The last section concludes and proofs are delegated to the Appendix.

1.2 The Model

Consider a continuous-time principal-agent model, in which investors (the principal) hire a manager (the agent) to operate the firm. Both the principal and the agent are risk-neutral and discount future cash flows at rate r and ρ respectively, where $\rho > r > 0$. The principal has unlimited wealth while the agent is protected by limited liability and has no initial wealth. The firm produces cash flows according to a jump-diffusion

$$dY_t = \mu(a_t)dt + \sigma dZ_t - LdN_t$$

where $\mu(a_t)$ is the drift of the cash flows, $\{a_t\}_{t\geq 0}$ is the action process to be described below, $\sigma > 0$ is the volatility, $L > 0$ is the size of the loss,[4] $Z = \{Z_t\}_{t\geq 0}$ is a standard Brownian motion and $N = \{N_t\}_{t\geq 0}$ is a standard Poisson process with intensity $\{\lambda(a_t)\}_{t\geq 0}$. The drift and the intensity are controlled by the agent's action a_t. Both Z and N are defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ that satisfies the usual condition[5] and they are independent. The principal can only observe the cash flows process $\{Y_t\}_{t\geq 0}$ but not the agent's action process $\{a_t\}_{t\geq 0}$. Since the cash flow path jumps downward when a loss occurs, the principal is able to identify all the loss events. This implies that the principal essentially observes the diffusion component $X = \{X_t\}$, where $dX_t = \mu(a_t)dt + \sigma dZ_t$ and the jump component N separately. In other words, the principal's information is modeled by taking the filtration $\mathcal{F}_t = \sigma(\{X_s, N_s\}_{s<t})$ generated by X and N.

[4]Random jump size is considered in section 5.

[5]That is, (i) the probability space $(\Omega, \mathcal{F}, P)$ is complete and $\mathcal{F}_0$ contains all P-null sets in $\mathcal{F}$, and (ii) the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous: $\mathcal{F}_t = \mathcal{F}_{t+}$ for all t.

Moral hazard is *two-dimensional.* At each moment of time, the agent can choose an amount of effort to supply *and* an action that relates to the firm's risk. Formally, $a_t^e \in \{a_L, a_H\}$ is the agent's effort choice, where a_L denotes low effort or "shirking" and a_H is high effort or "working". High effort increases the drift by $\mu > 0$ while low effort allows the agent to derive a private benefit $B > 0$. $a_t^s \in \{a_N, a_R\}$ is the agent's risk-action. a_N denotes "safe action" or "non-risk-taking" behavior and a_R stands for "risky action" or "risk-taking" behavior. The Poisson intensity is $\lambda > 0$ when the agent is not taking risks and $\lambda + \gamma$, where $\gamma > 0$ when the agent takes risks. Moreover, when the agent takes risk, the drift of the cash flows increases by $\alpha > 0$. Intuitively, risk-taking behavior leads to a higher short-term profit at a cost of increased likelihood of infrequent large losses. Note that there is no cost for the agent to take risk. I refer $\{a_t = (a_t^e, a_t^s)\}_{t\geq 0}$ as *action process.* The following table summarizes the above discussion

action	drift	intensity	private benefit
(a_L, a_N)	0	λ	B
(a_H, a_N)	μ	λ	0
(a_L, a_R)	α	$\lambda + \gamma$	B
(a_H, a_R)	$\mu + \alpha$	$\lambda + \gamma$	0

Notice that the riskiness choice admits multiple interpretations. One can think of the agent is choosing between a safe project or a risky project. Another interpretation is that the agent is choosing among financial assets with different riskiness.[6] It should also be noted that risk-taking leads to a higher cash flow variance: if $a_t^s = a_R$, $\text{var}(dY_t) = (\sigma^2 + L^2(\lambda + \gamma))dt$. and if $a_t^s = a_N$, $\text{var}(dY_t) = (\sigma^2 + L^2\lambda)dt$.[7]

[6]To see this more formally. One can rewrite the cash flow process as follows: $dY_t = 1_{\{a_t^s = a_R\}} dY_t^R + 1_{\{a_t^s = a_N\}} dY_t^N$, where $dY_t^R = (\mu(a_t^e) + \alpha)dt + \sigma dZ_t - LdN_t$ with Poisson intensity $\lambda + \gamma$, and $dY_t^N = \mu(a_t^e)dt + \sigma dZ_t - LdN_t$ with Poisson intensity λ. Hence the choice of safe/risky action is equivalent to the choice of different cash flow structure, ie., different projects. Moreover, note that the quadratic variation of Y^R and Y^N are the same: for any t, $[Y^R]_t = [Y^N]_t$, therefore the principal cannot observe the project choice.

[7]Sung (1995) assumes that the agent controls the volatility of cash flow process (which is an arithmetic Brownian motion). However, if the principal can continuously observe the cash flows, she can recover the agent's action by estimating the quadratic variation of the cash flows. Therefore Sung assumes that the principal only observes the terminal cash flows to obtain a non-trivial moral hazard problem of project

At time 0, the principal offers a contract to the agent. The contracting party can fully commit to the contract. A contract specifies a liquidation time τ, payments $I = \{I_t\}_0^\tau$ that are functions of past cash flows, and a recommended action process $a = \{a_t\}_{t\geq 0}$. Denote a contract as $\Gamma = (I, \tau, a)$. At the liquidation time, the principal receives the liquidation value of the project $l \geq 0$ and the agent receives nothing. Because of limited liability, the cumulative payment process $I = \{I_t\}_0^\tau$ is non-negative and increasing. At any time t prior to liquidation, the sequence of events during a short time interval $[t, t+dt)$ can be described as follows:

1. The agent takes an action $\hat{a}_t$.
2. The cash flow dY_t realizes. And with probability $\lambda(\hat{a}_t)dt$, there is a loss, in which case $dN_t = 1$; otherwise $dN_t = 0$.
3. The agent receives a non-negative compensation dI_t.
4. The principal decides either to continue or liquidate the project.

Based on this timing, I require formally that the action process a is $\mathcal{F}_t$-predictable, the payment process I is $\mathcal{F}_t$-adapted and the liquidation time τ is $\mathcal{F}_t$-stopping time that can take the value ∞. Note that an action process a induces a unique probability measure P^a over the paths of the cash flows $\{Y_t\}_{t\geq 0}$. Therefore, I let $E^a(.)$ denote the expectation operator under the measure P^a.

Given a contract $\Gamma = (I, \tau, a)$ and an action process $\hat{a}$, the agent's total expected discounted utility at time 0 is given by

$$E^{\hat{a}}\left[\int_0^\tau e^{-\rho t}(dI_t + \mathbf{1}_{\{\hat{a}_t^e = a_L\}} B dt)\right] \tag{1.2.1}$$

selection. In the current setup, the quadratic variation of the cash flows on interval $[0, t]$ is $[Y]_t = \int_0^t \sigma^2 ds + L^2 \int_0^t N_s ds$. As the agent's action does not map one-to-one to the quadratic variation process, my setup is free of this problem.

and the principal's total expected discounted profit at time 0 is given by

$$E^a\left[\int_0^\tau e^{-rt}(dY_t - dI_t) + e^{-r\tau} l\right] \tag{1.2.2}$$

Facing any contract Γ, the agent's decision problem is to choose an action process $\hat{a}$ that maximizes her expected utility (**??**). Denote the solution to this maximization problem as $a^* = \{a_t^*\}_0^\tau$. This solution needs not be the same as the recommended action a. If for a given contract $\Gamma = (I, \tau, a)$, the agent's choice satisfies $a_t^* = a_t$ for all $t \in [0, \tau)$, then the contract Γ is said to be *incentive compatible*. The optimal contracting problem of the principal is to find an incentive compatible contract Γ that maximizes her expected discounted profit (**??**) subject to delivering the agent a required level of expected utility W_0.

Throughout the paper, I maintain the following assumptions

- **Assumption 1:** $\frac{\mu+\alpha-L(\lambda+\gamma)}{r} > l$

The first assumption says that liquidation is inefficient. In particular, even if the agent is taking risks all the time and large losses occur quite frequently, as long as the agent works hard forever, the firm still generates a discounted cash flows that is strictly greater than the project's liquidation value[8].

- **Assumption 2:** $\alpha - L\gamma \leq 0$

The second assumption says that risk-taking is potentially harmful to the principal in the sense that it reduces the NPV of the firm. This implies in the optimal contract, the principal would try to avoid implementing risk-taking behavior as it leads to lower profits. Assumption 1 and 2 together imply $\frac{\mu-L\lambda}{r} > l$.

- **Assumption 3:** $\mu > B$

[8] Essentially this requires the downside risks not to be too bad. If this assumption fails to hold, the principal prefers to shut down the firm instead of providing incentive for risk-taking.

The last assumption says that when the agent shirks, the loss in the expected profits of the project is larger than the agent's private benefit of shirking. The last two assumptions suggest that risk-taking and shirking at the same time is not socially efficient. When there is no agency issue, namely, a is contractible, then under the above assumptions, the first best contract can be characterized as follows: by assumption 2 and 3, maximal effort and safe action $a_t = (a_H, a_N)$ are implemented at every moment of time. The expected discounted cash flow is $\frac{\mu - L\lambda}{r}$. Since the agent is more impatient than the principal, the principal pays an amount W to the agent at time 0. By assumption 1, the principal never liquidates the project. Principal's value under the first-best contract subject to delivering the agent an amount W is $F^{FB}(W) = \frac{\mu - L\lambda}{r} - W$.

1.3 Model Solution

1.3.1 Incentive compatibility.

This section provides a necessary and sufficient condition for incentive compatibility of any contract. Fix any arbitrary contract Γ and an action process a, the agent's continuation value at time t is defined by

$$W_t(\Gamma) \equiv E_t^a \left[\int_t^{\tau} e^{-\rho s} (dI_s + \mathbf{1}_{\{a_s^e = a_L\}} B ds) \right] \tag{1.3.1}$$

In other words, $W_t(\Gamma)$ is the agent's promised utility at time t given contract Γ and when she plans to follow the continuation of action process a from t onward. Note that the continuation value at t reacts to information and hence summarizes all past performance of the agent. The principal could control how the continuation value responds to cash flow realizations and loss events in order to provide incentives. Using the Martingale Representation Theorem, the following proposition characterizes the evolution of the continuation value process $\{W_t(\Gamma)\}_{t \geq 0}$ in terms of observed cash flows and loss events. In addition, it provides a key condition for the agent to follow the recommended actions. Let $W_{t-} = \lim_{s \uparrow t} W_s$.

Proposition 1.1. *Given any contract* $\Gamma = (I, \tau, a)$*, there exists a* $\mathcal{F}_t$*-predictable process* $\{(\beta_t, \psi_t)\}_0^\tau$ *such that for any* $t < \tau$*,*

$$dW_t(\Gamma) = (\rho W_{t-}(\Gamma) - 1_{\{a_t^e = a_L\}} B)dt - dI_t + \beta_t(dX_t - \mu(a_t)dt) - \psi_t(dN_t - \lambda(a_t)dt) \quad (1.3.2)$$

Moreover, a necessary and sufficient condition for the contract Γ *to be incentive compatible for high effort is that,* $a_t = (a_H, a_N)$ *if and only if* $\beta_t \geq \frac{B}{\mu}$ *and* $\psi_t \gamma \geq \beta_t \alpha$*, and* $a_t = (a_H, a_R)$ *if and only if* $\beta_t \geq \frac{B}{\mu}$ *and* $\psi_t \gamma \leq \beta_t \alpha$*, for any* $t \in [0, \tau)$*.*

In the optimal contract, the principal seeks to implement high efforts all the time but allow the agent's strategy to vary. Therefore, proposition 1 states that the dynamics of the agent's continuation value follows

$$dW_t = \rho W_{t-} dt - dI_t + \beta_t \sigma dZ_t^a - \psi_t dM_t^a$$

where $Z^a = \{Z_t^a\}_{t \geq 0}$ is a Brownian motion and $M^a = \{M_t^a\}$ is a compensated Poisson process under probability measure P^a. The drift of W contains two components: the first term $\rho W_{t-} dt$ corresponds to the promise-keeping constraint. As the principal promises a balance W_t at time t to the agent, the promised amount grows at rate ρ since the agent is impatient and discounts future cash flows at rate ρ. The second component is an immediate cash payment dI_t. Paying dI_t reduces the promise at time t exactly by dI_t.

The diffusion and jump part of W_t are connected to the incentive component of an contract because deviation from the recommended action affects continuation value through its effects on both Z^a and M^a. To understand the incentive constraints, suppose at time t, the principal recommends working. If the agent shirks, she obtains a private benefit Bdt, but loses $\beta_t \mu dt$ in compensation. This is because by shirking, the drift of dX_t reduces by μdt, and that translates to a reduction in continuation value by $\beta_t \mu dt$. Thus, the agent will choose high efforts if and only if $\beta_t \mu \geq B$. Similarly, suppose at time t, the principal recommends the agent not to take risk. If the agent deviates to take risk, her marginal benefit is $\beta_t \alpha dt$ because the drift of dX_t is increased by αdt. And the marginal cost is $\psi_t \gamma dt$ since by taking risks, the probability for a large loss to arrive within $[t, t+dt)$ is increased by γdt, and when a loss occurs, the agent is punished by a reduction in the promised utility by an amount ψ_t. Hence, the principal can prevent risk-taking if and only if $\psi_t \gamma \geq \beta_t \alpha$.[9]

1.3.2 Derivation of the Optimal Contract.

In this section, I study the Hamilton-Jacobi-Bellman (HJB) equation for the principal. Since there is only one state variable: the agent's continuation value W_t, I let $F(W)$ denote the principal's value function. In particular, after any history of the firm which is summarized by W_t, $F(W_t)$ is the principal's continuation value in an optimal contract. I assume for the moment that $F(W)$ is concave and twice differentiable.

Note that there are some key and standard properties for the value function $F(W)$. First, it is obvious that the value function cannot exceed the first best, that is $F(W) \leq F^{FB}(W)$ for all W. Second, the principal can always choose to pay the agent a lump-sum $dI > 0$ immediately and moving down the promise from W to $W - dI$. Therefore, $F(W) \geq F(W - dI) - dI$, which implies $F'(W) \geq -1$ for all W. The right-hand side of the inequality is the marginal cost of an immediate payment and $F'(W)$ can be interpreted as the marginal benefit of deferring cash payment. As the agent has a higher discount rate $\rho > r$, the benefit of deferring compensation will decrease as the promised balance grows. This implies that it is optimal to pay the agent when her continuation value hits a payment threshold defined by $W^p = \min\{W : F'(W) = -1\}$. Thus, the optimal amount to be paid is

$$dI = \max\{W - W^p, 0\}$$

It follows that for $W_t > W^p$, $F(W_t) = F(W^p) - (W_t - W^p)$. Third, because of limited liability, the agent's continuation payoff has to be non-negative. So for any t, $W_t \in [0, W^p]$. Moreover, because W_t has positive volatility regardless of choice of the risk strategy, thus

[9] The agent could deviate to low effort and risk-taking simultaneously. But under the above two constraints, she has no incentive to do so. See the appendix for a proof.

when $W_t = 0$, liquidation is necessary.

Now, I turn to the characterization of $F(W)$ for $W \in [0, W^p]$. Since the principal can implement different risk-taking strategies, I first write down the equations that the principal's profit has to satisfy if the principal recommends different choices. Let $F_N(W)$ be the principal's profit when implementing $a_t = (a_H, a_N)$ for all t. By proposition 1.1, the required incentive constraints are $\beta_t \geq \frac{B}{\mu}$ and $\psi_t \gamma \geq \beta_t \alpha$. Suppose F_N is concave, then it is optimal to set $\beta_t = \frac{B}{\mu}$ and $\psi_t = \frac{B}{\mu}\frac{\alpha}{\gamma}$ for all t such that $a_t = (a_H, a_N)$. From now on, let $\beta = \frac{B}{\mu}$ and $\psi = \frac{B}{\mu}\frac{\alpha}{\gamma}$. The dynamics of the agent's continuation value is

$$dW_t = (\rho W_t + \psi\lambda)dt + \beta\sigma dZ_t - \psi dN_t \tag{1.3.3}$$

Then the principal's profit satisfies the *safe action delay differential equation (DDE)* :

$$rF_N(W) = \mu - L\lambda + F_N'(W)(\rho W + \psi\lambda) + \frac{1}{2}F_N''(W)\beta^2\sigma^2 + \lambda(F_N(W-\psi) - F_N(W)) \tag{1.3.4}$$

To simplify notation, let $\mathcal{L}_N$ be the infinitesimal generator of the process (**??**). That is $\mathcal{L}_N F(W) = F'(W)(\rho W + \psi\lambda) + \frac{1}{2}F''(W)\beta^2\sigma^2 + \lambda(F(W-\psi) - F(W))$.

Denote by $F_R(W)$ the principal's profit when she implements $a_t = (a_H, a_R)$ for all t, the associated incentive constraints are $\beta_t \geq \frac{B}{\mu}$ and $\psi_t\gamma \leq \beta_t\alpha$. Assume that F_R is concave, then it is optimal to set $\beta_t = \frac{B}{\mu}$ and $\psi_t = 0$ for all t such that $a_t = (a_H, a_R)$. The dynamics of the agent's continuation value is

$$dW_t = \rho W_t dt + \beta\sigma dZ_t \tag{1.3.5}$$

Then the principal's profit satisfy the *risky action ordinary differential equation (ODE)* :

$$rF_R(W) = \mu + \alpha - L(\lambda + \gamma) + F_R'(W)\rho W + \frac{1}{2}F_R''(W)\beta^2\sigma^2 \tag{1.3.6}$$

Similarly, let $\mathcal{L}_R$ be the infinitesimal generator of the process (**??**) so that $\mathcal{L}_R F(W) =$

$F'(W)\rho W + \frac{1}{2}F''(W)\beta^2\sigma^2$.

Notice that the functions $F_N(W)$ and $F_R(W)$ only tell us how the principal's profit would evolve when implementing different actions, but not the optimal profit. However, based on these two differential equations, the HJB equation for the principal is given by

$$rF(W) = \max\left\{ \mu - L\lambda + \mathcal{L}_N F(W),\ \mu + \alpha - L(\lambda+\gamma) + \mathcal{L}_R F(W) \right\} \tag{1.3.7}$$

The structure of the HJB-equation suggests that the optimal policy can be obtained by comparing the components in the max operator. Because the comparison only depends on the state variable W, the action policy is Markovian. More formally, first consider the continuation value in the interval $[\psi, W^p]$. If $\mu - L\lambda + \mathcal{L}_N F(W) \geq \mu + \alpha - L(\lambda+\gamma) + \mathcal{L}_R F(W)$, then the principal would like to choose $a_t^s = a_N$. If otherwise, the principal chooses $a_t^s = a_R$. Notice that the previous inequality is equivalent to $F'(W)\psi\lambda + \lambda(F(W-\psi) - F(W)) \geq \alpha - L\gamma$. Define the following function on $[\psi, W^p]$

$$\mathcal{A}(W) = F'(W)\psi\lambda + \lambda(F(W-\psi) - F(W)) \tag{1.3.8}$$

Intuitively, the first term in (**??**) is the increase in the drift of the agent's continuation value if the principal implements the safe action. When W is low, an increase in drift is beneficial in the sense that the continuation value will move faster away from the liquidation boundary. But when W is higher, a higher drift is costly because it implies more frequent cash payments are made. The second term in (**??**) works in the reverse direction. When W is low, a downward jump in continuation value implies a greater chance of liquidation and for high W, it reduces the likelihood of making cash compensation. However, given liquidation is inefficient, the principal essentially exhibits risk-averse behavior. Incentive-compatible safe action thus implies a higher volatility on the agent's continuation value: the agent is exposed to Poisson risks. Therefore $\mathcal{A}(W) < 0$ for all $W \in [\psi, W^p]$ and one can interpret $|\mathcal{A}(W)|$ as the *agency cost of preventing risk-taking*. In addition, $\mathcal{A}(W)$ is strictly increasing in W, meaning that the agency cost decreases as the firm moves away from the liquidation boundary. Because the expected monetary loss is constant, one expects that there exists a threshold $W^* \in [\psi, W^p]$ such that for $W < W^*$, risk-taking will be optimal as the agency cost of preventing the agent to take downside risks is too high, and when $W \geq W^*$, the safe action will be optimally implemented as the monetary loss from the risky action is too high. Lemma 9 in the appendix proves these properties of $\mathcal{A}(W)$ and establishes the existence of W^*.

Next, consider the continuation value in the interval $[0, \psi)$. A critical observation is that W_t cannot enter into the interval $[0, \psi)$ *if the principal implements* $a_t^s = a_N$ *at all times.*[10] Intuitively, this happens because to motivate the agent to choose the safe action, incentive compatibility requires a large enough punishment following an observable large loss. However, when the agent's continuation value is too low, such a punishment cannot be imposed. That is, incentive compatibility is in conflict with limited liability. In other words, there is no incentive compatible contract that implements the safe action on the interval $[0, \psi)$. Therefore, for the firm's operation to continue, the principal would allow the agent to take risk on this region.

The risky action ODE, equation (**??**), is a second-order ordinary differential equation. I have argued above that limited liability requires risk-taking on the interval $[0, \psi)$. On this interval, the dynamics of the principal's profit is represented by equation (**??**) when the agent chooses the safe action. When $W_t = 0$, liquidation is necessary, hence $F_R(0) = l$ is a required boundary condition.

Combining the above analysis, the optimal action policy can be summarized as follows

$$a_t = a(W_t) = \begin{cases} (a_H, a_R) & \text{if } W_t \in [0, W^*] \\ (a_H, a_N) & \text{if } W_t \in (W^*, W^p] \end{cases}$$

[10]If the principal seeks to implement $a_t^s = a_N$ for all t, the suitable initial history function is $F_N(W) = \frac{m}{\psi} W + l$. It can be shown that there exists a maximal m such that the value function satisfies this initial history and boundary conditions $F_N'(W^p) = -1$ and $F_N''(W^p) = 0$. This specification is equivalent to stochastic liquidation when W_t crosses ψ from above. However, this policy is not optimal because the risky action ODE starts out concave and could lie above the linear part of F_N.

In the case where $W^* \in (\psi, W^p)$, W^* is implicitly defined by

$$\mathcal{A}(W^*) = \alpha - L\gamma \tag{1.3.9}$$

Applying the optimal policy, it follows that the HJB-equation (**??**) can be viewed as a differential equation

$$rF(W) = \mu - L\lambda + 1_{\{W \le W^*\}}(\alpha - L\gamma) + F'(W)(\rho W + 1_{\{W > W^*\}}\psi\lambda) \tag{1.3.10}$$

$$+\frac{1}{2}F''(W)\beta^2\sigma^2 + \lambda(F(W - 1_{\{W>W^*\}}\psi) - F(W))$$

Based on the discussion in the previous paragraphs, I impose the standard boundary conditions: $F(0) = l$, $F'(W^p) = -1$, and $F''(W^p) = 0$, in order to solve this differential equation. The main result for the optimal contract is stated as proposition 1.2.

Proposition 1.2. *The principal's value function is given by the solution F to equation (**??**) with boundary conditions $F(0) = l$, $F'(W^p) = -1$, and $F''(W^p) = 0$. The optimal contract (I, τ, a) that implements high efforts all the time and delivers the agent initial utility W_0 takes the following form: (i) Incentive loadings are $\beta_t = \frac{B}{\mu}$ and $\psi_t \in \{\frac{B}{\mu}\frac{\alpha}{\gamma}, 0\}$. There exists a switching point $W^* \in [\psi, W^p]$ such that when $W_t \in (0, W^*]$, the principal implements the risky action, $a_t^s = a_R$ and imposes no punishment, $\psi_t = 0$. The agent's continuation value follows $dW_t = \rho W_t dt + \beta\sigma dZ_t$. When $W_t \in (W^*, W^p]$, the principal implements the safe action, $a_t^s = a_N$ and imposes positive punishment $\psi_t = \frac{B}{\mu}\frac{\alpha}{\gamma} > 0$. The agent's continuation value follows $dW_t = (\rho W_{t-} + \psi\lambda)dt + \beta\sigma dZ_t - \psi dN_t$. (ii) Cash payments: $dI_t = \max\{W_t - W^p, 0\}$ reflects W_t back to W^p. (iii) Liquidation: $\tau = \min\{t | W_t = 0\}$, at which time the principal receives l and the agent gets 0.*

To summarize, when $W_t \in (0, W^*]$, because either the agency cost of preventing risk-taking is too high or exposing the agent to Poisson risk violates limited liability constraint, the principal implements the risky action $a_t^s = a_R$ and the agent's continuation value follows $dW_t = \rho W_t dt + \beta(dX_t - (\mu + \alpha)dt)$. On this interval, no cash payment is made, the principal only rewards or punishes the agent by changing the continuation promises. Termination of the contract occurs when the continuation value W_t hits 0 for the first time. In which case the principal receives l and the agent gets nothing. When $W_t \in (W^*, W^p]$, because the firm stands far away from the liquidation boundary and it is cheaper for the principal to let the agent bearing Poisson risks, she implements the safe action $a_t^s = a_N$ and the agent's continuation value follows $dW_t = (\rho W_t + \psi\lambda)dt + \beta(dX_t - \mu dt) - \psi dN_t$. The principal imposes a severe punishment when she observes a large loss and compensates the agent with cash when $W_t \geq W^p$.

From a computational perspective, an alternative way to understand the solution F to (**??**) is the following: On $[0, W^*]$, the value function is given by the solution to (**??**), $F_R(W)$. The risky-action ODE is solved by imposing boundary conditions $F_R(0) = l$ and $F_R(W^*) = F_N(W^*)$. On $[W^*, W^p]$, the value function is given by the solution to (**??**), $F_N(W)$. The safe-action DDE is solved by specifying an initial history $F_R(W)$ on $[W^* - \psi, W^*]$, together with the conditions on payment boundary $F'_N(W^p) = -1$ and $F''_N(W^p) = 0$. The solution to the HJB-equation $F(W)$ on $[0, W^p]$ is thus defined by

$$F(W) = \begin{cases} F_R(W) & \text{if } W \in [0, W^*] \\ F_N(W) & \text{if } W \in (W^*, W^p] \end{cases} \tag{1.3.11}$$

Lastly, notice that, depending on parameter values, it is possible to have an optimal contract that implements the risky action all the time, that is, $W^* = W^p$. This happens when the agency cost of risk-taking is sufficiently high or the expected monetary loss is sufficiently low. At the extreme when risk-taking does not destroy values, $\alpha - L\gamma = 0$, risk-taking is always optimal. Because the contract with an interior switching point $W^* < W^p$ is the most interesting case, in the rest of the paper, I assume that $\alpha - L\gamma < 0$ and the following lemma pins down conditions that deliver $W^* < W^p$. Although the conditions are not entirely stated in terms of the model's primitives, it offers additional insights to the contracting problem. To state the lemma, let $F_{N,m}$ denote the solution to (**??**) with initial history $F_{N,m}(W) = \frac{m}{\psi}W + l$, where m is the maximal initial slope such that $F_{N,m}$ satisfies the relevant boundary conditions.

Lemma 1.3. *Let W^p be the payment boundary under the optimal contract. Suppose that $F_{N,m}(W^p) + W^p > \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$, then $W^* < W^p$. For $W^* > \psi$, a necessary condition is $F_{N,m}(\psi) + \psi \leq \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$.*

The first part of the lemma provides a sufficient condition for the safe-action to be optimal. The quantity $F_{N,m}(W)+W$ is the social value of a contract that restricts attention to the safe-action. When the social value of the safe-action is strictly larger than firm's expected discounted cash flows under risk-taking, it will be efficient to switch from the risky action to the safe one. The second part of the lemma states the similar, the social value of implementing the safe action alone is smaller than the firm's expected discounted cash flows under risk-taking, it will be *potentially* efficient for the principal to assign the risky-action. The contrapositive of the second part suggests that if $F_{N,m}(\psi) + \psi > \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$, then under the optimal contract, it must be that $W^* = \psi$.

1.3.3 Discussion.

This subsection provides a short discussion of the optimal contract. I first discuss the economic intuition behind the optimal contract result. Then I will compare my model to DS and BMRV.

In the optimal contract, in order to induce the agent to work hard, the principal sets $\beta = \frac{B}{\mu}$. The incentive component admits standard interpretation: β is proportional to the effort dimension of moral hazard. A higher private benefit B, the agent is tempted to shirk, hence a higher-powered incentive is required. For a higher marginal effect of effort on cash flows μ, the principal can detect shirking easily, hence a smaller incentive loading. On the other side, high-powered incentive also drives the agent to take risks, as by doing so, she will earn $\beta\alpha$. So to prevent the agent from gambling, incentive compatibility requires an upper bound on β, which gives the constraint $\psi \geq \beta\frac{\alpha}{\gamma}$, so that the punishment is proportional to working incentives and the agent is exposed to the Poisson risk. The stick ψ also admits standard interpretation: A higher β implies a heavier punishment is needed to deter risk-taking. As γ increases, losses come more frequently and it is easier for the principal to detect risk-taking, hence ψ will be smaller. An increase in α tempts the agent to risk-shift, therefore, more severe punishment is needed.

On the equilibrium path, the principal would like to implement the risky action when the agent performs poorly, that is, when the agent's continuation value is low. This is in sharp contrast to the first-best in which only the safe action is implemented because by assumption 2 the risky action is value-reducing. To understand this result, notice that by proposition 1.1, to motivate the agent to choose the risky action, it suffices for the principal to set the incentive loadings to be $\psi_t \gamma \leq \beta_t \alpha$. By assumption 1, liquidation of the firm is inefficient and any increase in the volatility of the agent's continuation value process is costly to the principal, since it increases the chance for W_t to hit the liquidation boundary. Hence, in the optimum $\psi_t = 0$ when the principal recommends the risky action. Although risk-taking is costly to the principal as it destroys values, interestingly, from the principal's point of view, the benefit of implementing the risk-taking behavior is a reduction in volatility of the agent's continuation value, hence a smaller chance of inefficient liquidation. In other words, to rule out hidden gambling, the principal has to bear additional agency cost $|\mathcal{A}(W)|$ as defined in (**??**). The optimal action policy is obtained by weighting the expected monetary loss $|\alpha - L\gamma|$ of risk-taking and the agency cost of preventing risk-taking $|\mathcal{A}(W)|$. The logic of volatility effect implies $|\mathcal{A}(W)|$ is strictly decreasing in W: as the agent's performance accumulates, a downward jump in the agent's continuation value no longer brings the firm sufficiently close to the liquidation boundary. Hence the principal seeks to implement the safe action at the top and the risky action at the bottom.

These observations are closely related to "gambling for resurrection"-type behavior. Optimal contracting rationalizes this type of behavior. What is interesting is that although risk-taking destroys values of the firm, from the principal's point of view, hidden gambling at the bottom is constrained efficient (whenever $W^* > \psi$), and on the equilibrium path, the principal optimally allows the agent to take risky gambles by imposing suitable incentive loadings, and thus letting the agent to "resurrect" her from liquidation.[11]

1.3.3.1 Comparison to DS and BMRV

In DS, the cash flows generated by the firm follow an arithmetic Brownian motion and the principal implements high effort at all times. The optimal incentive structure is similar: to motive the agent, the principal has to expose the agent to a certain amount of risk (the volatility β). In my setting, since moral hazard is two dimensional, there are two sets of incentive constraints facing the agent. However, if the principal finds it optimal to implement the risky action forever, by proposition 1.2 and 1.3, the dynamics of the continuation value and the principal's profit exhibit the same structure as in DS, with expected cash flows $\mu + \alpha - L(\lambda + \gamma)$ replacing their drift term μ.

In BMRV, similar to my model, the firm's cash flows suffer negative shocks modeled by a Poisson process and the agent is also protected by limited liability. More formally, in BMRV, the firm generates cash flows at any moment of time according to

$$K_t dY_t = (\mu dt - L dN_t) K_t$$

where K_t is the firm's size or capital stock. Assume there is no investment but costless partial liquidation is possible. That is, let $\kappa_t \in [0, 1]$ denote the fraction of capital stock to liquidate at t. When the agent shirks, the intensity of the Poisson process increases from λ to $\lambda + \Delta\lambda$, but the agent accrues private benefits at rate BK_t. The agent's action does not affect the cash flow drift. Suppose that the principal finds it optimal to implement the high effort, then a result similar to proposition 1.1 applies, and the agent's continuation value follows

$$dW_t = \rho W_t dt - dI_t - \Psi_t(dN_t - \lambda dt)$$

[11] Notice that the principal is assumed to have deep pockets. Thus she can absorb all the losses generated by hidden risk-taking and all she cares is liquidation. It would also be interesting to study what will happen when the principal is credit-constrained.

where the sensitivity to Possion risk Ψ_t satisfies an incentive constraint: $\psi_t \geq \frac{BK_t}{\Delta\lambda}$. In BMRV, when the agent's continuation value is sufficiently low, incentive can be provided by partially liquidating the firm's capital. In particular, when continuation value is such that $\frac{BK_t}{\Delta\lambda} > W_{t-} - \Psi_t$, downsizing at time t, i.e., $\kappa_t < 1$, will be necessary to provide incentives. This is because $W_{t-} - \Psi_t$ is the continuation value of the agent after a loss occurs at t. And $\frac{BK_t}{\Delta\lambda}$ is the minimal required punishment when another loss occurs sufficiently close to time t. The strict inequality implies limited liability will be violated following two consecutive losses.[12] Therefore, in terms of size-adjusted continuation value $w_t = \frac{W_{t-}}{K_t}$ and punishment $\psi_t = \frac{\Psi_t}{K_t}$, and note that the incentive constraint is binding in equilibrium: $\psi = \frac{B}{\Delta\lambda}$, the principal partially liquidates the firm when $w_t \in (\frac{B}{\Delta\lambda}, \frac{2B}{\Delta\lambda})$. Downsizing leads to a reduction in the agent's private benefit of shirking and thus a smaller punishment is needed to incentivize the agent. In other words, downsizing could be viewed as an alternative way to provide incentives.

The key differences between BMRV and my model are that (i) I do not consider capital stock and partial liquidation. (ii) risk-prevention is costly in BMRV while in my model, risk-taking is a strategy which is costless.[13] Incentives can be provided by implementing a different risk-action. As the choice of risky action does not require punishment on the equilibrium path, the conflicts between incentive constraints and limited liability can be resolved by switching actions.

A natural question arises here: does the introduction of capital stock and downsizing relaxes the contract space and help eliminate risk-shifting problems in my setting? Suppose in the current setting the cash flows follow

$$K_t dY_t = K_t(\mu(a_t)dt + \sigma dZ_t - LdN_t)$$

[12]Formally, let $K_{t+} = \lim_{s \downarrow t} K_s \in [0, K_t]$ and $K_{t+} = \kappa_t K_t$. Limited liability requires at all t, $W_{t-} - \psi_t \geq \frac{BK_{t+}}{\Delta\lambda}$. If the strict inequality is true, then $K_t > K_{t+}$, which implies $\kappa_t < 1$.

[13]As long as risk-taking increases the drift of the cash flows, introducing costly risk prevention does not change the main result. To see this, suppose it costs the agent c_N to avoid risk-taking, i.e., take the strategy $a_t^s = a_N$. The required incentive constraint changes to $\psi_t \gamma \geq \beta_t \alpha - c_N$. The constraint is relaxed by a constant but a fixed-sized punishment is still needed to deter risk-taking. Therefore the key features of the optimal contract would be the same, but potentially with a different switch point W^*.

so that the operating profit is proportional to the size of capital stock K_t and partial liquidation is allowed. Also, assume that the private benefit is proportional to capital stock, BK_t. Applying the martingale approach, the required incentive constraints for $a_t = (a_H, a_N)$ are $\mu\beta_t K_t \geq BK_t$ and $\Psi_t\gamma \geq \beta_t K_t \alpha$. Thus to induce the agent to choose the safe action in a larger firm, a heavier punishment is required. This suggests that the logic of BMRV applies: whenever $\frac{B\alpha}{\mu\gamma}K_t > W_{t-} - \psi_t$, there is downsizing $\kappa_t < 1$. Letting $b = \frac{B}{\mu}\frac{\alpha}{\gamma}$.[14] In terms of size-adjusted variables, $\kappa_t = \frac{w_t - b}{b} < 1$ whenever $w_t \in (b, 2b)$, *if the principal implements the safe action on this interval.* However, downsizing needs not occur on the equilibrium path in my model. Both partial liquidation and switching to the risky action lead to reduction in the firm's value. The contractual term to be use depends on the relative losses of the options. The equilibrium could be of two types: (i) The optimal size-adjusted switching point $w^* > 2b$, or (ii) $w^* \in (b, 2b)$. In the former equilibrium where $w^* > 2b$. At $w_t \leq w^*$, the principal implements the risky action and hence punishment becomes irrelevant. Even though $w_t \leq 2b$, downsizing does not occur. In the latter one, for continuation value $w_t \in [w^*, 2b)$, the principal implements the safe action and downsizing occurs before switching to the risky action. Given the role of downsizing is to adjust incentives, optimal switching between actions replaces partial liquidation of capital stock. Therefore, downsizing and action-switching are two different ways to provide continuation incentives.

1.4 Implication and Analysis

In this section, I discuss some implications of the model. First, I provide some numerical analysis of the optimal contract. Then I turn to study the implementation of the optimal contract using a capital structure. Lastly, I show that the incentive structure of the optimal contract is the same as the high-water mark contracts and address the risk-shifting problems.

[14] Note that $\frac{B}{\mu}\frac{\alpha}{\gamma}$ is analogous to $\frac{B}{\Delta\lambda}$ in BMRV.

1.4.1 Quantitative Analysis.

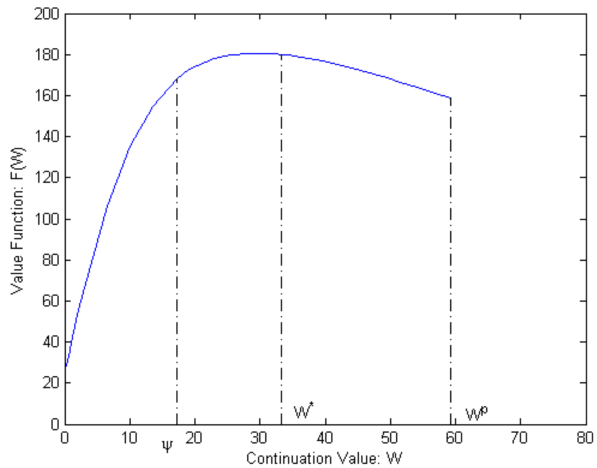

Figure 1.4.1: **The Principal's Value Function.**
Parameter values: $r = 0.1$, $\rho = 0.15$, $\mu = 35$, $\alpha = 20$, $L = 50$, $\lambda = 0.2$, $\gamma = 0.5$, $B = 15$, $\sigma = 16$, $l = 25$.

Figure 1.4.1 provides an example of the optimal contract. The blue line is the principal's value as a function of the agent's continuation utility, which is smooth and strictly concave on $[0, W^p)$, with a payment boundary is $W^p = 59.28$. In this contract, the incentive loadings are $\beta = 0.43$ and $\psi = 17.14$. The punishment is imposed when the principal implements the risky action. Whenever any loss events occur, the punishment causes a relatively large downward jump of the continuation value. The optimal action policy is characterized by an interior switching point $W^* = 33.31$. To the left of W^*, risk-taking is the second-best action and to the right of it, the safe action is optimal. This example is particularly interesting in the sense that the region of continuation value such that the principal assigns the risky action optimally is non-degenerate. To compare the numerical result to the theoretical analysis in the previous section, first notice that the expected monetary loss of risk-taking is $|\alpha - L\gamma| = 5$, which is 20% of the expected cash flows of the firm under the safe action $\mu - L\lambda = 25$. The percentage loss is reasonable and acceptable.

Now, consider the firm is currently sitting at continuation value $W = 25$ and suppose that the principal implements the safe action at this point. Whenever there is any large loss, incentive-compatible punishment brings the agent's continuation value down to $W - \psi \approx 8$, which is sufficiently close to the liquidation boundary. In fact, the numerical result shows that the agency cost of preventing risk-taking is $|\mathcal{A}(W = 25)| = 7.6$, which is 30% of the firm's expected profit. Therefore, the principal chooses to implement the risky action at W=25.

Now I perform comparative statics by adjusting the key parameters related to the expected monetary loss of risk-taking (α, L, γ). The results are reported in the following table. The parameter values used to compute the principal's value in figure 1.4.1 are served as the baseline parameters.

parameter	value	β	ψ	W^*	W^p	$\frac{W^*}{W^p}$
Baseline		0.43	17.14	33.31	59.28	0.56
α	15	0.43	12.86	12.86	54.82	0.23
	18	0.43	15.43	24.35	58.63	0.41
	22	0.43	18.86	36.36	59.16	0.61
γ	0.45	0.43	19.05	39.10	57.01	0.69
	0.6	0.43	14.29	21.29	57.53	0.37
	0.7	0.43	12.25	12.25	53.80	0.23
L	45	0.43	17.14	37.57	54.78	0.69
	55	0.43	17.14	26.43	61.43	0.43
	60	0.43	17.14	17.14	62.37	0.27

Table 1.1: **Comparative Statics of α, γ, L.**
The baseline parameter values: $r = 0.1$, $\rho = 0.15$, $\mu = 35$, $\alpha = 20$, $L = 50$, $\lambda = 0.2$, $\gamma = 0.5$, $B = 15$, $\sigma = 16$, $l = 25$.

In the table, the effort incentive remains unchanged because β is independent of these parameters. The general pattern revealed in the table is consistent with the theoretical analysis: (i) a higher α and a lower γ imply a more severe punishment. (ii) As the expected monetary loss of risk-taking increases (increase in γ, L and decrease in α), the optimal switching point W^* decreases. The changes in the switching point are much more sensitive to the change in payment boundary. The last column of the table computes the ratio $\frac{W^*}{W^p}$. The drastic change in the ratio suggests that the principal implements the safe project more often as the loss of risk-taking magnifies. In the extreme where $\alpha = 15$, $L = 60$, or $\gamma = 0.7$, the switching point W^* coincides with ψ, meaning that the expected monetary loss is so large that risk-taking is no longer optimal, yet it is still necessary for incentive purposes.[15]

parameter	value	β	ψ	W^*	W^p	$\frac{W^*}{W^p}$
Baseline		0.43	17.14	33.31	59.28	0.56
μ	30	0.5	20	33.34	65.94	0.50
	40	0.38	15	29.13	54.23	0.53
	45	0.33	13.33	27.32	49.78	0.55
l	0	0.43	17.14	32.82	60.02	0.54
	50	0.43	17.14	29.20	58.94	0.50
	100	0.43	17.14	25.60	56.86	0.45
σ	12	0.43	17.14	29.87	53.62	0.56
	20	0.43	17.14	33.18	65.47	0.51
	24	0.43	17.14	32.52	71.56	0.45
B	10	0.29	11.43	21.71	42.89	0.51
	20	0.57	22.86	43.44	75.24	0.58
	25	0.71	28.58	50.27	89.79	0.56

Table 1.2: **Comparative Statics of** μ, l, σ, B.
The baseline parameter values: $r = 0.1$, $\rho = 0.15$, $\mu = 35$, $\alpha = 20$, $L = 50$, $\lambda = 0.2$, $\gamma = 0.5$, $B = 15$, $\sigma = 16$, $l = 25$.

Table 1.2 considers comparative statics with respect to some other parameters. Concerning risk-taking, the table shows that the ratio $\frac{W^*}{W_p}$ is relatively insensitive to the change in these parameters which have no direct effect on the expected monetary loss of risk-taking. The comparative statics also confirms the intuition obtained from the theoretical analysis: a larger liquidation value implies liquidation is less inefficient, hence the agency cost of preventing risk-taking is less severe and thus the principal could deter risk-taking more often. An increase in cash flows volatility pushes the payment boundary to the right. This

[15] Although not reported in the table, when $\alpha = 25$, $\gamma = 0.8$, or $L = 40$, the expected monetary loss is 0. The numerical results show that the optimal policy is a_R for all $W \in [0, W^p]$.

is because higher volatility implies a higher chance for the firm to liquidate, in order to keep the firm to stay away from liquidation, a higher payment boundary is required. An increase in the private benefit of shirking B implies a more high-powered incentive scheme is needed to motivate the agent. This implies the agent is exposed to the Brownian risk more heavily. To prevent inefficient liquidation, the payment boundary moves up a lot. Meanwhile, a high β makes hidden gambling more tempting and a heavier punishment is needed to deter risk-taking. This leads to a higher agency cost in preventing risk-taking and thus pushes the optimal switching point W^* by a substantial amount in order to prevent inefficient liquidation when a punishment takes place.

1.4.2 Capital structure implementation.

In this subsection, I show how the optimal contract derived in section 3 can be implemented using standard securities (e.g., debt and equity) and an insurance contract.[16] To implement the contract, first recall that the optimal contract exhibits memory and its dynamics depends on the evolution of the agent's continuation value W_t. The state variable W_t can be positively linked to another variable that measures "financial slackness" of the firm. Financial slackness may correspond to the firm's cash reserves as in BMPR, or a credit line as in DS, or a combination of both. I choose to interpret the financial slackness as cash balance in this paper. Second, the insurance contract is relevant in the present set up in the sense that it allows the firm to hedge against large risks. It will be seen below that the terms of the insurance contract also affect the agent's incentive to take risks.

Specifically, I consider the following securities in the implementation.

***Cash Reserve*:** The firm maintains a level of cash reserves M_t at time t to meet any needs for short-term financing. The cash balances are deposited on a bank account and earn the market interest rate r. Daily operating cash flows are injected into the account. Any coupon, dividend payments, or insurance premium paid to the outside investors are

[16] In general, there are many different ways to implement the optimal contract. Nevertheless, the securities designed in this section is intuitive and suggestive.

drawn from this account. Any change in the level of cash reserves thus reflects the firm's performance and the occurrence of loss events.

***Equity*:** Equity is issued in the form of *common stocks*, in which equity holders receive dividend payments dD_t. The agent holds a fraction $\beta = \frac{B}{\mu} \in (0,1)$ of common stocks and is prohibited from selling the stocks.[17] By compensating the agent with a proportion of firm's equity, she has the appropriate incentive to work hard.

***Debt*:** The long-term debt is a performance pricing bond with no maturity that pays coupon at rate c_t. In particular, consider the coupon payment:

$$c_t = \begin{cases} \mu + \alpha - L(\lambda + \gamma) - (\rho - r)M_t & \text{if } M_t \leq M^* \\ \mu - L\lambda - (\rho - r)M_t & \text{if } M_t > M^* \end{cases}$$

where M^* is a cash threshold to be specified (and is endogenously determined) later. The first component of the coupon is positively related to the firm's expected cash flows and the second component $(\rho - r)M_t$ reflects the "performance pricing" component. The size of the coupon payment depends on the firm's current performance. When the cash balance is low so that $M_t \leq M^*$, the firm is in distress and the coupon is "downsized" by an amount $|\alpha - L\gamma|$.

***Insurance Contract*:** There is an insurance company in a competitive market that provides insurance services to the firm. The contract payout at rate dR_t and is fairly priced. In particular, consider:

$$dR_t = \begin{cases} (\lambda + \gamma)Ldt - LdN_t & \text{if } M_t \leq M^* \\ \lambda(L - \frac{\alpha}{\gamma})dt - (L - \frac{\alpha}{\gamma})dN_t & \text{if } M_t > M^* \end{cases}$$

The insurance company is liable for the firm's large risks. Whenever there is a large loss ($dN_t = 1$), the insurance contract delivers contractual payments that allow the firm either

[17] Recall by assumption 3, $\mu > B$.

to cover the entire loss L or to obtain partial coverage $L - \frac{\alpha}{\gamma}$.[18] As the insurance contract is fairly priced, the actuarially fair premium is $(\lambda + \gamma)L$ when $M_t \leq M^*$ and is $\lambda(L - \frac{\alpha}{\gamma})$ when $M_t > M^*$. This is true because the insurance company has rational expectation, it correctly anticipates that the firm's asset risk and the manager's project choice, in which case a loss occurs with probability $(\lambda+\gamma)dt$ during the time interval $[t, t+dt)$. Similarly, if the insurance company expects the firm's asset risk to be low, the premium is proportional to the probability of loss λdt for $M_t > M^*$. As the insurance premium is fair, $E_t(dR_r) = 0$.

In the implementation, the agent has discretion over the firm's investment policy, dividend payout policy, and when to accumulate cash balances. In addition to the choice of efforts, the agent is delegated by the principal to choose over safe or risky projects. Under this security design prescription, the cash reserves grow according to

$$dM_t = rM_t dt + dY_t - dD_t - c_t dt - dR_t \tag{1.4.1}$$

Proposition 1.4. *Consider the capital structure specified above and suppose the cash balances satisfying $M_t = \frac{1}{\beta}W_t$ and $M^* = \frac{1}{\beta}W^*$. Then it is incentive compatible for the agent to provide high efforts all the time, choose the risky action when $M_t \leq M^*$ and the safe action when $M_t > M^*$. The agent distributes dividends when M_t hits $M^p = \frac{1}{\beta}W^p$ and terminates the firm when M_t hits 0 for the first time.*

The incentive properties of the implementation can be understood as follows. First, the agent has incentives to work hard because she is compensated by a fraction of the firm's equity. The fraction is exactly equal to the incentive loading on the Brownian motion, $\beta = \frac{B}{\mu}$. Thus when dividends are paid, the agent consumes $dI_t = \beta dD_t$. Second, to ensure the agent does not pay dividends too early or too late, I set $M^p = \frac{1}{\beta}W^p$. The logic behind this equality is the same as DS's analysis of the length of the credit line. Suppose the current cash balance is M_t, and the agent's continuation utility is W_t. The agent can always pay a total dividend $dD_t = M_t$ and terminate the business, in which case the agent

[18] Recall by assumption 2, $\frac{\alpha}{\gamma} < L$.

obtains $\beta dD_t = \beta M_t$. However, as we set $M_t = \frac{1}{\beta}W_t$, the agent's payoff from this action is W_t. So the agent has no incentive to deviate.

Turning to the use of insurance. The insurance contract is structured in a way to provide the appropriate incentive for the agent to choose the optimal risk-related action. In particular, when the firm is performing well such that $M_t > M^*$, the investors agree to adopt incomplete hedging. That is, when there is a loss of size L, the insurance company only covers part of the loss of size $L - \frac{\alpha}{\gamma}$. In other words, the firm absorbs some of the loss, which is of size $\frac{\alpha}{\gamma}$. This reduces the cash balance, moves the firm away from the dividend payout boundary, and the agent's consumption will be delayed. Hence, the partial coverage serves as the punishment in the optimal contract. As a result, by setting $M^* = \frac{1}{\beta}W^*$, the insurance contract provides appropriate incentives for the agent to choose the optimal project. Suppose $M^* > \frac{1}{\beta}W^*$, and the current state of the firm is $M_t \in (\frac{1}{\beta}W^*, M^*)$. The optimal contract requires the agent to take the safe project. But the insurance contract provides full coverage on this interval and hence the agent will deviate to take risks as she is protected against any downside risk. Similarly, if $M^* < \frac{1}{\beta}W^*$ and the current cash balance is $M_t \in (M^*, \frac{1}{\beta}W^*)$. The optimal contract assigns the risky action, but the insurance contract provides partial coverage. Hence the agent would think that there is not enough protection against the risky project and deviate to take the safe project. As a result, by setting $M^* = \frac{1}{\beta}W^*$, the agent has the right incentive to adopt the optimal project allocation[19].

The analysis of the incentive property of the insurance contract implies that risk-management cannot be disentangled from project selection. Different projects carry different amount of risks. While project selection is delegated to the agent, suitable incentives have to be provided to the agent in order for the selection to be constrained efficient. Hence, incomplete hedging is driven by the agency conflicts with respect to risk-taking. While a

[19] In a working paper version of BMRV, they also consider implementation using an insurance contract in an environment with Poisson losses. However, the insurance coverage and the degree of hedging are the same for all histories because there is only a single task in their model and they do not allow any form of switching.

number of dynamic corporate finance models yield prescription on risk-management policies, typically they treat the firm as having a single project.[20] The security design exercise suggests that dynamic risk-management and project selection should be considered altogether.

As M_t measures the firm performance, the variable can be linked to the firm's credit rating. In fact, the threshold M^* slices the performance space into two halves. A natural assignment of credit rating is: when the firm performs poorly, $M_t \leq M^*$, the firm receives a speculative grade; and when the firm is performing well, $M_t > M^*$, the firm gets an investment grade. The coupon payment thus depends on the firm's credit rating and performance. As the firm is downgraded, there is a large reduction in coupon payments.

Summarizing the discussion, the model delivers a number of testable implications.

1. An investment grade firm generates higher expected cash flows.
2. A speculative grade firm invests more in risky projects, and has higher cash flows volatility.
3. A firm with better credit quality tends to adopt incomplete hedging.
4. A larger firm pays more coupons and pays dividends more often.
5. A firm with a manager who can generate a higher α exhibits more volatile cash flows.

Security pricing.

Given the capital structure implementation, I now consider the market value of these securities. Assume that in the case of default, the priority structure is such that the debt holders acquires all the liquidation value l. Using the specified security payments and

[20]See, for examples, Bolton, Chen, Wang (2011) and Rampini and Viswanathan (2010). The former considers agency cost of free cash flows, and the financial frictions come from limited enforcement in the latter paper. Both models predict that a more financially constrained firm hedges less.

equation (**??**), the dynamics of cash balances reduces to

$$dM_t = \rho M_t dt - dD_t + \sigma dZ_t - 1_{\{M_t > M^*\}}(\frac{\psi}{\beta}(dN_t - \lambda dt))$$

Notice that in this alternative representation, the downward adjustment on cash balances remains on $[M^*, M^p]$ because there is imperfect hedging on this region.

The market value of securities are defined as follows:

Stock Price

$$\mathcal{S}_t = E_t\left[\int_t^\tau e^{-r(s-t)} dD_s\right]$$

Bond price

$$\mathcal{B}_t = E_t\left[\int_t^\tau e^{-r(s-t)} c_s ds + e^{-r\tau} l\right]$$

Let $\mathcal{I}(M_t)$ be the market value of the insurance policy. It is easy to see that for any t and M_t,

$$F(\beta M_t) + M_t = (1-\beta)\mathcal{S}(M_t) + \mathcal{B}(M_t) + \mathcal{I}(M_t)$$

The left-hand side is the present value of the firm's underlying cashflow-generating asset plus the stock of cash. The right-hand side is the market value of the securities held by outside investors. Equity holders hold a fraction $1-\beta$ of common stocks and each share has a stock price $\mathcal{S}(M_t)$. Bondholders value the debt at price $\mathcal{B}(M_t)$ and the insurance contract has a value $\mathcal{I}(M_t) = 0$. Given the dynamics of the cash balance, the stock price and bond price can be computed easily by solving the appropriate differential equations.

Remark. In the implementation described above, the continuation value of the agent is linked to the firm's cash balance. Given the existence of an endogenously determined optimal switching point, another natural interpretation of financial slack is a combination of the firm's cash reserve and available credit. For example, as in DS, suppose the credit line limit is C^L, which satisfies $C^L = \frac{1}{\beta}W^*$. Any outstanding balance on the credit line is charged an interest rate r. With a suitable modification of the capital structure, one can see that the firm starts to accumulate cash after repaying all the outstanding balance on the credit line. In other words, when the internal cash reserve is exhausted, the firm turns to external funding and starts drawing down the credit line in order to support risky ventures. The firm is forced to terminate once the credit limit is reached. Note that this interpretation of financial slack is consistent with the dynamic pecking order theory.[21]

1.4.3 Compensation in Hedge Fund Industry.

Hedge funds nowadays manage $2.13 trillion of assets. Although hedge funds are set up as limited partnership, the fund management (the general partner) essentially acts as an agent who manages the fund's operation and makes use of various investment strategies to invest the assets on behalf of their principal. Hedge fund management compensation contracts typically include both management fees and performance-based incentive fees. The management fee is charged as a fixed fraction of the fund's asset under management, with values ranging from 1% to 4% per annum. The performance fee is a fraction of the profits made by the fund over a year, with values ranging from 15% to 50%. The combinations that are often observed in practice is 2-20 contract. The performance fee is meant to incentivize the fund manager and typically includes a "high-water mark" provision. The high-water mark keeps track of the historical maximum of the invested capital in the fund. Performance fee will only be paid if the current asset size exceeds the high-water mark. In other words, the fund manager has to recoup any cumulative losses over the previous year before she is paid the incentive fee. One of the key issues in the study of hedge fund compensation is whether risk-taking is induced by the high-water mark contract, as the implied payoff to the manager is convex.[22]

As observed by Biais *et al.* in an arithmetic Brownian motion framework, with a suitable capital structure implementation, the cumulative dividends admit a representation that

[21] For an empirical analysis of credit lines, see Sufi (2009).

[22] In reality, risk-taking is difficult to observe. The fund management investment strategies are typically very complex as they involve long-short positions, the use of leverage and derivative contracts to hedge against market risks, and targeting of macro or political events. Therefore, the strategies are costly to communicate to the investors. This justifies unobserved risk-taking in my model.

allows us to interpret the optimal compensation contract as a high-water mark contract. In my model, the agent's continuation value display similar dynamics as in Biais *et al.* (2013) except for a downward jump. A natural question to ask is: does the optimal contract derived in section 3 also admits such a representation? The answer turns out to be yes and the result is stated in the proposition below. To begin with, I maintain the capital structure stated in proposition 1.4.

Proposition 1.5. *At any time $t < \tau$, the cumulative dividends satisfy*

$$D_t = \sup_{s \in [0,t]} \{\Phi_s\} \tag{1.4.2}$$

where $\Phi_s = \max\{M_0 + \int_0^s ((rM_u - c_u)dt - dR_u) + Y_s - M^p, 0\}$.

To interpret equation (**??**), notice that $M_0 + \int_0^s ((rM_u - c_u)du - dR_u) + Y_s$ measures the firm's cumulative performance at time s, net of contractual payments to outsiders and before common stock dividends are distributed. This quantity also measure the fund size in a hedge fund. Once the cumulative performance is higher than the payout boundary M^p, the agent can issue dividends and is compensated, and below which, no bonus payments are made. In the latter case, $dD_t = 0$. After any dividends payout, a new round of dividends will only be paid when the cumulative performance reaches a new running maximum. In particular, when the agent is performing well, dividends will be delivered in which case she will be paid by a fraction β of it. On the other hand, when bad performance accumulates, the agent has to make up for earlier losses before any compensation takes place again. This compensation contract resembles high-water mark contracts with a performance fee equal to $\beta \in (0,1)$.[23]

High-water mark contracts possess an interesting option-like feature. One can think of the current running maximum as the strike price of an option. When fund return exceeds the running maximum, the fund manager can "exercise" the option and claim a

[23] Note that there is no management fee implied by the model. However, as suggested in the last paragraph, it can be easily accommodated by including a fixed cost of efforts.

bonus payment. When the fund performance is below the past maximum, the option is out-of-the-money and the manager cannot get paid. It is well-known that such a convex compensation structure will induce risk-shifting problems in the sense that the manager will be more willing to pursue risky projects or hold risky assets. The theoretical results derived in proposition 2 suggest that high-water mark contracts do not completely eliminate risk-shifting problem. There will still be a certain degree of risk-taking. However, the propositions highlight the fact that the amount of risks taken by the fund manager will critically depend on the current fund performance. Specifically, when fund performance is bad, the high-water mark contract induces more risk-taking. Intuitively, when the fund performs poorly, the option is far out-of-the money. This implies that the manager will have a stronger incentive to gamble in order to reach the strike price earlier. In contrast, if the fund performance is good, the option is closer to the exercise boundary. Taking risk means a higher chance to move away from this boundary and therefore, the manager does not want to gamble in this case. The analysis suggests that in a dynamic context, convex payoff structure does not necessarily induce hidden risk-taking. In particular, "distance-to-threshold", that is, the distance of current fund size to the high-water mark is the key to understand whether the manager will engage in risk-shifting.

In a recent paper, Shelef (2013) empirically studies the impact of incentive contracts on risk-taking and performance in hedge fund industry. By merging data from Lipper-TASS and Hedge Fund Research, he obtains a data set that includes month assets and returns of about 9000 hedge funds from 1994 to 2006. He then estimates the effects of "distance-to-threshold" on the fund returns and volatility. In particular, his estimation provides causal evidence that a fund that is 15% below its high-water mark, the expected returns over the next year is reduced by 2.1% and the riskiness of the fund is increased by about 50%. The new empirical evidence Shelef provides is consistent with the prediction of my model.[24][25]

[24]Shelef (2013) also documents that a hedge fund which is 50% below the incentive threshold takes few risks and exhibits much lower volatility. Based on the analysis of Zhu (2012), I conjecture that this empirical evidence can be captured in my model by allowing the agent to shirk at the bottom. This is because to implement shirking, $\beta_t = 0$ and the agent will have no strict incentive to take hidden risks.

[25]Kolokolova and Mattes (2013) studies 714 hedge funds over the period 2001-2011. They also find

Panageas and Westerfield (2009) address the question of whether high-water mark contracts would induce risk-taking in a portfolio choice setting. They find that the a risk-neutral manager who is compensated by a high-water mark will choose a portfolio with bounded volatility. To understand their result, one can imagine that if a manager chooses a more riskier portfolio, it would increase the chance for the fund performance to cross the high-water mark, which is beneficial to the manager. Meanwhile, it also increases the probability that the fund value will decrease tomorrow. The latter effect reduces the value of the option. Trading-off these effects, the manager will only take a bounded position in risky assets and behave as if he is an investor with constant relative risk aversion. In fact the optimal portfolio is a fixed mean-variance portfolio. Although Panages and Westerfield show that a fund manager will not choose unboundedly large position in risky assets, the contract in their analysis is given. My results suggest that (i) high-water mark contract is indeed an optimal contract. And (ii) the position in risky assets will change depending on the current fund performance.

1.5 Extensions

In this section, I consider further extensions of the model.

1.5.1 Replacing the agent.

In the baseline model, the principal rationally expects the agent to take inefficient downside risk when her continuation value. However the principal would optimally allow her to do so because the agent's skill is essential for the firm. The question I consider here is: can the principal deter risk-taking by threatening to fire the current manager and replace her with a new manager? Suppose the principal can fire the agent and replace her at a cost c_a . This cost captures, for example, searching and negotiation cost with the new agent. The principal initiates another optimal contract with the new agent from a competitive

that fund close enough to the high-water mark exhibits no risk-shifting and fund substantially below the threshold has a significant increase in risks.

market. With the possibility of replacement, the principal's liquidation value $l_R(W)$ is endogenously determined

$$l_R(W) = \max_{\tilde{W}} F(\tilde{W}) - c_a - W$$

where $\max_{\tilde{W}} F(\tilde{W})$ is the value of a new contract, c_a is the replacement cost, and W is the severance pay made to the fired agent when she is fired at point W. The optimal contract with the replacement option depends on the replacement cost.

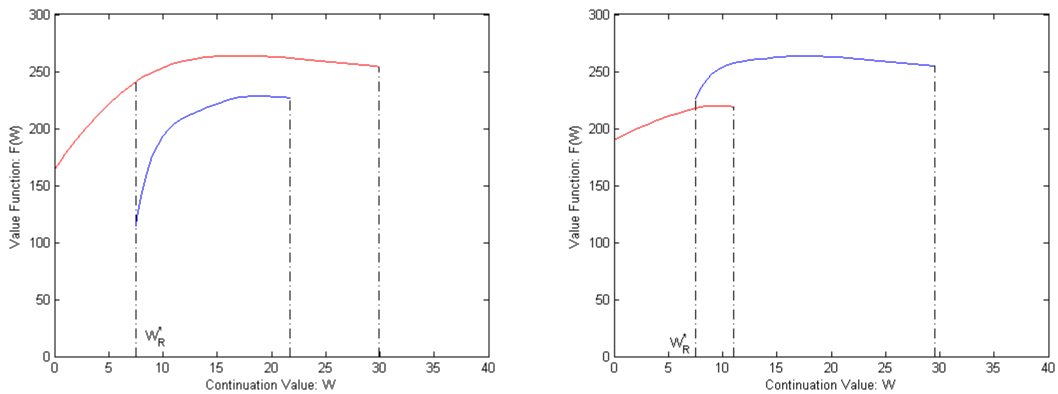

Figure 1.5.1: **Replacing the agent**.
When the principal can fire and replace the agent at a cost c_a, liquidation value is determined endogenously. Parameter: $c_a = 100$ (left panel) $c_a = 30$ (right panel). $r = 0.1$, $\rho = 0.5$, $\mu = 40$, $\alpha = 15$, $L = 50$, $\lambda = 0.2$, $\gamma = 0.5$, $B = 10$, $\sigma = 12$, $l = 25$.

In figure 5.1, the pictures compare two options available to the principal: (i) allowing the agent to take risk and fire the agent when continuation value hits 0, and (ii) do not allow the agent to take risk and fire the agent when the limited liability starts to bind. The profit from the first option is represented by the red lin in both figures and the profit from the second option is represented by the blue line. On the left panel of figure 5.1, the replacement cost, $c_a = 100$, is the same as the expected discounted monetary loss $\frac{|\alpha - L\gamma|}{r} = 100$. Thus, if the principal fires the agent too early, she will be paying too much replacement cost and severance pay. So the second option is dominated by the first one and in this case the principal would rather allow the agent to take up inefficient risk. The optimal contract terms are similar to the baseline result stated in proposition 1.2. On the right panel of figure 5.1, the replacement cost, $c_a = 30$, is low relative to the expected discounted monetary loss. Therefore, the principal is able to fire her agents more frequently. In particular, the principal fires the agent when the limited liability constraint starts to bind and she cannot stop the agent from taking inefficient downside risk. This occurs at the point $W_R^* = 7.5$. As the principal can prevent inefficient risk-taking, the second option dominates the first option. The analysis here thus delivers a prediction: when there is strong search frictions or when the manager has developed firm specific capital, which make the investors more costly to replace the manager, then we shall see that the firm's profit volatility becomes higher.

1.5.2 Multiple Projects.

Suppose there are multiple risky projects $i = 1, ..., n$. Let $N_i = \{N_{it}\}_{t\geq 0}$, $i = 1, ..., n$ be n independent standard Poisson processes with intensity $\{\lambda(a_t^s)\}_{t\geq 0}$ respectively. The cash flows of the firm is given by the dynamics

$$dY_t = \left(\mu(a_t^e) + \sum_{i=1}^{n} \alpha_i\right) dt + \sigma dZ_t - \sum_{i=1}^{n} L_i dN_{it}$$

Different projects carry different degree of risks Specifically, taking project i leads to an increase in the drift by α_i and intensity of loss is increased from λ_i to $\lambda_i + \gamma_i$. Extend the action profile to $a_t = (a_t^e, a_{1t}^s, ..., a_{nt}^s)$ where a_t^e is the effort as before and $a_{it}^s \in \{0, 1\}$ denote the risk action. In particular, $a_{it}^s = 1$ means risk-taking at time t and $a_{it}^s = 0$ means not taking risk. As in the baseline model, I assume risk-taking destroys values, that is, $\alpha_i - L_i\gamma_i < 0$ for all i.

By the same methodology, the dynamics of the continuation value of the agent satisfies

$$dW_t = \rho W_{t-} dt - dI_t + \beta_t \sigma dZ_t - \sum_{i=1}^{n} \psi_{it}(dN_{it} - (\lambda_i + \gamma_i a_{it}^s)dt)$$

and the relevant incentive constraints that restrict risk-taking behavior are $\psi_{it} \geq \beta_t \frac{\alpha_i}{\gamma_i}$ if

$a^s_{it} = 1$ and no restriction is needed if $a^s_{it} = 0$. Also, $\beta_t \geq \frac{B}{\mu}$ is the incentive constraint on high efforts. It follows that the HJB-equation is

$$rF(W) = \mu - \sum_{i=1}^{n} L_i\lambda_i + \sum_{i=1}^{n} a^s_i(\alpha_i - L_i\gamma_i) + F'(W)(\rho W + \sum_{i=1}^{n}(1 - a^s_i)\psi_i\lambda_i) \qquad (1.5.1)$$

$$+\frac{1}{2}F''(W)\beta^2\sigma^2 + \sum_{i=1}^{n} \lambda_i(F(W - (1 - a^s_i)\psi_i) - F(W))$$

As in proposition 5, there exists a C^2 and concave function F which is a solution to the above differential equation with boundary conditions $F(0) = l$, $F'(W^p) = -1$, and $F''(W^p) = 0$. At optimum, to induce high efforts, $\beta_t = \frac{B}{\mu}$ and to deter the agent from taking the risky project i, $\psi_{it} = \frac{B}{\mu}\frac{\alpha_i}{\gamma_i}$ for all t. There also exists a set of switching points $\{W^*_1, ..., W^*_n\}$ with $W^*_i \in [\psi_i, W^p]$ for all i such that $a^s_{it} = 0$ if $W_t \leq W^*_i$ and $a^s_{it} = 1$ if otherwise. Therefore, the optimal contract in the benchmark model is robust to the multiple projects extension.

Some natural questions arise: how does the project dynamics look like? how do agency costs affect the project choice? Is there any structure on the set $\{W^*_1, ..., W^*_n\}$? To isolate the effect of agency costs on project choices, I assume that the marginal expected monetary loss of each project and the baseline intensity are the same. Suppose two projects $\{i, j\}$ have different agency-related parameters, in particular, $\frac{\alpha_i}{\gamma_i} > \frac{\alpha_j}{\gamma_j}$. In words, the agent has higher incentives to take project i than project j either because of a higher increase in drift $\alpha_i > \alpha_j$ or it is more difficult for the principal to detect hidden risk-taking in project i, $\gamma_i < \gamma_j$. To deter the agent from taking risks, the punishment $\psi_i > \psi_j$ are needed and hence the agency cost of preventing the agent from taking project i is higher than that of project j. Therefore, it is expected that $W^*_i > W^*_j$. The following proposition formalizes this intuition.

Proposition 1.6. *Suppose $\alpha_i - L_i\gamma_i = K$ and $\lambda_i = \lambda$ for all $i = 1, ..., n$. Then $\frac{\alpha_i}{\gamma_i} \geq \frac{\alpha_j}{\gamma_j}$ if and only if $W^*_i \geq W^*_j$ for any i, j.*

Based on the proposition, if we order the projects according to their agency-related parameters, without loss of generality, let $\frac{\alpha_1}{\gamma_1} \geq \frac{\alpha_2}{\gamma_2} \geq ... \geq \frac{\alpha_n}{\gamma_n}$, then we have $W_1^* \geq W_2^* \geq ... \geq W_n^*$. This implies that when the firm moves away from the payment boundary W^p, the agent will start taking project 1 first as it entails the highest agency cost. Then followed by project 2 and so on. As a result the total cash flows become more volatility as the firm performance keeps deteriorating.

1.5.3 Random Jump Size.

In the benchmark model, the size of the loss is fixed at L. How would the contract be affected if the losses are random? Suppose the loss is modeled by a compound Poisson process $\{\sum_{k=1}^{N_t} J_k\}_{t\geq 0}$ where $\{N_t\}_{t\geq 0}$ is a standard Poisson process with intensity $\{\lambda(a_t^s)\}_{t\geq 0}$ and J_k's are i.i.d. random variables with distribution H on $(-\infty, 0)$. The cash flow is[26]

$$dY_t = \mu(a_t)dt + \sigma dZ_t + \sum_{k=1}^{N_t} J_k$$

The effect of risk-taking is the same as the benchmark: risk-taking implies a higher drift of Y_t but also a higher intensity of N_t. Note that the benchmark is obtained by specifying a degenerate distribution H on $-L$.[27] The distribution is not affected by any of the moral hazard variables.

Because of the random jump size, there is more public information in the extended model than in the benchmark. Does the optimal contract use this information? More formally, continuous-time contracting relies on the martingale representation theorem to deliver the incentive loadings and thus the relationship between public information and the

[26]An alternative way to write the cash flow process is $dY_t = \mu(a_t)dt + \sigma dZ_t + \int_0^t J \cdot N(dt, dJ)$, where $N(\cdot, \cdot)$ is a Poisson random measure. The random measure counts the number of times up to t that the jump size falls in the Borel set B, i.e., $N(t, B) = \#\{s \in [0, t] : J \in B\}$. Fix a set B, $N(t, B)$ is a Poisson process with intensity ν, a Levy measure defined as $\nu(B) = E[N(1, B)]$. A general Levy process allows for infinitely many jumps over a compact time interval. By construction, a compound Poisson process has only a finite number of jumps with a finite time interval. The Levy measure associated with the compound Poisson process is $\nu = \lambda(a_t^s) \int_{-\infty}^0 JdH(J)$, which is finite. For more details, see Applebaum (2009) and Oksendal and Sulem (2005).

[27]That is, a Dirac measure $H(B) = \delta_{-L}(B)$ for any Boral set $B \in \mathcal{B}$(-∞,0)

agent's continuation value. The set up induces two martingales for which the martingale representation theorem can be applied to: (i) the compensated Poisson process $\{N_t - \int_0^t \lambda(a_u^s)du\}_{t\geq 0}$, and (ii) the process $\{\sum_{k=1}^{N_t} J_k - \int_0^t \int_{-\infty}^0 \lambda(a_t^s)JdH(J)\}_{t\geq 0}$. The former includes only the number of jumps while the latter, the compensated compound Poisson, includes both the number of jumps and the jump size. However, the jump size is not affected by moral hazard. In other words the jump size provides no information about when the agent is engaging in hidden gambling. Thus incorporating the jump size should be sub-optimal as the agent is exposed to more risks and her continuation value is more volatile. This is costly to the principal hence she does not want to write a contract on the jump size. The following proposition illustrates this point.[28]

Proposition 1.7. *The optimal contract with random jump size is the same as the optimal contract in the benchmark. That is, the contract does not use information about the jump size.*

1.6 Conculsion

I consider a dynamic environment in which an agent who works for her principal faces a multi-task problem. On the one hand, the agent must exert costly effort to improve the profitability of the firm. On the other hand, she has discretion over undertaking projects/assets with different expected returns and riskiness. Risky projects generate lower expected returns in the sense that large risks occur more frequently. By applying the continuous-time techniques, I provide a clean characterization of the optimal contract. In the contracting equilibrium, severe punishment deters agent's hidden gambling but the principal has to bear additional agency cost. Certain degree of NPV-reducing risk-taking is second-best optimal, and the firm's project selection dynamics is completely characterized. The framework can be applied into project choices in standard business firms, or selection of defaultable assets in financial institutions.

[28] In fact, the argument show that Holmstrom's informativeness principle holds in the above environment. See Holmstrom (1979)

I further study the implementation of the optimal contract and apply my analysis to the hedge fund industry. Compare to the literature on continuous-time dynamic contract theory, two main additional implications emerge. First, the model supports incomplete hedging due to the existence of agency cost. In the capital structure implementation, insurance contracts are used to cover losses. Incomplete hedging against downside risks is optimal because partial coverage implies the agent, who has discretion over project selection, is exposed to the Poisson risk and thus has the right incentive to select the safe project. Second, high-water mark contract is optimal and "distance-to-threshold" is important in understanding hedge fund manager's risk-taking behavior. The fund manager has a higher incentive to gamble and take risks as the current fund performance lies further away from the bonus threshold.

My results raise several interesting questions. For example, what is the optimal contract if the agent can also control the size of the losses and how are risks managed in this case? In a delegated portfolio choice environment, where the price of risky assets could jump downward, what is the optimal portfolio choice? How would the compensation contract of the fund manager look like? The recent financial crisis is partly driven by default of various financial assets. What is the aggregate implications of hidden gambling? These questions are left for future research.

Chapter 2

Dynamic Team Incentives

2.1 Introduction

This paper generalizes an optimal contracting problem between shareholders and a manager in He (2011) to a multi-agent environment. In my model, the shareholders (the principal) hires a team of managers (the agents) to run the business. The firm comprises multiple divisions and agents exert costly efforts to improve the divisional cash flows. The firm size evolves stochastically based on the agents' efforts and the aggregate cash flows. As in He (2011), I embed the agency problem into a structural model of capital structure in corporate finance (Leland (1994)). The exercise further generalizes the agency framework of corporate finance and hence allows us to study the effect of moral hazard problems on firm valuation, investment policy and capital structure.

The first half of the paper sets up the model and provides a general analysis of the optimal contracting problem with many agents. In the model, the agents are assumed to have exponential utility (CARA). To solve for the optimal contract in continuous-time, I adopt the martingale method developed by Sannikov (2008) and Williams (2009). The method employs the agent's continuation value as a state variable. The method applied in a n-agent setting leads to n state variables. The absence of wealth effect in exponential utility allows us to break the firm's value into the shareholders value and the team's value. The team's value is the sum of the certainty equivalent of individual agents, which is a function of continuation values. The separation allows us to characterize the optimal contract by an ordinary differential equation (ODE). I derive the firm's optimal value, the agents' optimal effort dynamics, and the required incentive schemes. One of the insights is that the contract could reward the agents for high outputs based on joint performance or relative performance. The analysis shows that the evaluation scheme to be used depends on the production technology and the noise structure, but not on the agent's risk preferences.

In the second half of the paper I embed the multi-agent moral hazard problem into Leland's structural model of debt. The embedding allows us to endogenize the firm's growth rate, which is the drift of the geometric Brownian cash flows process in Leland's model. The firm growth is thus affected by the agents' effort. Following Leland, the firm only issues consol bonds that pay a constant stream of coupon until default. In designing the optimal debt contract, both the shareholders and the debt holders take into account of the agency problem in teams once the firm starts running. Hence the principal writes down the optimal compensation contract as a best response to the capital structure.

In the quantitative analysis, I show that the firm's investment policy, optimal capital structure (leverage ratio), and default threshold depend on the number of agents working in the firm. As in He (2011), debt overhang problem generates a negative relationship between leverage and incentives. However, as the firm hires more and more agents, the convexity of the incentive cost allows the principal to smooth out the effort investment cost and thus mitigate the debt-overhang. Therefore, the firm with more agents has a higher endogenous growth rate and accumulates capital stock at a faster rate. These firms will stack up enough financial resources and they are less prone to insolvency. As a best response, these firms will be able to issue a higher amount of debt to take advantage of the tax shields.

This paper belongs to the growing continuous-time optimal contracting literature. Sannikov (2008) and Williams (2009) develop a general methodology for solving dynamic principal-agent problems. There are many applications of the martingale method in various economic environments, for example, DeMarzo and Sannikov (2006) study the use of credit lines in a model where a risk-neutral agent is protected by limited liability; Biais, Mariotti, Rochet, and Villeneuve (2010) study firm dynamics with a Poission risk; Piskoroski and Tchistyi (2010) study the optimal mortgage design. However, all papers employ a model with a single agent. My paper is the first in the literature to study team production and I contribute to the literature by showing that the same martingale method can be applied to model with multiple agents. I also contribute to the analysis of team production in a dynamic environment. Holmstrom (1982) shows that team members free ride when they obtain only a fraction of the output and a budget-breaker is needed to achieve efficiency. Mookherjee (1984) characterizes the optimal contract with many agents and highlights the multiple equilibria problem. Che and Yoo (2001) study the nature of performance evaluation contract in a discrete-time principal-agent set up. Their model assumes that agents are risk-neutral and in my paper I allow for risk aversion. Bonatti and Horner (2011) and Georgiadis (2014) both study team dynamics. They analyze effort provision as the state of the project evolves and how efforts vary with the number of agents in the team. None of them focus on the optimal contract between the principal and the team.

The rest of the paper is organized as follows. Section 2.2 develops the model. Section 2.3 describes the model's solution and characterizes the optimal contract. I also study the use of joint versus relative performance evaluation contract in this section. Section 2.4 focuses on the impact of agency issue in capital structure and provides a quantitative analysis of the team effect on optimal capital structure. The last section concludes and proofs are delegated to the Appendix.

2.2 The Model

Consider the following infinite horizon, continuous-time contracting environment. There is a risk-neutral principal and n risk-averse agents. Both the principal and the agents discount cash flows at the market interest rate $r > 0$. The agents form a team to operate m tasks for the principal. Let X_t be the firm size at time t and suppose the firm produces cash flows at rate X_t. The firm size evolves according to

$$dX_t = X_t(\sum_{k=1}^{m} dA_{kt})$$

where A_{kt} is the productivity of task k up to time t and there are m tasks. The agents can affect the productivity of each of the task by exerting efforts non-cooperatively. Let $Z = (Z_1, ..., Z_d)$ be a d-dimensional independent standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ that satisfies the usual condition. The dynamics of $A_t = (A_{1t}, ..., A_{nt})$ satisfies the following the following system of stochastic differential equations

$$dA_{kt} = \mu_k(a_t)dt + \sum_{l=1}^{d} \sigma_{kl} dZ_{lt} \; k = 1, ..., n$$

where $a_t = (a_{1t}, ..., a_{nt})$ is the agents' effort profile at time t with $a_{it} \in [0, \bar{a}_i]$ for all $i = 1, ..., n$. $\mu_k : [0, \bar{a}_i]^n \to \mathbb{R}$ is the drift function for task k. I assume that $\frac{\partial \mu_k}{\partial a_i}(a) > 0$ and $\frac{\partial^2 \mu_k}{\partial a_i^2}(a) \leq 0$ for all k and i. $\sigma = [\sigma_{kl}]$ is a $m \times d$ volatility matrix with each of the element being a constant. Define a process $B = (B_1, ..., B_n)$ with $B_{kt} = \sum_{l=1}^{d} \sigma_{kl} dZ_{lt}$ for $k = 1, ..., m$, then B is a n-dimensional *correlated* Brownian motion. The associated $n \times n$ correlation matrix is $\rho = \sigma \cdot \sigma^T$. Moreover, the firm size X_t and productivity A_t are publicly observable and contractible. Therefore, the principal's information is the filtration generated by the productivity process $A = (A_t)_{t\geq 0}$, that is, I take $\mathcal{F}_t = \sigma(\{A_{ks} : s \leq t\}_{k=1}^{m})$.

Agent i effort process is $a_i = (a_{it})_{t\geq 0}$. Effort exerted by agent i is neither observable to the principal nor to any other agents. The agents' monetary effort cost is $h_i(a_{it}, X_t)$, where the cost function $h_i : [0, \bar{a}_i] \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\frac{\partial h_i}{\partial a_i}(a_i, X) > 0$ and $\frac{\partial^2 h_i}{\partial a_i^2}(a_i, X) > 0$. Let $c_i = (c_{it})_{t\geq 0}$ be agent i's consumption process with $c_{it} \in \mathbb{R}$. The instantaneous utility function $u_i : \mathbb{R} \times [0, \bar{a}_i] \to \mathbb{R}$ is assumed to be in CARA form

$$u_i(c_{it}, a_{it}) = -\frac{1}{\gamma_i} \exp(-\gamma_i(c_{it} - h_i(a_{it}, X_t)))$$

where $\gamma_i > 0$ is the coefficient of absolute risk aversion. Moreover, the agents can privately save at the risk-free interest rate r. The account balance S_{it} is not observable to the principal and to any other agents. This implies that the actual consumption is unknown to the principal.

2.2.1 The Contracting Problem.

At time 0, the principal offers a contract $\Gamma = \{(c_{it}, a_{it})_{t\geq 0}\}_{i=1}^{n}$ to the agents. The contracting party can fully commit to the contract. The contract specifies, for each agent i, the wage process $(c_{it})_{t\geq 0}$ and the recommended effort level $(a_{it})_{t\geq 0}$. The contract is written on publicly observable productivity and firm size. At any time t, the sequence of events that occur during the small time interval $[t, t+dt)$ are

1. The agents takes their action $\hat{a}_t = (\hat{a}_{1t}, ..., \hat{a}_{nt})$ simultaneously.

2. The tasks productivity $(dA_{1t}, ..., dA_{nt})$ realize, hence the cash flow X_t.

3. The agents are paid wage $c_t = (c_{1t}, ..., c_{nt})$ and choose their actual consumption $\hat{c}_t = (\hat{c}_{1t}, ..., \hat{c}_{nt})$.

Given this timing, formally the process of effort profile a is $\mathcal{F}_t$-predictable, and the consumption process c is $\mathcal{F}_t$-adapted. Note that the team effort process a induces a unique probability measure P^a over the paths of productivity $(A_t)_{t\geq 0}$. Therefore, I let $E^a(\cdot)$ to denote the expectation operator under the measure P^a.

Given a contract $\Gamma = \{(c_{it}, a_{it})_{t\geq 0}\}_{i=1}^{n}$ and the effort process of the team members $a_{-i} = (a_{-it})_{t\geq 0}$, agent i's decision problem is

$$W_0(\Gamma) = \max_{(\hat{c}_i, \hat{a}_i)|a_{-i}} E^{(\hat{a}_i, a_{-i})}\left[\int_0^\infty e^{-rt} u_i(\hat{c}_{it}, \hat{a}_{it})dt\right] \tag{2.2.1}$$

subject to

$$dA_{kt} = \mu_k(\hat{a}_{it}, a_{-it})dt + \sum_{l=1}^{d} \sigma_{kl} dZ_{lt},\ \forall k = 1, ..., m$$

$$dS_{it} = rS_{it}dt + (c_{it} - \hat{c}_{it})dt,\ S_{i0} = 0$$

where $W_0(\Gamma)$ is agent i's time-0 value derived from the contract Γ. Note that when the agent chooses his effort level, he is making his choice given that he expects his teammates to follow a_{-i}. The last constraint reflects the agents' ability to privately save. I assume that all agents have no initial wealth in the sense that the initial saving balance is $S_{i0} = 0$. Having received a wage c_{it} at time t, the agent could consume a different amount $\hat{c}_{it}$. Any unspent money $c_{it} - \hat{c}_{it} > 0$ is saved in a personal account that grows according to the market interest rate r.

Facing any contract Γ, if (i) the agents find it optimal to follow the recommended effort, given the team members follow the recommended effort; and (ii) the agent takes the recommended consumption, that is, when the solution to the agent's problem is $(\hat{c}_{it}, \hat{a}_{it})_{t\geq 0} = (c_{it}, a_{it})_{t\geq 0}$ for all $i = 1, ..., n$. Then the contract Γ is said to be *incentive-compatible* and *no-savings*. Since the principal could also save at interest rate r and the party can fully commit, the Revelation Principle applies:

Lemma 2.1. *without loss of generality, I can focus on incentive-compatible and no-savings contract.*

Suppose agent i's reservation utility is W_{i0}, then the principal's problem is

$$\max_{\{(c_{it}, a_{it})_{t\geq 0}\}_{i=1}^n} E^a \left[\int_0^\infty e^{-rt} (X_t - \sum_{i=1}^n c_{it}) dt \right]$$

subjec to

$$E^a \left[\int_0^\infty e^{-rt} u_i(c_{it}, a_{it}) dt \right] \geq W_{i0} \ \forall i = 1, ..., n$$

$$E^a \left[\int_0^\infty e^{-rt} u_i(c_{it}, a_{it}) dt \right] \geq E^{(\hat{a}_i, a_{-i})} \left[\int_0^\infty e^{-rt} u_i(\hat{c}_{it}, \hat{a}_{it}) dt \right] \ \forall (\hat{c}_i, \hat{a}_i), \forall i = 1, ..., n$$

$$dA_{kt} = \mu_k(\hat{a}_{it}, a_{-it})dt + \sum_{l=1}^{d} \sigma_{kl} dZ_{lt},\ \forall k = 1, ..., m$$

The first constraint is the individual rationality constraint and the second one is the incentive-compatibility constraint. Note that the wage payment c induces a repeated game of imperfect public monitoring among the agents. The agents could deviate from the assigned effort level and based their choice on publicly observed productivity. The incentive constraints require that the recommended effort a being an equilibrium of this game.

2.2.2 First-Best Solution.

Suppose there is no agency issue, that is, both efforts and consumption are contractible, then incentive constraints are not required. Let $F^{FB}(X, W_{10}, ..., W_{n0})$ denote the principal's first-best value function. Since the principal does not need to provide incentives, and as the agents are risk-averse, the principal only concerns is risk-sharing. The optimal risk-sharing contract is solved as follows: The principal can pay the agents an amount at any time t that delivers a constant stream of instantaneous utility. Define CE_i^{FB} as follows

$$\frac{1}{r}(-\frac{1}{\gamma_i}\exp(-\gamma_i CE_i^{FB}) = W_{i0}$$

then $CE_i^{FB} = -\frac{1}{\gamma_i}\ln(-\gamma_i r W_{i0})$. Hence, $c_{it} = h_i(a_{it}, X_t) - \frac{1}{\gamma_i}\ln(-\gamma_i r W_{i0})$ and by construction, the participation constraint is satisfied. It follows that the principal's first-best value can be written as $F^{FB}(X, W_{10}, ..., W_{n0}) = f^{FB}(X) - \sum_{i=1}^{n} -\frac{1}{\gamma_i r}\ln(-\gamma_i r W_{i0})$ where $f^{FB}(X)$ satisfies the following HJB-equation

$$rf^{FB}(X) = \max_a\{X - \sum_{i=1}^{n} h_i(a_i, X) + \frac{\partial f^{FB}(X)}{\partial X}\sum_{k=1}^{m}\mu_k(a)X + \frac{1}{2}\frac{\partial^2 f^{FB}(X)}{\partial X^2}\sum_{l=1}^{d}(\sum_{k=1}^{m}\sigma_{kl})^2 X^2\}$$

Assume that $h_i(a_i, X) = \tilde{h}_i(a_i)X$, then f^{FB} is proportional to X. Guess that $f^{FB}(X) = QX$, then

$$rQ = \max_{(a_1,..,a_n)} \{1 - \sum_{i=1}^{n} \tilde{h}_i(a_i) + Q \sum_{k=1}^{m} \mu_k(a)\}$$

Denote a^* the optimal effort, then $Q = \frac{1-\sum_{i=1}^{n} \tilde{h}_i(a_i^*)}{r-\sum_{k=1}^{m} \mu_k(a^*)}$. Note that the optimal effort is constant over time as the above maximization is independent of X, which is dynamic.

Remark. The first-order conditions with respect to a_i is $Q \sum_{k=1}^{m} \frac{\partial \mu_k}{\partial a_i}(a) = \frac{\partial \tilde{h}_i}{\partial a_i}(a_i)$ if a_i is an interior solution. If we further specify: $\tilde{h}_i(a_i) = \frac{1}{2}\theta_i a_i^2$, with $\theta_i > 0$ and $\mu_k(a) = \sum_{i=1}^{n} \tilde{\mu}_{ki} a_i$, with $\tilde{\mu}_{ki} > 0$. Then an interior solution for effort is $a_i^* = \frac{1}{\theta_i}(Q \sum_{k=1}^{m} \tilde{\mu}_{ki})$.

2.3 Model Solution

This section discusses the model solution. The solution method follows the martingale approach developed by Sannikov (2008). Applying the martingale approach, I obtain the dynamics of the agents' continuation utility. Then I provide the necessary and sufficient conditions for the contract to be incentive-compatible and no-saving. These conditions impose restrictions on the continuation utility dynamics. With such dynamics in hand, I represent the principal's optimal contracting problem as a stochastic control problem where the principal's value function satisfies an ordinary differential equation that can be solved.

2.3.1 The Dynamics of Continuation Utility.

Consider agent i. Fix any arbitrary contract Γ and effort process of other team members a_{-i}, agent i's continuation utility at time t if i follows continuation strategy $(c_{is}, a_{is})_{s \geq t}$ is defined as

$$W_{it}(\Gamma) \equiv E_t^a \left[\int_t^{\infty} e^{-r(s-t)} u_i(c_{is}, a_{is}) ds \right] \tag{2.3.1}$$

this is the promised utility at time t under contract Γ when agent i follows the continuation consumption and effort $(c_{is}, a_{is})_{s \geq t}$. As the continuation value $W_t = (W_{1t}, ..., W_{nt})$ reacts to public information and summarizes the performance of all the agents. The principal could control how the continuation values responses to realized productivity in order to provide incentives. Following the continuous-time contracting literature, the martingale representation theorem helps to deliver the dynamics of W_t.

Lemma 2.2. *Given any contract* $\Gamma = \{(c_{it}, a_{it})_{t\geq 0}\}_{i=1}^{n}$, *there exists a* $\mathbb{R}^{n\times m}$*-valued* $\mathcal{F}_t$*-progressive measurable process* $\beta = (\beta_t)_{t\geq 0}$ *such that for any* $i = 1, ..., n$ *and any* t

$$dW_{it}(\Gamma) = (rW_{it}(\Gamma) - u_i(c_{it}, a_{it}))dt + (-\gamma_i r W_{it})\sum_{k=1}^{m}\beta_t^{ik}(\sum_{l=1}^{d}\sigma_{kl}dZ_{lt}) \tag{2.3.2}$$

As is now standard in the literature, the dynamics of continuation captures the promise-keeping and incentive components of the contract. On expectation, $E_t\left[dW_{it} + u_i(c_{it}, a_{it})dt\right]$ $= rW_{it}dt$, hence the drift plays the role of promise-keeping. The volatility plays the role of incentive provision: $\sum_{l=1}^{d}\sigma_{kl}dZ_{lt}$ reflects the total shock to task k, and it connects continuation value to the observed productivity: $\beta_t^{ik}(-\gamma_i r W_{it})(dA_{it} - \mu_k(a_t)dt)$. Hence β_t^{ik} measures agent i's the dollar incentive to task k and $-\gamma_i r W_{it}$ converts money to utility. As there are n tasks, the total volatility is the sum incentive loading on individual task.

2.3.1.1 No-Saving Condition.

The no-saving condition states that agent i's marginal utility of consumption equals to his marginal utility of saving. The condition is derived from the equation stated in the following lemma......

To state the result formally, consider agent i's decision problem at time t with accumulated saving S

$$W_{it}(S;\Gamma) = \max_{(\hat{c}_i,\hat{a}_i)|a_{-i}} E_t^{(\hat{a}_i, a_{-i})}\left[\int_t^{\infty} e^{-r(s-t)}u_i(\hat{c}_{is}, \hat{a}_{is})ds\right] \tag{2.3.3}$$

subject to

$$dA_{ks} = \mu_k(\hat{a}_{is}, a_{-is})ds + \sum_{l=1}^{d}\sigma_{kl}dZ_{ls},\ \forall k = 1, ..., m$$

$$dS_{is} = rS_{is}dt + (c_{is} - \hat{c}_{is})ds,\ \ S_{it} = S \text{ for} s > t$$

Therefore, $W_{it}(S;\Gamma)$ is agent i's continuation value when she deviates to off-equilibrium with positive savings. Let $W_{it}(0;\Gamma)$ be agent i's continuation value (**??**) along the no-saving path.

Lemma 2.3. *At any time $t \geq 0$, for the agent i with saving S, we have*

$$W_{it}(S;\Gamma) = e^{-\gamma_i r S} \cdot W_{it}(0;\Gamma) \tag{2.3.4}$$

The equation (**??**) for the continuation value with saving S pins down the agent's marginal utility of saving. Fix the effort policy, the optimality of the agent's consumption-saving problem in (**??**) requires her marginal utility of consumption equals to her marginal value of saving (on the no-saving path):

$$\frac{\partial}{\partial c} u_i(c_{it}, a_{it}) = \frac{\partial}{\partial S} W_{it}(S;\Gamma)|_{S=0}$$

By equation (**??**), to rule out private saving, we need:

$$e^{-\gamma_i(c_{it} - h_i(a_{it}, X_t))} = \frac{\partial}{\partial c} u_i(c_{it}, a_{it}) = \frac{\partial}{\partial S} W_{it}(S;\Gamma)|_{S=0} = -\gamma_i r W_{it}$$

where the last equality follows from equation (**??**). Therefore $u_i(c_{it}, a_{it}) = rW_{it}$. Since this holds for every agent, each agent's continuation value is thus a martingale

$$dW_{it} = (-\gamma_i r W_{it}) \sum_{k=1}^{m} \beta_t^{ik} (\sum_{l=1}^{d} \sigma_{kl} dZ_{lt}) \tag{2.3.5}$$

In addition, the fact that the flow continuation value equals to the agent's instantaneous utility implies each agent's wage process has to satisfy

$$c_{it} = h_i(a_{it}, X_t) - \frac{1}{\gamma_i} \ln(-\gamma_i r W_{it}) \tag{2.3.6}$$

2.3.1.2 Incentive-Compatibility condition.

Now I turn to incentive-compatibility condition and characterize the incentive loadings β_t in (**??**)

Lemma 2.4. *Agent i's effort process a_i is a best response to a_{-i} if and only if for all t*

$$a_{it} \in \arg\max_{\hat{a}_{it}} \left\{ \sum_{k=1}^{m} \beta_t^{ik}(-\gamma_i r W_{it})\mu_k(\hat{a}_{it}, a_{-it}) + u_i(c_{it}, \hat{a}_{it}) \right\}$$

Therefore, a is a PPE if and only if the above condition holds for all $i = 1, ..., n$.

The proposition delivers restrictions on the incentive loadings β that has to satisfy when the principal seeks to implement a as an equilibrium. The first-order conditions are for all $i = 1, ...n$,

$$\sum_{k=1}^{m} \beta_t^{ik} \frac{\partial \mu_k(a_t)}{\partial a_i} = \frac{\partial h_i(a_{it}, X_t)}{\partial a_i}$$

where the left-hand side represents the marginal benefits of exerting effort. Since a marginal increase in effort affects the expected productivity in each task, and the agent is compensated based on the observed outcome in each task. The marginal benefit of effort is the summation of the change in continuation value due to the higher productivity of each task. The right-hand side is the marginal cost of effort.

2.3.2 HJB-equation.

This section derives the HJB-equation for the principal value function and adopt a dynamic programming approach to solve the optimal contracting problem. The relevant state variables in the problem is the firm size X_t and the agents' continuation value W_{it}, $i = 1, ..., n$. Upon rearrangement, their dynamics are

1. $dX_t = X_t \sum_{k=1}^{n} \mu_k(a_t)dt + X_t \sum_{l=1}^{d}(\sum_{k=1}^{m} \sigma_{kl})dZ_{lt}$

2. $dW_{it} = (-\gamma_i r W_{it}) \sum_{l=1}^{d}(\sum_{k=1}^{m} \beta_t^{ik}\sigma_{kl})dZ_{lt}$ for $i = 1, ..., n$

the system is driven by the d-dimensional Brownian motion. Alternatively, a $n+1$-dimensional correlated Brownian motion with the following correlation matrix drives the system

$$\begin{pmatrix} \tilde{B}_0 \\ \tilde{B}_1 \\ \vdots \\ \tilde{B}_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m \sigma_{k1} & \sum_{k=1}^m \sigma_{k2} & \cdots & \sum_{k=1}^m \sigma_{kd} \\ \sum_{k=1}^m \beta^{1k}\sigma_{k1} & \sum_{k=1}^m \beta^{1k}\sigma_{k2} & \cdots & \sum_{k=1}^m \beta^{1k}\sigma_{kd} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m \beta^{nk}\sigma_{k1} & \sum_{k=1}^m \beta^{nk}\sigma_{k2} & \cdots & \sum_{k=1}^m \beta^{nk}\sigma_{kd} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_d \end{pmatrix}$$

The principal's value function hence satisfies the following HJB-equation

$$rF(X, W_1, ..., W_n) = \max_{(a_1,...,a_n)} \{X - \sum_{i=1}^{m} c_i + \frac{\partial F}{\partial X} X \sum_{k=1}^{m} \mu_k(a) + \frac{1}{2}\frac{\partial^2 F}{\partial X^2} X^2 \sum_{l=1}^{d} (\sum_{k=1}^{m} \sigma_{kl})^2 +$$

$$+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{d} \frac{\partial^2 F}{\partial W_i \partial W_j}(-\gamma_i r W_i)(-\gamma_j r W_j)(\sum_{k=1}^{m}\beta^{ik}\sigma_{kl})(\sum_{k=1}^{m}\beta^{jl}\sigma_{kl})$$

$$+\sum_{i=1}^{n}\sum_{l=1}^{d} \frac{\partial^2 F}{\partial X \partial W_i}(-\gamma_i r W_i)(X\sum_{j=1}^{n}\sigma_{jk})(\sum_{j=1}^{n}\beta^{ij}\sigma_{jk})\}$$

where the effort choices have to satisfy the incentive constraint stated in proposition 2.2. With CARA utility, and following He (2011), guess

$$F(X, W_1, ..., W_n) = f(X) - \sum_{i=1}^{n} -\frac{1}{\gamma_i r}\ln(-\gamma_i r W_i)$$

where $-\frac{1}{\gamma_i r}\ln(-\gamma_i r W_{it})$ is agent i's certainty-equivalent. The guess is intuitive as it says the principal value equals the firm's value plus the "team's certainty equivalent. Moreover, as there is no consumption externality or cost synergies, the balance owing to the team is the sum of the individual balance (certainty-equivalent). This guess naturally generalizes a result in a single agent setup to a multi-agent environment. The guess implies $\frac{\partial F}{\partial X} = f'(X)$, $\frac{\partial^2 F}{\partial X^2} = f''(X)$, $\frac{\partial F}{\partial W_i} = \frac{1}{\gamma_i r W_i}$, $\frac{\partial^2 F}{\partial W_i^2} = -\frac{1}{\gamma_i r W_i^2}$, $\frac{\partial^2 F}{\partial X \partial W_i} = \frac{\partial^2 F}{\partial W_i \partial W_j} = 0$. Plugging in this quantities, f has to satisfies the ODE:

$$rf(X) = \max_{(a_1,...,a_n)} \{X - \sum_{i=1}^{n} h_i(a_i, X) + f'(X)\sum_{k=1}^{m}\mu_k(a)X + \frac{1}{2}f''(X)X^2\sum_{l=1}^{d}(\sum_{k=1}^{m}\sigma_{kl})^2$$

$$-\frac{1}{2}\sum_{i=1}^{n}\gamma_i r\sum_{l=1}^{d}(\sum_{k=1}^{m}\beta^{ik}(a)\sigma_{kl})^2\}$$

The interpretation is straightforward: on the left-hand side is the firm's flow value. On the right-hand side, the first term is the instantaneous cash flow X, the third and forth term capture the expected instantaneous change in firm's value due to the dynamics of X. The second and last term are the monetary cost and incentive cost of efforts. The optimal effort a^* is characterized by:

$$a_i^* \in \arg\max_{a_i \in [0, \bar{a}_i]} \left\{ f'(X) \sum_{k=1}^{m} \mu_k(a) X - h_i(a_i, X) - \frac{1}{2}\gamma_i r \sum_{l=1}^{d}(\sum_{k=1}^{m}\beta^{ik}(a)\sigma_{kl})^2 \right\}$$

for all $i = 1, ..., n$. The first term in bracket captures the expected total productivity of i's effort. The second term is again monetary effort cost. Because the agent is risk-averse the last term is the incentive cost of exposing again to productivity uncertainty.

2.3.3 Model Implication

Two-agent case.

In this section, I specialize the contracting environment to a two-agent team in order to obtain clean analytic solution for the agents' efforts and incentives. Let $n = m = 2$ and suppose there are two tasks. The productivity dynamics are

$$dA_{1t} = (\mu_{11}a_1 + \mu_{12}a_2)dt + \sigma_1 dZ_{1t} + \sigma_c dZ_{ct}$$

$$dA_{2t} = (\mu_{21}a_1 + \mu_{22}a_2)dt + \sigma_2 dZ_{2t} + \sigma_c dZ_{ct}$$

Hence $d = 3$. The volatility matrix $\sigma = [\sigma_{kl}]$ and the correlated Brownian motion can be put in the following matrix form

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & \sigma_c \\ 0 & \sigma_2 & \sigma_c \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_c \end{pmatrix}$$

with the correlation matrix $\rho = \sigma \cdot \sigma^T$

$$\begin{pmatrix} \sigma_c^2 + \sigma_1^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_2^2 \end{pmatrix}$$

I think of agent i is assigned to task i. Hence dA_{it} is i's output. But by exerting effort, i can also "help" his teammate j. With this interpretation, (Z_1, Z_2) are idiosyncratic shocks and Z_c is a common shock where its variance σ_c^2 represents the correlation between the two agents' output.

Let $h_i(a_i, X) = \frac{1}{2}\theta_i a_i^2 X$. Applying the lemmas in the previous section, the incentive constraint requires $a_i = \frac{1}{\theta_i X}(\beta^{i1}\mu_{1i} + \beta^{i2}\mu_{2i})$ when a_i is an interior solution. Notice that due to additive drift, β^j does not affect a_i. Hence β^1 and β^2 are determined using a separate set of equations. The ODE for $f(\cdot)$,

$$rf(X) = \max_{\beta^1, \beta^2}\{X - \frac{1}{2}\theta_1 a_1^2 X - \frac{1}{2}\theta_2 a_2^2 X + f'(X)X\left((\mu_{11} + \mu_{21})a_1 + (\mu_{12} + \mu_{22})a_2\right) +$$

$$+\frac{1}{2}f''(X)X^2(4\sigma_c^2 + \sigma_1^2 + \sigma_2^2) - \frac{1}{2}\gamma_1 r\{\sigma_c^2(\beta^{11} + \beta^{12})^2 + (\beta^{11}\sigma_1)^2 + (\beta^{12}\sigma_2)^2\}$$

$$-\frac{1}{2}\gamma_2 r\{\sigma_c^2(\beta^{21} + \beta^{22})^2 + (\beta^{21}\sigma_1)^2 + (\beta^{22}\sigma_2)^2\}\}$$

Consider agent $i = 1$, differentiating the right-hand side of the ODE with respect to β^{11}

and β^{12}, the following system of equations determines their values.

$$\begin{pmatrix} \mu_{11}^2 + \gamma_1 r\theta_1 X(\sigma_c^2 + \sigma_1^2) & \mu_{21}\mu_{11} + \gamma_1 r\theta_1 X\sigma_c^2 \\ \mu_{11}\mu_{21} + \gamma_1 r\theta_1 X\sigma_c^2 & \mu_{21}^2 + \gamma_1 r\theta_1 X(\sigma_c^2 + \sigma_2^2) \end{pmatrix} \begin{pmatrix} \beta^{11} \\ \beta^{12} \end{pmatrix}$$

$$= \begin{pmatrix} f'(X)X(\mu_{11} + \mu_{21})\mu_{11} \\ f'(X)X(\mu_{11} + \mu_{21})\mu_{21} \end{pmatrix}$$

Therefore, the solutions are

$$\beta^{11} = D_1 \cdot \big((\mu_{11} - \mu_{21})\sigma_c^2 + \mu_{11}\sigma_2^2\big)$$

$$\beta^{12} = D_1 \cdot \big((\mu_{21} - \mu_{11})\sigma_c^2 + \mu_{21}\sigma_1^2\big)$$

where $D_1 = \frac{1}{|H|} \cdot f'(X)X(\mu_{11} + \mu_{21})\gamma_1 r\theta_1 X$ with H being a positive semi-definite matrix.

Say that the incentive scheme is *Joint Performance Evaluation* (JPE) if $\beta^{ij} > 0$ and *Relative Performance Evaluation* (RPE) if $\beta^{ij} < 0$. Depending on parameter values, there are a few notable and interesting cases. The analysis below apply to both agents due to symmetry.

- If $\tilde{\mu}_{21} = 0$ and $\sigma_c = 0$, there is no interaction among agents. The agents are contracted "separately":

  $$\beta^{11} = \frac{f'(X)X}{1 + \gamma_1 r\theta_1 X(\frac{\sigma_1}{\mu_{11}})^2} \text{ and } \beta^{12} = 0$$

  incentive loading to the agent's task goes back to the single-agent case. As $\beta^{12} = 0$, the agent is not exposed to any risk of the task that she is not responsible for.

- If $\tilde{\mu}_{21} = 0$, then $\beta^{11} > 0$ and $\beta^{12} < 0$. Agent 1 does not help for task 2. With the presence of common shock, a high dA_{2t} implies a_2 is high, the principal would infer that a_1 is not high enough and hence penalize agent 1 by lowering her continuation value.

- If $\sigma_c = 0$, then $\beta^{11} > 0$ and $\beta^{12} > 0$. With no common shock, it is difficult for the principal to infer the efforts of both agents. As the agents' efforts are productive in both tasks, the principal rewards agent 1 when the output in task 2 is high.

Note that the sign of β^{i}'s are independent of the degree of risk aversion. Therefore, it is not the risk attitude of the agents that affect the use of evaluation scheme. Instead, it is the underlying technological structure that drives the nature of contracts. In particular, the key condition for a relative performance scheme, $\beta^{12} < 0$, is that

$$\frac{\sigma_c^2 + \sigma_1^2}{\tilde{\mu}_{11}} < \frac{\sigma_c^2}{\tilde{\mu}_{21}} \text{ or } \sigma_c^2 > \sigma_1^2 \frac{\tilde{\mu}_{21}}{\tilde{\mu}_{11} - \tilde{\mu}_{21}}$$

The condition states that the common shock has to be large enough so that the principal will be able to use the differences in productivity to infer the differences in individual efforts. Furthermore, in the case of $\tilde{\mu}_{21} = 0$, I obtain the following closed-form solutions for the incentive loadings.

$$\beta^{11} = \frac{f'(X)X\mu_{11}^2(\sigma_c^2 + \sigma_2^2)}{\mu_{11}^2(\sigma_c^2 + \sigma_2^2) + \gamma_1 r\theta_1 X((\sigma_c^2 + \sigma_1^2)(\sigma_c^2 + \sigma_2^2) - \sigma_c^4)}$$

$$\beta^{12} = \frac{-f'(X)X\mu_{11}^2\sigma_c^2}{\mu_{11}^2(\sigma_c^2 + \sigma_2^2) + \gamma_1 r\theta_1 X((\sigma_c^2 + \sigma_1^2)(\sigma_c^2 + \sigma_2^2) - \sigma_c^4)}$$

2.4 Capital Structure

In this section, I embed the dynamic agency problem into Leland's (1994) capital structure model. Following Leland, I assume the firm's cash flows evolve according to a geometric Brownian motion

$$dX_t = (\phi + \sum_{i=1}^{n} a_{it})X_t dt + \sigma X_t dZ_t$$

where ϕ and σ are constants. In other words, $m = d = 1$, $\mu(a) = (\phi + \sum_{i=1}^{n} a_{it})$ and there is a single aggregate output. The parameter ϕ captures the baseline growth rate and is independent of agents' efforts. By exerting effort, the team pushes up the firm growth rate and the effort cost is assumed to be $h_i(a_i, X) = \frac{1}{2}\theta_i a_i^2 X$.

In addition, I assume that in multi-agent firms, the agents split the task equally on its single-agent counterpart in the sense that the drift of the cash flows are the same in a single and a multi-agent firm without agency cost. Intuitively, one can think of a firm that consists of different production segments, or divisions. Each segment produces a small part of a final product. Investors are not able to observe divisional profits, but they do observe the firm aggregate profits. Hence it make sense to assume a single-dimensional cash flows process. Technically, let $\bar{a}$ remains to be the upper bound of efforts in a single-agent firm. Now let $\bar{a}(n) = \frac{\bar{a}}{n}$ so that when all agents exert the highest effort, the expected outputs in firms with different number of agents are the same. Moreover, to focus on how the number of agents affect firm's capital structure, investment policy, default policy, I assume that agents are symmetric in the sense that $\gamma_i = \gamma$, $\theta_i = \theta$, and $\bar{a_i} = \bar{a}(n)$ for all $i = 1, ..., n$.

2.4.1 Optimal Contracting in an Unlevered Firm

Before introducing debt into the firm, I apply the optimal contracting result in section 3 to an unlevered firm. To implement effort $a_t = (a_{1t}, ..., a_{nt})$, by lemma 2.4, the appropriate incentive loadings are $\beta_{it} = \theta a_{it} X_t$ for all i. As the agents are symmetric, the equilibrium efforts are the same so that $a_i(\cdot) = a(\cdot)$ for all i. Then HJB equation becomes

$$rf_n(X) = \max_{a \in [0, \bar{a}(n)]} \left\{ X - \frac{n}{2}\theta a^2 X + f_n'(X)(\phi + an)X + \frac{1}{2}f_n''(X)X^2\sigma^2 - \frac{1}{2}n\gamma r\theta^2 a^2 X^2 \right\} \tag{2.4.1}$$

where $f_n(X)$ denote the value of the shareholder when she contracts with n agents. Optimal effort is given by

$$a_t^*(n) = \min\left(\frac{f_n'(X_t)}{\theta(1 + \theta\gamma r\sigma^2 X_t)}, \frac{\bar{a}}{n} \right) \tag{2.4.2}$$

The optimal effort function displays a similar form as in He's (2011) model with a single agent. However, the optimal effort in my model depends on the number of agents because the team size potentially affect shareholder's value. In order to highlight the differences, let's examine equation (**??**). Since agents are identical, the parameter n enters directly into equation (**??**). First, the term $anf_n'(X)X$ captures the total effect of effort a on firm's value. With more agents in the firm, the marginal effect of effort is higher due to agents' cooperation. Second, the number of agents also affect the cost of operating the firm. The cost is captured by the second term and the last term in the bracket of equation (**??**) and they reflect the effort cost and the incentive cost respectively. Intuitively, both terms should be increasing in n. However, the forthcoming analysis shows that when the principal chooses the efforts optimally, both the equilibrium effort costs and incentive costs are decreasing in n.

2.4.2 Optimal Contracting in a Levered Firm.

Next, I introduce debt into the firm. Following Leland (1994) I focus only on the consol bond with a constant coupon rate C. Shareholders can default when the firm is in financial distress. An implicit timing assumption is maintained here. First, the principal chooses the coupon rate to maximize the firm value. Second, given the amount of debt outstanding, the principal contracts optimally with the agents and determines the default policy. In the default event, the principal renders the firm to the debt holders. The debt holders then run the firm as an unlevered firm using the same set of agents. By the logic of backward induction, I first solve for the effort and default policy given a fixed amount of coupon. Then taking these policy functions as given, I let the principal chooses the optimal coupon.

Equity Value and default policy.

As in the case of an unlevered firm, the shareholders' value function is $F^E(X, W_1, ..., W_n) = f_n^E(X) + \frac{n}{\gamma r}\ln(-\gamma r W)$, where $f_n^E(X)$ denote the value of a levered firm, or equity value. Notice that the team value $\sum_{i=1}^{n} -\frac{1}{\gamma r}\ln(-\gamma r W_i)$ becomes n times the agent's certainty equivalent because in equilibrium, every agent exerts the same amount of effort and they face the same incentives. With a single aggregate output, their continuation value evolves exactly in the same way. Hence $W_i = W$ for all i.

The HJB equation that characterizes the firm's equity value with n agents is

$$rf_n^E(X) = \max_{a \in [0, \frac{\bar{a}}{n}], X_B} \{X - (1-\tau)C - \frac{n}{2}\theta a^2 X$$

$$+f_n^{E'}(X)(\phi + an)X + \frac{1}{2} f_n^{E''}(X) X^2 \sigma^2 - \frac{1}{2} n\gamma r \theta^2 a^2 X^2\} \tag{2.4.3}$$

where C is the coupon rate and τ is the corporate tax rate. X_B is the default boundary to be chosen optimally. Equation (??) parallels equation (??) except that the latter has additional cash outflow $(1-\tau)C$. Optimal effort is given by

$$a_t^*(n) = \min\left(\frac{f_n^{E'}(X_t)}{\theta(1+\theta\gamma r\sigma^2 X_t)}, \frac{\bar{a}}{n}\right) \tag{2.4.4}$$

which is similar to equation (??). Next, the optimal default policy is characterized by the value-matching condition

$$f_n^E(X_B(n)) = 0 \tag{2.4.5}$$

and the smooth-pasting condition

$$f_n^{E'}(X_B(n)) = 0 \tag{2.4.6}$$

Both conditions are standard [1] and they pin down the default boundary X_B as a function of n.

Debt Value and Leverage Ratio.

The debt holders rationally expect that the shareholders and the team sign an optimal contract, and hence they know the implemented effort policy $a^*(n)$ and the firm's default

[1] See Dixit (1994) for relevant discussions.

policy $X_B(n)$. The value of debt $D_n(X)$ is a solution to the following ODE

$$rD_n(X) = C + D_n^{'}(X)(\phi + na^*(n;X))X + \frac{1}{2}D_n^{''}(X)\sigma^2 X^2 \tag{2.4.7}$$

with two boundary conditions: (i) $D_n(X_B) = (1-\alpha)f_n(X_B(n))$ where $\alpha < 1$ is the percentage bankruptcy cost, and (ii) $D_n(X) \to \frac{C}{r}$ as $X \to \infty$. The first condition states that the debt value at default boundary is the value of an unlevered firm less the bankruptcy cost. This is because, at default event, the debt holders take over the firm and run the firm themselves. As in Leland, the debt holders cannot issue new debts. The second condition states that as firm size grows, default is unlikely and the debt holders receive the discounted value of coupon payment.

The principal choose coupon C to maximize the total firm value with leverage, given the initial size of the firm X_0 , the number of agents n, and the investment policies. Hence the optimal coupon $C^*(X_0, n)$ is defined as

$$C^*(X_0, n) = \arg\max_C \left\{ f_n^E(X_0; C) + D_n(X_0; C) \right\}$$

Given the optimal choice of coupon payment, I can compute the optimal leverage ratio as

$$LR(X_0; n) \equiv \frac{D_n(X_0; C(X_0, n))}{f_n^E(X_0; C^*(X_0, n)) + D_n(X_0; C^*(X_0, n))}$$

In order to directly compare my results with He (2011), I adopt the parameter values as in his numerical analysis. Interest rate $r = 5\%$, baseline growth rate $\phi = -0.05$, effort cost $\theta = 35$, degree of risk aversion $\gamma = 10$, volatility $\sigma^2 = 6.25\%$, marginal tax rate $\tau = 20\%$, bankruptcy cost $\alpha = 0.25$, upper bound effort $\bar{a} = 0.05$, initial firm size $X_0 = 20$.[2]

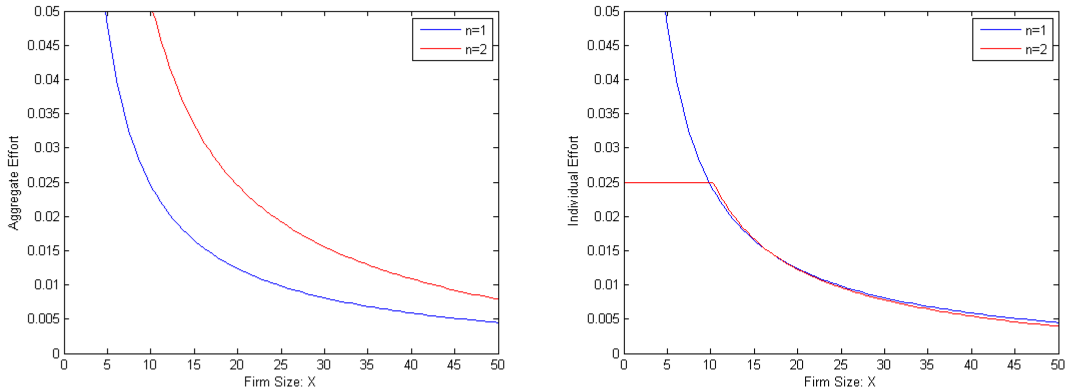

Figure 2.4.1: **Effort policy in unlevered firms.**
The red line is for $n = 2$ and the blue line is for $n = 1$. Left panel: Aggregate effort. Right panel: Individual effort

2.4.3 Discussion.

2.4.3.1 Effort policy in an unlevered firm.

Figure 4.1 plots the effort policy in unleverled firms as a function of the firm size. The red line is the effort when the firm has two agents ($n = 2$) and the blue line is for a single-agent firm ($n = 1$). The left panel shows the aggregate effort and the right panel plots the individual effort. Efforts in both firms are decreasing with respect to firm size. This is because the increasing cost associated with risk compensation. For the individual efforts, the initial effort in a two-agent firm is lower than the firm with one agent. This happens because the agents split the task. As the firm grows, the effort profiles overlap. This indicates in unconstrained region, the principal in different firms would like to assign roughly the same amount of effort investment. In terms of aggregate effort, the efforts in a two-agent firm is thus higher than its single-agent counterpart.

2.4.3.2 Effort Policy in levered firms.

Figure 4.2 plots the effort policy in levered firms as a function of firm size. The red line is for two-agent firm and the blue line for a single-agent firm. The figure displays a few

[2] For more discussions about parameterization, please see He (2011).

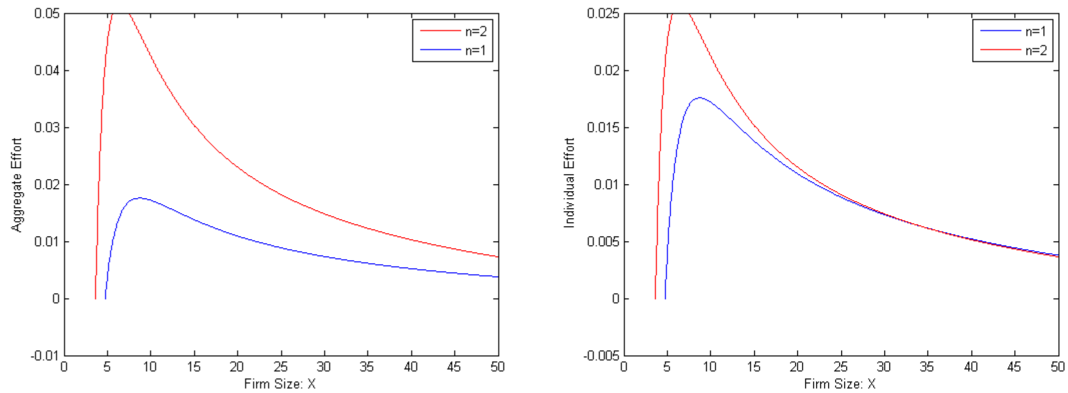

Figure 2.4.2: **Effort Policy in levered firms**.
The red line is for $n = 2$ and the blue line is for $n = 1$. The left panel: aggregate effort. The right panel: individual effort.

features. First, in both firms, the effort has a bell shape. In contrast with unlevered firms, as the firm size decreases, effort level will drop and eventually reach zero. This happens because in levered firms, when firm size is small, the firm is close to the default boundary. The principal provides less incentives as she expects the firm will default with high probability. That is, the marginal value of effort is too low in that region. Second, in contrast to unlevered firms, the individual efforts in two-agent firms is much higher than its single-agent counterpart when firm size is small. To understand this result, note that as the firm hires more agents, these agents split the task. In particular, the effort cost and the incentive cost are convex function. The principal can therefore allocate the efforts to different agents and thus smooth out the costs. Convexity implies by doing so, the total cost will be lower. This in turn implies the principal will choose higher effort investment in multi-agent firms.

This effect helps to alleviate the debt overhang problem as discussed in He (2011). Recall that with the presence of debts, the debt overhang effect tends to lower effort provisions. This is because effort is a form of investment, the principal pays the costs, but the possibility of default implies part of the investment returns will go to the debt holders. Now, the principal could mitigate the debt overhang problem by smoothing out

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
Coupon	8.969	7.361	9.399	10.751	11.625
Leverage Ratio	0.537	0.469	0.551	0.594	0.617
First-Best leverage ratio	0.627	0.627	0.627	0.627	0.627
Default Policy	4.73	3.625	4.523	5.078	5.407
Equity Value	107.960	134.397	121.727	115.249	112.451
Debt Value	124.945	119.027	149.262	168.881	181.397
Levered Firm Value	232.905	253.424	270.989	284.130	293.848
Unlevered Firm Value	219.476	236.334	249.925	260.476	268.429

Table 2.1: **Capital Structure and default policy with different number of agents.**

the investment cost in a multi-agent firm. Therefore, more effort investment is implied.

2.4.3.3 Capital Structure.

Table 1.1 reports the optimal debt policy, equity and debt value, and default policy of levered firms. The first-best leverage ratio is computed using equation (24) in Leland (1994).

Default Policy.

In a levered firm with two agents, the firm has a smaller default boundary ($\delta_B = 3.63$) than the firm with a single agent ($\delta_B = 4.73$). Consistent with figure 4.2, the effort vanishes earlier with less agents. The intuition is that, in a firm with two agents, the production technology is more valuable to the shareholders. Therefore it is costly for such a firm to default and thus default will tend to be delayed. However, as the firm continues to hire more agents, the default threshold increases and eventually exceeds the single-agent boundary.

Coupon and leverage ratio.

Turning to the capital structure, the firm with two agents issues a smaller amount of coupon ($C = 7.36$) than the single-agent firm ($C = 8.97$). The intuition of the result also follows from the debt overhang analysis. As point out in He (2011), to implement a high effort profile, the principal cannot issue too much debt as debt will "crowd-out" effort investment due to debt overhang. The mechanism plays the same role here. In the previous analysis of effort invesment, we have seen a two-agent firm will implement a higher effort profile. Therefore, in order not to crowd-out effort investment, less debt is issued in the first stage.

However, as the firm hires more agents, debt overhang becomes a less important issue. With more and more agents, the total cost to implement a certain effort profile will be lower because the principal could spread the costs across agents. Moreover, the marginal returns of effort investment will also be higher. This implies in firms with more agents, the firm endogenous growth rate is higher. As the firm is able to produce more cash flows, the principal worries less about insolvent state and bankruptcy cost. Therefore, the principal will be able to issue more debts to take advantage of the tax shield.

This analysis also shows that the optimal leverage ratio is non-monotonic in the number of agents. First, with a smaller amount of coupon payment, the debt value is lower. But as the firm issues a larger amount of coupon, debt values increases again. Second, the equity value moves in an opposite direction because the equity value is shifted to the debt holders as the coupon varies. However, the levered firm value increases strictly with the number of agents because with more agents, the firm's growth rate is higher. The combined effect is that the shape of the leverage ratio follows exactly the pattern the debt issuance.

2.5 Conclusion

This paper generalizes He's (2011) model to a multi-agent set-up and studies the model implications in corporate finance. The absence of wealth effect in exponential utility deliver allows me to simplify the optimal contracting problem by disentangling the firm's value into the principal's value and the team's value, where the team's value is the summation of the certainty equivalent of all the agents working inside the firm. Due to this separation, I characterize the optimal contract by an ODE and study the contractual implications on compensation.

I further embed the agency model in Leland (1994). The application allows me to endogenize the expected firm's growth rate in his classical geometric Brownian cash flows set-up. Due to debt-overhang problem, there is under-investment in efforts and the optimal leverage ratio is lower than the first best. A new implication is that the optimal leverage ratio is non-monotonic in the number of agents. As a single-agent firm hires the second agent, the leverage ratio drops sharply. However, the leverage ratio increases monotonically as more agents are added into the firm afterwards. This interesting implication can be further tested empirically.

Chapter 3

Multi-armed Bandit Problems with Ambiguity

3.1 Introduction

The multi-armed bandit problem is a statistical decision problem in which a decision-maker experiments different choices and optimizes his decision based on the trade-off over returns and information. As a concrete example, consider a doctor prescribing drugs to patients. Patients consult the doctor sequentially, and the doctor has two different types of drugs that she can assign to the patients. A newly-invented drug of which the effects on patients is not completely known. An old drug with known effects. By assigning the new drug, the doctor is able to learn more about its effect and gain more information, but the patient may suffer losses because of the potential side effect of the drug. By treating the patient with the old drug, the doctor learns nothing new, but she is certain about the return. The drugs in the example are "bandit processes". The doctor can assign one and only one treatment to each patient. The trade-off facing the doctor involves information (experimenting with the new drug) and return (certain payoff from the old drug).

The classic result of the multi-armed bandit problem is the so-called Gittins Index Theorem. The decision-maker will assign an "Index", which takes into account of the trade-off between exploration and exploitation, to each of the arm, and play the arm with the highest index at any point of time. The classic analysis assumes that the decision-maker has probabilistic beliefs. That is, she is assumed to know how returns are generated and the objective probability law that governs the state process. However, the assumption is too strong in a lot of contexts. In the drug prescription example, if the patient dies after taking the new drug, can the doctor be sure that the patient dies because of the drug but not due to other medical factors? Information, in a lot of environments, is too imprecise to be summarized by a single probability law. Instead, the decision-maker could have in mind a multiple probability law when she looks at the world. The situation is often referred as *ambiguity.* In this paper, I extend the multi-armed bandit problem to situations where the decision-maker does not have a single complete theory about the world, i.e., she faces ambiguity.

To incorporate ambiguity, I adopt the recursive multiple-priors utility model developed by Epstein and Wang (1994). In that model, the decision-maker is assumed to have a set of beliefs over the state of nature. In particular, beliefs are modeled by a collection of sets of one-step ahead conditionals. The set of one-step-ahead conditionals captures the degree of ambiguity aversion. The nice feature of modeling beliefs in this way is that when we paste, by way of backward induction, the collection of the set of one-step-ahead conditionals together, we obtain a set of priors which is rectangular and the decision-maker's behavior is dynamically consistent.

I therefore take a standard multi-armed bandit problem and generalize the decision-maker's preferences, from subjective expected utility in the Bayesian literature, to recursive multiple-priors utility, in order to incorporate ambiguity. In deriving the optimal arm pulling strategy, I adopt Whittle's (1980) retirement option approach. The approach is divided into two parts. First, an artificial retirement option is introduced and we study an optimal stopping problem when the decision-maker faces a single bandit process versus this retirement option. By analyzing the stopping problem, we can derive the Gittins Index, which is the return that the decision-maker obtains when she pulls the arm optimally and expects to stop and retire in some states in the future. After we obtain the Gittins Index for each of the arm, the second part of the analysis shows that if the decision-maker follows the index strategy, her optimal value will satisfy the Bellman equation for the entire bandit process and hence by the uniqueness of the value function, we can establish the optimality of the index strategy.

Both steps of the analysis carry over to the situation with ambiguity. This is because

Whittle's argument is essentially a dynamic programming analysis. Recursive multiple-priors utility allows us to use dynamic programming. In handling the optimal stopping problem in the first step, I borrow techniques developed in Miao and Wang (2011). Their paper introduces Epstein-Wang's type ambiguity into a standard real-option model and provides a neat formulation of optimal stopping problems with ambiguity. The main result shows that the classic Gittins Index Theorem generalizes to Multiple-Priors Index Theorem (proposition 3.3). The generalization comes from the fact that when the decision-maker evaluates an ambiguous arm, instead of computing the Gittins Index, she computes a Multiple-Priors (MP) Index. The MP-Index takes into account of a set of probability laws that the decision-maker has in mind. As the decision-maker is ambiguity averse, she evaluates arms using the worst-case measure and the MP-Index is thus the lower envelope of the Gittins Index (proposition 3.2).

The literature on multi-armed bandit problem is huge. Gittins and Jones (1974) prove the celebrated Gittins Index Theorem. Banks and Sundaram (1992) generalizes the theorem to countable arms. Gittins, Glazer and Brook (2010) is a comprehensive monograph on bandit problem. For application in economics, see Bergemann and Valimaki (2006). There are many studies on preferences under ambiguity in decision theory. Gilboa and Schmeidler (1989) provide an axiomatic foundation of the static multiple-priors utility. Epstein and Wang (1994) and Epstein and Schneider (2003) bring the multiple-priors utility to dynamics. The former paper formulates the recursive multiple-priors utility and applies it to asset pricing while the latter axiomatizes the utility model. To the best of my knowledge, my work is the first to study ambiguity and the multi-armed bandit problem in an integrated framework. I contribute to theliterature by proving a version of Gittins Index theorem when thedecision-maker is ambiguity averse and provide a stepping stone for further applications of the bandit problem with ambiguity.

3.2 The Model

Consider an infinite horizon bandit problem. The primitive is $(\mathcal{K}, (A_i)_{i\in\mathcal{K}}, \beta)$. There are K independent arms and the set of arms is denoted by $\mathcal{K} = \{1, ..., K\}$, with generic element i. Each arm i can be described by a tuple $A_i = (X_i, R_i, \mathcal{P}_i^{+1})$. X_i denote the state space for arm i, it is a complete and separable metric space with Borel σ-algebra $\mathcal{B}_i$. Let $\Delta(X_i)$ be the set of all Borel probability measures, endowed with the topology of weak convergence. The reward function is $R_i : X_i \to \mathbb{R}$. Thus the reward $R_i(x_i)$ is what the decision-maker obtains when he plays arm i at state x_i. I assume that $R_i's$ are uniformly bounded, continuous, and $\mathcal{B}_i$-measurable.

$$\exists C \text{ s.t. } |R_i(x_i)| \leq C \ \forall x_i \ \forall i$$

When the decision-maker observes that arm i is in state x_i, and if the arm is played, the state x_i evolves to state $x_i^{'}$. Beliefs about how the state evolves are ambiguous. We model beliefs about the state evolution by a *probability kernel correspondence* $\mathcal{P}_i^{+1} : X_i \rightsquigarrow \Delta(X_i)$, which is a nonempty, continuous, compact-valued, and convex-valued correspondence. For each $x_i \in X_i$, we think of $\mathcal{P}_i^{+1}(x_i)$ as the set of beliefs about the next period's state, if arm i is activated. In particular, each $p_i \in \mathcal{P}_i^{+1}(x_i)$ is a transition probability that satisfies (i) for all $x_i \in X_i$, $p_i(\cdot, x_i)$ is a measure on X_i, and (ii) for any $B \in \mathcal{B}_i$, $p_i(B, \cdot) : X_i \to [0, 1]$ is $\mathcal{B}_i$-measurable. Therefore, we view $\{\mathcal{P}_i^{+1}(x_i)\}$ as the set of one-step-ahead conditionals. Finally, the decision-maker discounts the future rewards geometrically, using the discount factor $\beta \in [0, 1)$.

Denote the period state space by $X = \mathop{\times}\limits_{i=1}^{K} X_i$, endowed with the product topology. In obvious notations, $x_t = (x_{1t}, ..., x_{Kt}) \in X$ refers to period t state of all arms. In each period t, the decision-maker observes x_t, she must decide which arm to be played in that period. Let the choice at time t be $a_t \in \mathcal{K}$. If $a_t = i$, arm i is played at t, its state evolves from x_{it} to x_{it+1}. If arm j is not pulled, $a_t \neq j$, its state remains frozen, thus $x_{jt} = x_{jt+1}$. A t-history for the bandit is a record of the arm played up to t, states and rewards observed. Formally, write the *t-history* as $h_t = (a_0, x_0, R_{a_0}, ..., a_{t-1}, x_{t-1}, R_{a_{t-1}}, x_t)$, where x_0 is the initial state of the problem. We can write the set of all t $\mathcal{H}_t = H^t \times X$, where $H = \mathcal{K} \times X \times \mathbb{R}$. Its Borel σ-algebra is $\mathcal{B}_t$. A *strategy* $a = \{a_t\}_{t=0}^{\infty}$ is a specification of the arm to be played in each period, depending on the information accumulated up to t. Specifically, for any t, $a_t : \mathcal{H}_t \to \mathcal{K}$ is a $\mathcal{B}_t$-measurable function. Let $\mathcal{A}$ denote the set of all strategies.

To accommodate ambiguity, I adopt the recursive multiple-priors utility model[1] in evaluating strategies. An important property of the utility function is dynamic consistency. To ensure the property is satisisfied, at any state x_t, the set of priors $\mathcal{P}(x_t)$ over X needs to be “rectangular”. The set of priors is constructed in the following way. Using the set of one-step-ahead conditionals $\{\mathcal{P}_i^{+1}(x_i)\}$ for each arm i, the set of arm i's prior $\mathcal{P}_i(x_{it})$ over X_i at any time t is defined by

$$\mathcal{P}_i(x_{it}) = \int \mathcal{P}_i(x_{it+1}) d\mathcal{P}_i^{+1}(x_{it}),\ x_{it} \in X_{it}$$

Epstein and Schneider (2003) show that when the set of priors is constructed by pasting marginals and foreign one-step-ahead conditionals, the prior set is rectangular. Any max-min decision-maker using such a set as her beliefs display dynamically consistent behavior.

Moreover, arms are stochastically independent. I adopt Gilboa and Schmeidler's notion of “stochastic independence”, define the set of priors $\mathcal{P}(x_t)$ at t as

$$\mathcal{P}(x_t) = \bar{co} \left\{ \mathop{\times}_{i=1}^{K} \mathcal{P}_i(x_{it}) \right\},\ x_t \in X$$

where $\bar{co}(\cdot)$ denote the closed and convex hull of a set. Note that $\mathcal{P}(x_t)$ is rectangular because closure operation preserves rectangularity. Finally, the utility delivered by a strategy a, given the initial state x_0, and the initial priors $\mathcal{P}(x_0)$, is the decision-maker's total

[1]For the axiomatic foundation of multiple-priors utility, see Gilboa-Schmeidler (1989) and Epstein-Schneider (2003).

discounted expected reward under this strategy

$$W(a;x_0) = \min_{p \in \mathcal{P}(x_0)} E_p(\sum_{t=0}^{\infty} \beta^t R_{a_t}(x_{a_t t}))$$

Notice that for any p, the expectation $E_p(\cdot)$ depends on the strategy a. The decision-maker's objective is to find an *optimal strategy* a^* that maximizes the expected rewards, namely, $W(a^*, x_0) \geq W(a, x_0)$ for all $a \in \mathcal{A}$. The value function for the multi-armed bandit problem under ambiguity can be written as

$$V(x_0) = \sup_{a \in \mathcal{A}} \min_{p \in \mathcal{P}(x_0)} E_p(\sum_{t=0}^{\infty} \beta^t R_{a_t}(x_{a_t t}))$$

The objective of this paper is to characterize the optimal strategy.

3.3 Main Results

Gittins and Jones (1974) prove that for multi-armed bandits when the decision-maker has *probabilistic beliefs* about the state evolution process, the optimal strategy coincides with an "index strategy" in the following way. Each arm i can be associated with a *Dynamic Allocation Index*, which depends on the current state on that arm. The index-type strategy requires the decision-maker to play the arm with the maximal index at each period. Moreover, this optimal strategy can be obtained by solving a family of stopping problems. These results are known as Gittins Index Theorem.

I expect that Gittins Index Theorem would extend to multi-armed bandit problem with ambiguity. The intuition of why the classic result carries over is that arms are *stochastically independent*–realizations of one arm do not affect the state of all the other frozen arms. Therefore, we expect ambiguity only affect how the indices are evaluated, but not the *form* of the optimal strategy.

This section is divided into two parts. In the first part, I solve for the index in the current set-up. In the second part, I characterize the optimal arm-pulling strategy.

3.3.1 The Multiple-Priors Gittins Index.

In this subsection, I consider a single bandit process. In deriving the Gittins Index under ambiguity, I adopt the solution procedure in Whittle (1980). Whittle considers a dynamic programming problem for a single-armed bandit problem and calculates the Gittins Index. In what follows, consider a single ambiguous bandit process and a retirement option. The retirement option can be viewed as a safe and constant arm, which state never change even if the decision-maker pulls it. Intuitively, since the state of the retirement option never changes, once the decision-maker switches from the ambiguous arm to the retirement option, the decision-maker "retires" and will continue to play this outside option forever. Hence, the decision-maker's problem essentially becomes an optimal stopping problem.

Suppose arm i is the ambiguous arm. Let $m \in \mathcal{M}=[-C', C']$ be the per-period terminal reward of the retirement option, where $C' = \frac{2C}{1-\beta}$. Let $V_i : X_i \times [-C', C'] \to \mathbb{R}$ be the value function of the problem. And let τ be a stopping time and $\mathcal{T}_t$ be the set of stopping times that starts at t. Then the decision problem when the initial state is x_{i0} is

$$V_i(x_{i0}, m) = \sup_{\tau \in \mathcal{T}_0} \min_{p_i \in \mathcal{P}_i(x_{i0})} E_{p_i}[\sum_{t=0}^{\tau-1} \beta^t R_i(x_{it}) + \beta^\tau m]$$

The value function satisfies the Bellman equation for the optimal stopping problem

$$V_i(x_i, m) = \max \left\{ m, R_i(x_i) + \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i, m) dp_i(\tilde{x}_i, x_i) \right\}$$

By standard argument, the existence of $V_i(\cdot, m)$ can be established for any terminal reward $m \in \mathcal{M}$. In fact, let $C(X_i)$ be the space of all real-valued continuous functions on X_i, endowed with the sup-norm topology. This renders $C(X_i)$ a complete metric space. Define an operator $T : C(X_i) \to C(X_i)$ by, for any $v \in C(X_i)$,

$$(Tv)(x_i) = \max \left\{ m, R_i(x_i) + \beta \min_{p_i \in \mathcal{P}_i(x_i)} \int_{X_i} v(\tilde{x}_i) dp_i(\tilde{x}_i, x_i) \right\}$$

Then we have

Proposition 3.1. *For any $m \in \mathcal{M}$, there exists a unique fixed point $V_i(\cdot, m)$ for the mapping $T : C(X_i) \to C(X_i)$.*

For the ambiguous arm i, define its *Multiple-Priors Gittins Index (MP-index)* at state x_i as

$$G_i(x_i) = \inf \{m \in \mathcal{M} | V_i(x_i, m) = m\}$$

For $m \geq \frac{C}{1-\beta}$, stopping immediately is optimal, we must have $V_i(x_i, m) = m$. For $m \leq -\frac{C}{1-\beta}$, continuation is optimal, we must have $V_i(x_i, m) > m$. As $V_i(\cdot, m)$ is a continuous function, it follows that $G_i(x_i)$ is well-defined and takes value in the compact set $[-\frac{C}{1-\beta}, \frac{C}{1-\beta}]$. The optimal stopping time is given by

$$\begin{aligned} \tau_i(x_i, m) &= \inf \{t : V_i(x_i, m) = m\} \\ &= \inf \{t : G_i(x_{it}) \leq m\} \end{aligned}$$

The analysis also shows that the optimal strategy in the single ambiguous arm bandit problem is to play the ambiguous arm continuously until its MP-index falls below the payoff the decision-maker obtains from switching to the retirement option.

Next, I provide an explicit form or *forward induction* characterization for the index. Define for arm i, $v_i : X_i \to \mathbb{R}$ by

$$v_i(x_i) = \sup_{\tilde{\tau} \in \mathcal{T}_{t+1}} v_i(x_i, \tilde{\tau}) = \sup_{\tilde{\tau} \in \mathcal{T}_{t+1}} \min_{p_i \in \mathcal{P}_i(x_{it})} \frac{E_{p_i}(\sum_{s=t}^{\tau-1} \beta^t R_i(x_{is}) | x_{it})}{E_{p_i}(\sum_{s=t}^{\tau-1} \beta^s | x_{it})}$$

where $v_i(x_i, \tau)$ is the minimum of the expected discounted reward per expected unit of discounted time, when the arm is played from the initial state x_i, for a positive duration $\tilde{\tau} > 0$. The index $v_i(x_i)$ is then the supremum over the set of all strictly positive stopping times. In fact, it can be shown that the supremum is achieved by a stopping time from state x_i

$$\tilde{\tau}_i(x_i) = \min \{t : v_i(x_{it}) < v_i(x_i)\}$$

Observe that $v_i(x_i)$ is well-defined since the reward function is bounded. Notice that $v_i(x_i)$ reduces to the original Gittins Index when beliefs about state evolution is probabilistic, namely, when $\mathcal{P}_i^{+1} = \{p_i^{+1}\}$ is a probability kernel function (so that $\mathcal{P}_i$ is a singleton). To simplify notation, let $F_i^{p,\tau}(x_{it}) = \frac{E_p(\sum_{s=t}^{\tau-1} \beta^t R_i(x_{is})|x_{it})}{E_p(\sum_{s=t}^{\tau-1} \beta^s|x_{it})}$.

Proposition 3.2. ***(Characterization)*** *(i) The Multiple-Priors Index admits the "forward induction" characterization: For any arm i and x_i, $G_i(x_i) = (1-\beta)v_i(x_i)$ and $\tilde{\tau}_i(x_i) = \tau_i(x_i, G_i(x_i))$.*

(ii) Minimax identity: The Multiple-Priors Index process is the lower envelope of the Gittins Index:

$$G_i(x_{it}) = \inf_{p \in \mathcal{P}_{it}} G_i^p(x_{it})$$

where $G_i^p(x_{it}) = \sup_{\tau \in \mathcal{T}_{t+1}} F_i^{p,\tau}(x_{it})$ is the (Bayesian) Gittins Index.

The first part of the proposition characterizes the MP-Index in terms of the worst-case return-to-cost ratio. With the scaling factor $1-\beta$, it also relates the flow value from playing the ambiguous arm $G_i(x_{it})$ to its stock value $v_i(x_{it})$. Since at any time t and state x_{it}, $G_i(x_{it}) = (1-\beta)v_i(x_{it})$ and the decision-maker continues to play the ambiguous arm as long as $G_i(x_{it}) > m$, the proposition implies that the decision-maker will play the ambiguous arm when

$$v_i(x_i) = \sup_{\tilde{\tau} \in \mathcal{T}_{t+1}} \min_{p_i \in \mathcal{P}_i(x_{it})} \frac{E_{p_i}(\sum_{s=t}^{\tau-1} \beta^t R_i(x_{is})|x_{it})}{E_{p_i}(\sum_{s=t}^{\tau-1} \beta^s|x_{it})} > \frac{m}{1-\beta}$$

In words, when the worst-case return-to-cost ratio is higher than the stock value $\frac{m}{1-\beta}$, the decision maker would not stop experimenting with the ambiguous arm.

The second part of the proposition, the minimax identity provides a convenient way to compute the MP-Index. It states the one can interchange the min and max operation in the definition of $v_i(x_{it})$. In particular, the MP-Index can be obtained by taking the worst case measure over a set of the classical Gittins Index. In the statistics literature, there are numerical procedures for computing the Gittins Index. The result here suggests that to compute the MP-Index, one can first calculate the Gittins Index associated with each probability measure and then take the minimum of the set of Gittins Index.

In terms of behavioral implication, the minimax identity implies that a Bayesian decision-maker, when her probabilistic belief is contained in the set of priors of an ambiguity averse decision-maker, will be more willing to experiment than the decision-maker who views the arms as ambiguous. [2] Therefore, in general, an ambiguity averse decision-maker will experiment less and acquire less information.

3.3.2 The Multiple-Priors Gittins Index Theorem

In this section, I characterize the optimal strategy in the multi-armed bandit problem under ambiguity. Proposition 3.3, which is the main result of this chapter, establishes the optimality of index strategy.

Proposition 3.3. ***(Multiple-Priors Index Theorem):*** *For the multi-armed bandit problem $(\mathcal{K}, (A_i)_{i\in\mathcal{K}}, \beta)$ with retirement option m, the uniquely optimal strategy is the index strategy: (a) If $m \geq G_i(x_i)$ for all $i \in \mathcal{K}$, retire. (b) Otherwise, play arm k for which $G_k(x_k) = \max_{i\in\mathcal{K}}\{G_i(x_i)\}$, except possibly at history with probability zero.*

The index strategy certainly is well-defined given finitely many arms. Observe that the optimal strategy for the multi-armed bandit *without* the retirement option can be obtained by setting $m \leq -\frac{C}{1-\beta}$, so that the retirement for the entire bandit process is not optimal. In that case, $V(x,m) = V(x)$.

The Multiple-Priors Index Theorem is a generalization of the Gittins Index Theorem (1974) since the MP-Index generalizes the Gittins Index. Should the decision-maker have probabilistic beliefs, the predictive content of the MP-Index Theorem will be the same as the classic result.The intuition of why the MP-Index theorem holds is that the arms are

[2] That is, consider the case: $G_i^p(x_{it}) > m \geq G_i(x_{it})$.

stochastically independent and that the MP-Index provides a sensible way to evaluate the return of pulling an arm in a single-ambiguous arm stopping problem.

3.4 Conclusion

This paper studies multi-armed bandit problem under ambiguity. The decision-maker views the stochastic processes underlying arms as ambiguous and she has ambiguous beliefs about how the state evolves. The decision-maker is assumed to have recursive multiple-priors utility. I show that the classic results in the bandit literature, the Gittins Index Theorem and the characterization of Gittins Index, generalize to the case with ambiguity. Due to ambiguity aversion, the decision-maker will experiment less and acquire less information about the bandit processes in general.

Future research can investigate whether the index theorem depends on the way we model preferences under ambiguity. For example, variational preferences (Maccheroni, Marinacci, and Rustichini (2006)) generalizes multiple-priors utility. Does the theorem still hold under such preferences? If it holds, one may conjecture that the specific utility model may only affect how the Gittins Index is computed, but not the index-nature of the optimal strategy. To begin with, it may be useful to first study Hansen-Sargent (2001) utility, which is a subclass of variational preferences.

Appendix A

Appendix to Chapter 1

Proof of Proposition 1.1.

Given any contract $\Gamma = (I, \tau, a)$, define the agent's lifetime expected utility conditional on the information available at time t,

$$V_t(\Gamma) = E_t^a \left[\int_0^\tau e^{-\rho s}(dI_s + 1_{\{a_s^e = a_L\}} B ds) \right]$$
$$= \int_0^t e^{-\rho s}(dI_s + 1_{\{a_s^e = a_L\}} B ds) + e^{-\rho t} W_t(\Gamma, a)$$

where $W_t(\Gamma)$ is defined in (**??**). The process $\{V_t(\Gamma)\}$ is a $\mathcal{F}_t$-martingale. Moreover, define a compensated Poisson process $M^a = \{M_t^a\}_{t \geq 0}$ by $M_t^a = N_t - \int_0^t \lambda(a_s) ds$ for all $t \geq 0$. Similarly, we have $Z^a = \{Z_t^a\}_{t \geq 0}$, where $dZ_t^a = dX_t - \mu(a_t)dt$ for all $t \geq 0$. Note that Z^a and M^a are $\mathcal{F}_t$-martingale under measure P^a. Applying the martingale representation theorem, there exists a $\mathcal{F}_t$-predictable process $\{(\beta_t, \psi_t)\}_0^\tau$ such that for all $t < \tau$, $V_t(\Gamma) = V_0(\Gamma) + \int_0^t e^{-\rho s} \beta_s \sigma dZ_s^a - \int_0^t e^{-\rho s} \psi_s dM_s^a$. Taking derivative with respect to time for both representation of $V_t(\Gamma)$, and combining, we obtain (**??**).

As for incentive compatibility of (Γ, a). Fix any action process a with $a_t = (a_H, a_N)$ at t. Let $\tilde{a}$ denote any other action process. First I prove sufficiency. Suppose the agent follows a. Then $\{V_t(\Gamma)\}_0^\tau$ is a martingale since by differentiating the definition of $V_t(\Gamma)$, $dV_t(\Gamma) = e^{-\rho t} dI_t - \rho e^{-\rho t} W_t(\Gamma) dt + e^{-\rho t} dW_t(\Gamma)$. And substituting (**??**), we have $dV_t(\Gamma) = e^{-\rho t}(\beta_t(dX_t - \mu dt) - \psi_t(dN_t - \lambda dt))$. If the agent deviates at t to $\tilde{a}_t = (\tilde{a}_L, \tilde{a}_N)$, $dV_t(\Gamma) = e^{-\rho t}(B - \beta_t \mu)dt + e^{-\rho t}(\beta_t dZ_t^{\tilde{a}} - \psi_t(dN_t - \lambda dt))$. Since $\beta_t \geq \frac{B}{\mu}$, the drift of V_t is nonpositive.

If the agent deviates at t to $\tilde{a}_t = (\tilde{a}_H, \tilde{a}_R)$, $dV_t(\Gamma) = e^{-\rho t}(\beta_t\alpha - \psi_t\gamma)dt + e^{-\rho t}(\beta_t(dX_t - (\mu + \alpha)dt) - \psi_t(dN_t - (\lambda + \gamma)dt))$. Since $\psi_t\gamma \geq \beta_t\alpha$, again, the drift of V_t is nonpositive. Hence the above two types of deviation are suboptimal. Note that under $\beta_t \geq \frac{B}{\mu}$ and $\psi_t\gamma \geq \beta_t\alpha$, the drift of V_t will be nonpositive if the agent deviates to $\tilde{a}_t = (\tilde{a}_L, \tilde{a}_R)$. Now, for necessity. If either $\beta_t < \frac{B}{\mu}$ or $\psi_t\gamma < \beta_t\alpha$ is true on a set of positive measure during $[0, \tau)$, the agent will deviate from $a_t = (a_H, a_N)$ since V_t will be a submartingale by the previous argument. A similar argument establishes the incentive compatibility condition for a contract that assigns $a_t = (a_H, a_R)$ at t. ■

Lemma A.1. *Suppose W_t evolves according to*

$$dW_t = (\rho W_t + \pi_t\psi\lambda)dt - dI_t + \beta\sigma dZ_t - \pi_t\psi dN_t$$

in the interval $[0, W^p]$ until time $\tau = \min\{t : W_t = 0\}$, where $\pi_t \in \{0, 1\}$ is a controlled process. I_t is a nondecreasing process that reflects W_t at W^p. Let $\eta \in \mathbb{R}$, $g : [0, W^p] \times \{0, 1\} \to \mathbb{R}$ be a bounded function, and $\lambda(W_t, \pi_t)$ be the intensity process of a Poisson process N_t. Then $f : [0, W^p] \to \mathbb{R}$, with $f \in C^2$, solves the following differential equation

$$rf(W) = \sup_{\pi} \{g(W, \pi) - \lambda(W, \pi)L + \mathcal{L}^{\pi} f(W)\} \tag{A.0.1}$$

where $\mathcal{L}^{\pi} f(W) = f'(W)(\rho W + \pi\psi\lambda(W, \pi)) + \frac{1}{2}f''(W)\beta^2\sigma^2 + \lambda(W, \pi)(f(W - \pi\psi) - f(W))$, with boundary condition $f(0) = l$ and $f'(W^p) = -\eta$, if and only if f admits the following representation

$$f(W_0) = \sup_{\{\pi_t\}} E\left[\int_0^{\tau} e^{-rt}(g(W_t, \pi_t) - \lambda(W_t, \pi_t)L)dt - \eta dI_t) + e^{-r\tau} l\right] \tag{A.0.2}$$

Proof of Lemma A.1.

First, for necessity. Suppose f solves (**??**) and take any $\{\pi_t\}$. Apply Ito's Lemma to $e^{-rs}f(W_s)$ from 0 to $T = \min(t, \tau)$,

$$\begin{aligned}
e^{-rT}f(W_T) = f(W_0) + \int_0^T e^{-rs}\left\{(\mathcal{L}^{\pi_s}f(W_s) - rf(W_s))ds - dI_s f'(W_s)\right\} \\
+ \int_0^T e^{-rs}\left\{f'(W_s)\beta\sigma dZ_s\right\} \\
+ \int_0^T e^{-rs}\{(f(W_s - \pi_s\psi) - f(W_s))(dN_s - \lambda(W_s, \pi_s)ds)\}
\end{aligned}$$

Since the expectations of the last two terms are zero, and by (**??**) and also the boundary condition $f'(W^p) = -\eta$ when $dI_t > 0$ and $dI_t = 0$ when $f'(W^p) \neq 0$, we then get

$$\begin{aligned}
E\left[e^{-rT}f(W_T)\right] =& f(W_0) + E\left[\int_0^T e^{-rs}(\mathcal{L}^{\pi_s}f(W_s) - rf(W_s))ds + \eta dI_s - dI_s(\eta - f'(W_s))\right] \\
\leq& f(W_0) - E\left[\int_0^T e^{-rs}(g(W_s, \pi_s) - \lambda(W_s, \pi_s)L)ds + \eta dI_s\right]
\end{aligned}$$

Since τ is a bounded stopping time, sending t to ∞, and using $f(W_\tau) = l$, we have for all $\{\pi_t\}$,

$$f(W_0) \geq E\left[\int_0^\tau e^{-rt}(g(W_t, \pi_t) - \lambda(W_t, \pi_t)L)dt - \eta dI_t) + e^{-r\tau}l\right]$$

Hence (**??**) holds with $\geq$. Now suppose $\{\pi_t^*\}$ is the optimal control, and by replicating the above argument, we have

$$f(W_0) = E\left[\int_0^\tau e^{-rt}(g(W_t, \pi_t^*) - \lambda(W_t, \pi_t^*)L)dt - \eta dI_t) + e^{-r\tau}l\right]$$

Therefore, f admits representation (**??**)

Conversely, suppose f is represented by (**??**). In particular, for all t

$$f(W_t) = \sup_{\{\pi_s\}} E_t\left[\int_t^\tau e^{-r(s-t)}((g(W_s, \pi_s) - \lambda(W_s, \pi_s)L)ds - \eta dI_s) + e^{-r(\tau-t)}l\right]$$

Fix $h > 0$ such that $t + h < \tau$. By Law of Iterated Expectation,

$$f(W_t) = \sup_{\{\pi_s\}} E_t \left[\int_t^{t+h} e^{-r(s-t)}((g(W_s, \pi_s) - \lambda(W_s, \pi_s))ds - \eta dI_s) + e^{-rh} f(W_{t+h}) \right] \tag{A.0.3}$$

Fix a control $\pi_s = \pi$ for $s \in [t, t+h]$, then (**??**) holds with $\geq$. Applying Ito's formula to $e^{-rs} f(W_s)$ from t to $t + h$,

$$\begin{aligned} e^{-rh} f(W_{t+h}) &= f(W_t) \\ &+ \int_t^{t+h} e^{-r(s-t)} \left\{ (\mathcal{L}^\pi f(W_s) - \lambda(W_s, \pi_s)L + LdN_s - rf(W_s))ds - dI_s f'(W_s) \right\} \\ &+ \int_t^{t+h} e^{-r(s-t)} \left\{ f'(W_s)\beta\sigma dZ_s \right\} \\ &+ \int_t^{t+h} e^{-r(s-t)} \left\{ (f(W_s - \pi_s\psi) - f(W_s) - L)(dN_s - \lambda(W_s, \pi_s)ds) \right\} \end{aligned}$$

where the time t expectation of the third and forth term are zero. Substituting out $e^{-rh} f(W_{t+h})$ in (**??**) with π replacing π_s, and noting that $f'(W_s) = -\eta$ if $dI_s > 0$ and $f'(W_s) \neq -\eta$ if $dI_s = 0$, then we get

$$0 \geq E_t \left[\int_t^{t+h} e^{-r(s-t)} (g(W_s, \pi) - \lambda(W_s, \pi)L + \mathcal{L}^\pi f(W_s) - rf(W_s))ds \right]$$

Dividing by h and sending h to 0, by mean value theorem, $rf(W_t) \geq g(W_t, \pi) - \lambda(W_t, \pi)L + \mathcal{L}^\pi f(W_t)$. Since this is true for all π, we have

$$rf(W_t) \geq \sup_\pi \left\{ g(W_t, \pi) - \lambda(W_t, \pi)L + \mathcal{L}^\pi f(W_t) \right\}$$

By the same arugment as above, if π_t^* is the optimal control, then $rf(W_t) = g(W_t, \pi_t^*) - \lambda(W_t, \pi_t^*)L + \mathcal{L}^{\pi_t^*} f(W_t)$. Combining with the last inequality, we then get

$$rf(W_t) = \sup_\pi \left\{ g(W_t, \pi) - \lambda(W_t, \pi)L + \mathcal{L}^\pi f(W_t) \right\}$$

Finally note that by the probabilistic representation, $f(W_\tau) = f(0) = l$, therefore, f solves (**??**) with the stated boundary conditions. ■

Lemma A.2. *Suppose F is twice differentiable, strictly concave on $[0, W^p)$, and with strictly convex first derivative. Then $\mathcal{A}(W) < 0$ and is strictly increasing in W. Moreover, there exists a unique cutoff $W^* \in [\psi, W^p]$ such that for $W \in [\psi, W^*]$, $a_t^s = a_R$ and for $W \in (W, W^p]$, $a_t^s = a_N$, subject to limited liability.*

Proof of Lemma A.2.

From the HJB-equation, $a_N = a^s(W_t)$ if $(\mu - L\lambda) + \mathcal{L}_N F(W) \geq (\mu + \alpha - L(\lambda + \gamma)) + \mathcal{L}_R F(W)$, which is equivalent to $F'(W)\psi\lambda + \lambda(F(W-\psi) - F(W)) \geq \alpha - L\gamma$. Let $\mathcal{A}(W) = F'(W)\psi\lambda + \lambda(F(W - \psi) - F(W))$. Rewrite $\mathcal{A}(W)$ as $-\lambda(F'(W)(-\psi) + F(W) - F(W - \psi))$, by strict concavity of F on $[0, W^p)$, $\mathcal{A}(W) < 0$. Observe that $\mathcal{A}'(W) = -\lambda(F\text{"}(W)(-\psi) + F'(W) - F'(W - \psi))$, by strict convexity of F', $\mathcal{A}'(W) > 0$. Now consider three cases: (i) $\mathcal{A}(W^p) > \alpha - L\gamma > \mathcal{A}(\psi)$, since $\mathcal{A}$ is continuous, by intermediate value theorem, there exists a unique W^* such that the claim is true. (ii) $\alpha - L\gamma \geq \mathcal{A}(W^p)$, take $W^* = W^p + \epsilon$ for some $\epsilon > 0$, then $a_t^s = a_R$ for all $W_t \in [0, W^p]$. (iii) $\mathcal{A}(\psi) \geq \alpha - L\gamma$, by limited liability, we can only take $W^* = \psi$.■

Lemma A.3. *There exists a solution $F \in C^2$ to the HJB-equation (**??**) with the boundary conditions $F(0) = l$, $F'(W^p) = -1$, and $F''(W^p) = 0$. The solution F is strictly concave on $[0, W^p)$ and extends linearly so that $F(W) = F(W^p) - (W - W^p)$ for $W \geq W^p$. Moreover, if the solution F is such that $\alpha - L\gamma \geq \mathcal{A}(W^p)$, then F is a solution to the risky action ODE (**??**) with the same boundary conditions. And if the solution F is such that $\mathcal{A}(\psi) \geq \alpha - L\gamma$, then F is defined by (**??**) with $W^* = \psi$.*

Proof of Lemma A.3.

Existence: Let $g(W_t, \pi_t) = \mu + \pi_t\alpha$, $\lambda(W_t, \pi_t) = \lambda + \pi_t\gamma$, and $\eta = 1$. Define $F : [0, W^p] \to \mathbb{R}$ by

$$F(W_t) = \sup_{\{\pi_s\}} E_t \left[\int_t^\tau e^{-r(s-t)}((\mu + \pi_s\alpha) - (\lambda + \pi_s\gamma))ds - dI_s) + e^{-r(\tau-t)} l \right]$$

Then $F(W_0)$ is the probabilistic representation (**??**) and hence F solves (**??**) with boundary conditions $F(0) = l$ and $F'(W^p) = -1$.

Strict concavity of F on $[0, W^p)$: By way of contradiction, suppose F is not strictly concave on $[0, W^p)$. First notice that at $W = 0$, we have $rF(0) = \mu + \alpha - L(\lambda + \gamma) + \frac{1}{2}F''(0)\beta^2\sigma^2$. With the boundary condition $F(0) = l$, thus $\frac{1}{2}F''(0)\frac{\beta^2\sigma^2}{r} = l - \frac{\mu+\alpha-L(\lambda+\gamma)}{r} < 0$ by assumption 1. Hence there exists $\hat{W} = \inf\{W \in (0, W^p) : F''(W) \geq 0\}$. By continuity of F, $F''(\hat{W}) = 0$. If $\hat{W} \leq W^*$, then differentiating (**??**), $\frac{1}{2}F'''(\hat{W}) = -(\rho - r)F'(\hat{W})$. There are two cases: (i) If $F(\hat{W}) \geq \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$, then F must be increasing on $[0, W^*]$, so $F'''(\hat{W}) \leq 0$. For $\epsilon > 0$ small, continuity implies $F''(\hat{W} - \epsilon) \geq F''(\hat{W}) = 0$, which contradicts the definition $\hat{W}$. (ii) $F(\hat{W}) < \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$. There is also a contradiction if $F'(\hat{W}) \geq 0$. So assume $F'(\hat{W}) < 0$, we have $F'''(\hat{W}) > 0$. Take $\epsilon > 0$, we have $F''(\hat{W} + \epsilon) > F''(\hat{W}) = 0$ and $F'(\hat{W} + \epsilon) \geq F'(\hat{W})$ (because $F''(\hat{W}) \geq 0$). Using (**??**), $F(\hat{W} + \epsilon) > F(\hat{W})$, contradicting $F'(\hat{W}) < 0$. Now suppose $\hat{W} > W^*$. Differentiating (**??**) at $\hat{W}$, we have $\frac{1}{2}F'''(\hat{W})\beta^2\sigma^2 = -(\rho - r)F'(\hat{W}) - \lambda(F'(\hat{W} - \psi) - F'(\hat{W}))$. By definition of $\hat{W}$, $F'' < 0$ on $[0, \hat{W}]$, thus $F'(\hat{W} - \psi) - F'(\hat{W}) > 0$. If $F'(\hat{W}) \geq 0$, then $F'''(\hat{W}) < 0$. Again for $\epsilon > 0$ small, continuity implies $F''(\hat{W} - \epsilon) > F''(\hat{W}) = 0$, which contradicts the definition $\hat{W}$. Assume now $F'(\hat{W}) < 0$ and $F'''(\hat{W}) \geq 0$. Choose a small $\epsilon > 0$, as $F''(\hat{W}) \geq 0$, we have $F'(\hat{W} + \epsilon) \geq F'(\hat{W})$, $F''(\hat{W} + \epsilon) \geq 0$, and because of $F'''(\hat{W}) \geq 0$ and continuity, together with strict concavity of F on the left of $\hat{W}$, $\mathcal{A}(\hat{W} + \epsilon) \geq \mathcal{A}(\hat{W})$. This implies $F(\hat{W} + \epsilon) \geq F(\hat{W})$, which contradicts $F'(\hat{W}) < 0$.■

Proof of Proposition 1.2.

The proof verifies that the contract described in proposition 1.2 in optimal. Take any incentive compatible contract $\Gamma = (I, \tau, a)$ and define a gain process $\{G_t\}$ by $G_t \equiv \int_0^t e^{-rs}(dY_s - dI_s) + e^{-rt}F(W_t)$. Let $r_t = 1_{\{a_t^s = a_R\}}$. Under this contract, the agent's continuation value satisfies $dW_t = \rho W_t - dI_t + \beta_t(dX_t - (\mu + \alpha r_t))dt - \psi_t(dN_t - (\lambda + \gamma r_t)dt)$ for $t < \tau$. Differentiating G_t and applying Ito's lemma,

$$\begin{aligned} e^{rt}dG_t =& (\mu + \alpha r_t - L(\lambda + \gamma r_t) + F'(W_t)(\rho W_t + \psi_t(\lambda + \gamma r_t)) + \frac{1}{2}F''(W_t)\sigma^2\beta_t^2 \\ &+ (\lambda + \gamma r_t)(F(W_t - \psi_t) - F(W_t)) - rF(W_t))dt - dI_t(1 + F'(W_t)) \\ &+ (1 + \beta_t F'(W_t))\sigma(dX_t - (\mu + \alpha r_t)dt) \\ &+ (F(W_t - \psi_t) - F(W_t) - L)(dN_t - (\lambda + \gamma r_t)dt) \end{aligned}$$

By the HJB equation (**??**), the first dt term is nonpositive. Also, $dI_t \geq 0$ and $F'(W_t) \geq -1$. So the drift term is nonpositive and $\{G_t\}$ is a supermartingale under alternative policy. Under the optimal contract $dI_t(1 + F'(W_t)) = 0$. Moreover, $r_t = 1$ implies $(\beta_t, \psi_t) = (\frac{B}{\mu}, 0)$ and $r_t = 0$ implies $(\beta_t, \psi_t) = (\frac{B}{\mu}, \frac{B}{\mu}\frac{\alpha}{\gamma})$, therefore, for $W_t \in [0, W^*)$,

$$\begin{aligned} e^{rt}dG_t =& (\mu + \alpha - L(\lambda + \gamma) + F'(W_t)\rho W_t + \frac{1}{2}F''(W_t)(\frac{B}{\mu})^2\sigma^2 - rF(W_t))dt \\ &- (1 + F(W_t))dI_t + (1 + \frac{B}{\mu}F'(W_t))\sigma dZ_t - L(dN_t - (\lambda + \gamma)dt) \end{aligned}$$

and for $W_t \in [W^*, W^p]$,

$$\begin{aligned} e^{rt}dG_t =& (\mu - L\lambda + F'(W_t)\rho W_t + \frac{1}{2}F''(W_t)(\frac{B}{\mu})^2\sigma^2 + \lambda(F(W_t - \frac{B}{\mu}\frac{\alpha}{\gamma}) - F(W_t)) - rF(W_t))dt \\ &- (1 + F(W_t))dI_t + (1 + \frac{B}{\mu}F'(W_t))\sigma dZ_t + (F(W_t - \frac{B}{\mu}\frac{\alpha}{\gamma}) - F(W_t) - L)(dN_t - \lambda dt) \end{aligned}$$

In either case, the drift is zero, together with the boundedness of F' and F, $\{G_t\}$ is a martingale. Now evaluate the principal's profit for any incentive compatible contract. For all $t < \infty$, since $F(W_\tau) = l$,

$$
\begin{aligned}
&E^a[\int_0^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}l] \\
&\quad = E^a\left[G_{\min(t,\tau)} + 1_{t\leq\tau}\left(\int_t^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}l - e^{-rt}F(W_t)\right)\right] \\
&\quad = E^a\left[G_{\min(t,\tau)}\right] + E^a\left[1_{t\leq\tau}\left(E_t^a(\int_t^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}l) - F(W_t)\right)\right]
\end{aligned}
$$

For any t, the first term is bounded by $G_0 = F(W_0)$ because $\{G_t\}$ is a supermartingale, the stopped process $\{G_{\min(t,\tau)}\}$ is also a supermartingale. As for the second term, we have $E_t^a(\int_t^\tau e^{-r(s-t)}(dY_s - dI_s) + e^{-r(\tau-t)}l) \leq \frac{\mu-L\lambda}{r} - W_t = F^{FB}(W_t)$. At any moment of time, the principal can always pay the agent W_t and terminate the contract to get l, this implies $l - W_t \leq F(W_t)$. It follows that $\frac{\mu-L\lambda}{r} - W_t - F(W_t) \leq \frac{\mu-L\lambda}{r} - l$, so the intergrand of the second term is bounded. Sending t to ∞, the second term vanishes and since $\min(t,\tau)$ is a bounded stopping time, $G_{\min(t,\tau)} \to G_t$ a.s. as $t \to \infty$, $E^a\left[\int_0^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}l\right] = E^a(G_t) \leq F(W_0)$. Under the optimal contract, $\{G_t\}$ is a martingale and profit $F(W_0)$ is achieved with equality.■

Proof of Lemma 1.3.

Suppose $F_{N,m}(W^p) + W^p > \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$. Since $F(W) \geq F_{N,m}(W)$ for all W, $-(\alpha - L\gamma) > \mu - L\lambda - r(F_{N,m}(W^p) + W^p) > \mu - L\lambda - r(F(W^p) + W^p)$. Substituting $F(W^p)$ using (**??**) with $F'(W^p) = -1$ and $F''(W^p) = 0$, $-(\alpha - L\gamma) > W^p(\rho - r) - \mathcal{A}(W^p) > -\mathcal{A}(W^p)$. By lemma A.2, $W^* < W^p$. For the second part, by contrapositive, suppose $F_{N,m}(\psi) + \psi > \frac{\mu+\alpha-L(\lambda+\gamma)}{r}$, repeating the above argument, $\mathcal{A}(\psi) > \alpha - L\gamma$. So by lemma A.2, $W^* = \psi$.■

Proof of Proposition 1.4.

Since $W_t = \beta M_t$, $dW_t = \beta dM_t$. Using equation (**??**),

$$\begin{aligned} dW_t =& r\beta M_t dt + \beta(dX_t - LdN_t) - \beta dD_t - \beta c dt - \beta p dt - \beta dR_t \\ =& \rho\beta M_t dt + \beta(dX_t - \mu dt) - \beta dD_t - 1_{\{M_t \geq M^*\}}(\psi(dN_t - \lambda dt) + \beta\alpha dt) \\ =& \rho W_t dt - dI_t + \beta(dX_t - \mu dt) - 1_{\{W_t \geq W^*\}}(\psi(dN_t - \lambda dt) + \beta\alpha dt) \end{aligned}$$

where the second line uses the definition of c_t and dR_t stated in the proposition. The third line is obtained by defining $dI_t = \beta dD_t$ and $W_t = \beta M_t$ with $M^* = \frac{1}{\beta}W^*$. This calculation shows that for $W_t \geq W^*$, $dW_t = (\rho W_t + \psi\lambda)dt - dI_t + \beta(dX_t - \mu dt) - \psi dN_t$, and for $W_t < W^*$, $dW_t = \rho W_t dt - dI_t + \beta(dX_t - (\mu + \alpha)dt)$. Therefore, incentive compatibility follows from proposition 1.1. ■

The proof of proposition 1.5 is equivalent to solving a Skorohod problem.

Definition. Let $K = (-\infty, k)$ for some $k > 0$. The *Skorohod Problem* for reflected jump diffusions into $\bar{K}$ (with respect to direction -1) is to find a pair (M_t, D_t) of RCLL and $\mathcal{F}_t$-adapted process such that the following conditions are satisfied:

$$\begin{aligned} dM_t =& \left(rM_t - \Psi_t + \mu + 1_{\{M_t < M^*\}}\alpha\right) dt + \sigma dZ_t - LdN_t - dD_t \\ M_{0^-} =& M_0 \in \mathbb{R} \\ M_t \in& \bar{K} \text{ for all } t \geq 0 \\ D_t \in& \mathbb{R} \text{ has finite variation and } dD_t = 0 \text{ if } M_t \in \bar{K} \end{aligned}$$

Proof of Proposition 1.5.

Let $\mathcal{C} = \{f | f : \mathbb{R}_+ \to \mathbb{R},\ \text{RCLL}\}$, i.e., $\mathcal{C}$ is a set of RCLL functions. Define $\mathcal{H} : \mathcal{C} \to \mathcal{C}$ by $\mathcal{H}(f)(t) = f(t) - \sup_{0 \leq s \leq t} f^k(s)$, where $f^k(s) = \max(f(s) - k, 0)$. Also define $\mathcal{Q} : \mathcal{C} \to \mathcal{C}$ by $\mathcal{Q}(f)(t) = \mathcal{H}(f)(t) - f(t) = \sup_{0 \leq s \leq t} f^k(s)$. First, I show that there exists a solution (L_t) to the SDE

$$dM_t = \left(rM_t - \Psi_t + \mu + 1_{\{M_t < M^*\}}\alpha\right) dt + \sigma dZ_t - LdN_t$$

with $L_0 = M_0$. Now define $M_t = \mathcal{H}(L)(t)$ and $D_t = \mathcal{Q}(L)(t)$. Then the pair (M_t, D_t) satisfies the jump diffusion in Skorohod problem, because

$$\begin{aligned} dM_t + dD_t =& d\left(\mathcal{H}(L)(t) - \mathcal{Q}(L)(t)\right) \\ =& d\left(\mathcal{H}(L)(t) - \mathcal{H}(L)(t) + L(t)\right) \\ =& dL(t) \\ =& \left(rM_t - \Psi_t + \mu + 1_{\{M_t < M^*\}}\alpha\right) dt + \sigma dZ_t - LdN_t \end{aligned}$$

By definition of $\mathcal{H}$, $M_t = \mathcal{H}(L)(t) = L(t) - \sup_{0 \le s \le t} L^k(s) \in \bar{K}$. And observe $D_t = \mathcal{Q}(L)(t) = \sup_{0 \le s \le t} L^k(s)$ is nondecreasing and hence has finite variation. Lastly, I need to show that if $M_t \in K$, then $dD_t = 0$. Suppose $M_t \in K$, $L(t) - \sup_{0 \le s \le t} L^k(s) < k$, and thus $L(t) < k + \sup_{0 \le s \le t} \{\max(L(s) - k, 0)\} = \max\left\{\sup_{0 \le s \le t} L(s), k\right\}$. This strict inequality implies both $L(t) < k$, in which case $dD_t = 0$, and $L(t) < \sup_{0 \le s \le t} L(s)$, which also implies $dD_t = 0$. Therefore, the pair (M_t, D_t) defined above solves the Skorohod problem. The proof is completed by setting $k = M^p$. ■

Proof of Proposition 1.6.

By equation (**??**), project i is chosen when $F'(W)\psi_i\lambda + \lambda(F(W-\psi) - F(W)) \ge \alpha_i - L_i\gamma_i$. Suppose $\alpha_i - L_i\gamma_i = K$ and $\lambda_i = \lambda$ for all $i = 1, ..., n$. Then W_i^* is defined by

$$\mathcal{A}(W_i^*; \psi_i) \equiv \lambda(F'(W_i^*)\psi_i + F(W_i^* - \psi_i) - F(W_i^*)) = K$$

Suppose $\frac{\alpha_i}{\gamma_i} \ge \frac{\alpha_j}{\gamma_j}$. By concavity of F, incentive constraints are binding and hency $\psi_i \ge \psi_j$. Again by strict concavity of F on $[0, W^p]$, $\frac{\partial \mathcal{A}(W;\psi)}{\partial \psi} = \lambda(F'(W) - F'(W-\psi)) < 0$. It follows that $\mathcal{A}(W;\psi_i) \le \mathcal{A}(W;\psi_j)$ for all W. By lemma A.2, $\mathcal{A}(W;\psi)$ is strictly increasing in W. As a result $W_i^* \geq W_j^*$ by the equation that defines the optimal switching point. Conversely, suppose not, $\frac{\alpha_j}{\gamma_j} > \frac{\alpha_i}{\gamma_i}$, then by the same argument, $W_j^* > W_i^*$. Thus the claim follows. ■

Proof of Proposition 1.7.

The proof goes as follows: write the contract on the compensated compound Poisson process, this generates a value function $\tilde{F}$ that satisfies an intergo-differential equation. Then I show that this value function is dominated by the solution to equation (**??**). Now apply martingale representation theorem to $\{\sum_{k=1}^{N_t} J_k - \int_0^t \int_{-\infty}^0 \lambda(a_u^s) J dH(J) du\}_{t\geq 0}$, the continuation value evolves according to

$$dW_t = \rho W_t - dI_t + \beta_t \sigma dZ_t + \tilde{\psi}_t (J dN_t - (\lambda + a_t^s \gamma) E(J) dt)$$

To implement $a_t^s = 0$, the required incentive constraint is $\tilde{\psi}_t \geq \beta_t \frac{\alpha}{\gamma} \frac{1}{E(-J)}$. Then resulting HJB-equation is

$$r\tilde{F}(W) = \mu - L\lambda + a^s(\alpha - L\gamma) + \tilde{F}'(W)(\rho W - (1-a^s)\tilde{\psi}\lambda E(-J)) + \frac{1}{2}\tilde{F}''(W)\beta^2\sigma^2 \quad \text{(A.0.4)}$$

$$+\lambda \int_0^\infty (\tilde{F}(W - (1-a^s)\tilde{\psi}(-J)) - \tilde{F}(W)) dH(J)$$

where $\tilde{F}$ is a solution with boundary condition $\tilde{F}(0) = l$, $\tilde{F}'(\tilde{W}^p) = -1$, and $\tilde{F}''(\tilde{W}^p) = 0$. Denote $\tilde{W}^p$ and $\tilde{W}^*$ as the optimal payment boundary and switching point with jump size incorporated respectively. Note that $\tilde{F}$ is concave and therefore the incentive constraints are binding: $\beta = \frac{B}{\mu}$ and $\tilde{\psi} = \frac{B}{\mu}\frac{\alpha}{\gamma}\frac{1}{E(-J)}$. Because $\tilde{F}$ strictly concave and $a^s = 0$ on $[\tilde{W}^*, \tilde{W}^p]$, by Jensen's inequality, $\int_0^\infty \tilde{F}(W - \tilde{\psi}(-J)) dH(J) < \tilde{F}(W - \tilde{\psi}E(-J))$. Therefore, with $\tilde{W}^p$, $\tilde{W}^*$ fixed,

$$r\tilde{F}(W) < \mu - L\lambda + 1_{\{W<\tilde{W}^*\}}(\alpha - L\gamma) + \tilde{F}'(W)(\rho W - 1_{\{W\geq\tilde{W}^*\}}\tilde{\psi}\lambda E(-J)) + \frac{1}{2}\tilde{F}''(W)\beta^2\sigma^2 \quad \text{(A.0.5)}$$

$$+\lambda(\tilde{F}(W - 1_{\{W \geq \tilde{W}^*\}}\tilde{\psi}E(-J)) - \tilde{F}(W))$$

Observe that $\tilde{\psi} = \psi\frac{1}{E(-J)}$, and so the right-hand side of (**??**) can be satisfied by a differential equation with solution $F_{\tilde{W}^*,\tilde{W}^p}$ that passes through $(0, l)$ and $(\tilde{W}^p, \tilde{F}(\tilde{W}^p))$. The solution also satisfies equation (**??**), in fact, as $\tilde{W}^p$ and $\tilde{W}^*$ are fixed, $F_{\tilde{W}^*,\tilde{W}^p}(W) \leq F(W)$ for all W because W^* and W^p are chosen optimally. ■

Appendix to Chapter 2

Proof of Lemma 2.1.

See lemma 2 in He (2011). Note that his lemma is a form of single-agent revelation principle. His proof directly extends to multi-agent environment. ■

Proof of Lemma 2.2.

Fix an agent i and an arbitrary contract Γ. Suppose all the agents follow a. Then $B_t^a = (B_{1t}^a, ..., B_{nt}^a)$ with $B_{kt}^a = A_{kt} - \int_0^t \sum_{k=1}^n \mu_k(a_s)ds$ for k is a n-dimensional Brownian motion under measure P^a. Define the value process for agent i, $(V_{it}(\Gamma, a_{-i}))_{t\geq 0}$ as follows

$$V_{it}(\Gamma, a_{-i}) \equiv E_t^a \left[\int_0^t e^{-rs}u_i(c_{is}, a_{is})ds\right]$$

then the value process is a square-integrable $\mathcal{F}_t$-martingale. By the martingale representation theorem, there exists for each i, a $\mathcal{F}_t$-progressive measurable process $\beta_t^i = (\beta_t^{i1}, ..., \beta_t^{in})$ in $\mathcal{L}^2$ such that for all t,

$$V_{it}(\Gamma, a_{-i}) = V_{i0}(\Gamma, a_{-i}) + \int_0^t e^{-rs}(-\gamma_i r W_{is})\sum_{k=1}^n \beta_s^{ik} dB_{ks}^a$$

Differentiating the above two representation of V_{it} with respect to t and using the relation $B_{kt} = \sum_{l=1}^{d} \sigma_{kl} dZ_{lt}$, we have

$$dW_{it}(\Gamma, a_{-i}) = (rW_{it}(\Gamma, a_{-i}) - u_i(c_{it}, a_{it}))dt + (-\gamma_i r W_{it}) \sum_{k=1}^{n} \beta_t^{ik} (\sum_{l=1}^{d} \sigma_{kl} dZ_{lt})$$

Therefore the claim follows. ■

Proof of Lemma 2.3.

See lemma 3 in He (2011). His proof follows by noting that the agent in my model takes the effort of all other agents as given when she makes a consumption-saving decision.■

Appendix to Chapter 3

Proof of Proposition 3.1.

We verify T is a self-map. Since R_i, v are bounded and $p_i \in [0,1]$, so as Tv. Moreover, $p_i \mapsto \int v dp_i$ is continuous under the topology of weak convergence. Thus, by the Maximum Theorem, $\min_{p_i \in \mathcal{P}_i(x_i)} \int v(\tilde{x}_i) dp_i(\tilde{x}_i, x_i)$ is continuous in x_i. Thus, Tv is continuous in x_i. Hence T is a self-map. Verification of monotonicity and discounting are routine. The claim follows from the contraction mapping theorem.■

Proof of Proposition 3.2.

Part (i): Define $g_i(x_i) = G_i(x_i)(1 - \beta)$. We argue that $g_i(x_i) = v_i(x_i)$.

Step 1: show that $v_i(x_i) \geq g_i(x_i)$.

Let $y < g_i(x_i)$, this implies $m = \frac{y}{1-\beta} < G_i(x_i)$. Given the terminal reward is m, the optimal action is to continue to play arm i. By definition of $G_i(x_i)$, $V_i(x_i, m) > m$. Let $\tau_i(x_i, m)$ be the optimal stopping time with arm i when terminal reward is m, then

$$\begin{aligned} V_i(x_i, m) &= \min_{p_i \in \mathcal{P}_i(x_i)} E_{p_i}(\sum_{t=0}^{\tau_i(x_i,m)-1} \beta^t R_i(x_{it}) + \sum_{t=\tau_i(x_i,m)}^{\infty} y) \\ &> m = \tfrac{y}{1-\beta} \end{aligned}$$

This implies for any $p_i \in \mathcal{P}_i(x_i)$

$$E_{p_i}(\sum_{t=0}^{\tau_i(x_i,m)-1} \beta^t R_i(x_{it}) + \sum_{t=\tau_i(x_i,m)}^{\infty} y) > \frac{y}{1-\beta}$$

and

$$E_{p_i}(\sum_{t=0}^{\tau_i(x_i,m)-1} \beta^t R_i(x_{it})) > yE_{p_i}(\sum_{t=0}^{\tau_i(x_i m)-1} \beta^t)$$

It follows that

$$v_i(x_i, \tau_i(x_i, m)) = \min_{p_i \in \mathcal{P}_i(x_i)} \frac{E_{p_i}(\sum_{t=0}^{\tau_i(x_i,m)-1} \beta^t R_i(x_{it}))}{E_{p_i}(\sum_{t=0}^{\tau_i(x_i,m)-1} \beta^t)} > y$$

Hence, $v_i(x_i) > y$ and $v_i(x_i) \geq g_i(x_i)$.

Step 2: show that $g_i(x_i) \geq v_i(x_i)$.

Let $y < v_i(x_i)$. By definition of v_i, there exists a stopping time σ such that $y < v_i(x_i, \sigma)$. Hence for any $p_i \in \mathcal{P}_i(x_i)$,

$$\frac{E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t R_i(x_{it}))}{E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t)} > y$$

and

$$E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t R_i(x_{it})) > E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t y)$$

Given a terminal reward $m = \frac{y}{1-\beta}$, the worth of σ is

$$\begin{aligned} w_\sigma(x_i, m) &= \min_{p_i \in \mathcal{P}_i(x_i)} E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t R_i(x_{it}) + \beta^\sigma m) \\ &> \min_{p_i \in \mathcal{P}_i(x_i)} E_{p_i}(\sum_{t=0}^{\sigma-1} \beta^t y + \beta^\sigma m) \\ &= m \end{aligned}$$

Taking the supremum over the set of stopping times, $V_i(x_i, m) > m$, thus continuation is optimal and $G_i(x_i) > m$ by definition. It follows that $g_i(x_i) > y$, so $g_i(x_i) \geq v_i(x_i)$.

The second claim follows from:

$$\begin{aligned}\tau_i(x_i) &= \inf\{t : v_i(x_{it}) < v_i(x_i)\} \\ &= \inf\{t : G_i(x_{it}) < G_i(x_i)\} \\ &= \tau_i(x_i, G_i(x_i))\end{aligned}$$

where the second equality follows from the first part of the proposition.

Part (ii): By definition of infimum, at any x_{it}, for any $\tau \in \mathcal{T}_{t+1}$ and $p \in \mathcal{P}_0$, we have $F_i^{p,\tau}(x_{it}) \geq \inf_{p\in\mathcal{P}_i(x_{it})} F_i^{p,\tau}(x_{it})$. This implies $\sup_{\tau\in\mathcal{T}_{t+1}} F_i^{p,\tau}(x_{it}) \geq \sup_{\tau\in\mathcal{T}_{t+1}} \inf_{p\in\mathcal{P}_i(x_{it})} F_i^{p,\tau}(x_{it})$ and hence $\inf_{p\in\mathcal{P}_i(x_{it})} G_i^p(x_{it}) \geq G_i(x_{it})$.

Let p^* and τ^* be the measure that achieves the infimum and the stopping time that achieves the supremum respectively. By definition of the MP-index, we have $\frac{G_i(x_{it})}{1-\beta} = E_{p^*}\left(\sum_{s=t}^{\tau^*-t}\beta^{s-t}R_i(x_{is}) + \beta^{\tau^*-t}\frac{G_i(x_{it})}{1-\beta}|x_{it}\right)$. Hence, $G_i(x_{it}) = F^{p^*,\tau^*}(x_{it})$. But for p^*, the Gittins Index $G_i^{p^*}(x_{it})$ satisfies $\frac{G_i^{p^*}(x_{it})}{1-\beta} = \sup_{\tau\geq t+1} E_{p^*}\left(\sum_{s=t}^{\tau-1}\beta^{s-t}R_i(x_{is}) + \beta^{\tau-t}\frac{G_i^{p^*}(x_{it})}{1-\beta}|x_{it}\right)$. By the minimax result in Riedel (2009), we know τ^* also achieves the supremum for the right-hand side of the last equality. This implies for any stopping time $\tau \in \mathcal{T}_{t+1}$, $F_i^{p^*,\tau^*}(x_{it}) \geq \sup_{\tau\geq t+1} F_i^{p^*,\tau}(x_{it})$. Collecting the results,

$$\begin{aligned}\sup_{\tau\geq t+1}\inf_{p\in\mathcal{P}_i(x_{it})} F_i^{p,\tau}(x_{it}) = G_i(x_{it}) &= F_i^{p^*,\tau^*}(x_{it}) \\ &\geq \sup_{\tau\geq t+1} F_i^{p^*,\tau}(x_{it}) \\ &\geq \inf_{p\in\mathcal{P}_i(x_{it})}\sup_{\tau\geq t+1} F_i^{p,\tau}(x_{it})\end{aligned}$$

This completes the proof. ∎

Proof of Proposition 3.3.

We imitate Whittle's (1980) retirement option argument. For the moment, we suppress the subscript i, lemma A.4 to lemma A.6 hold for any i.

For any τ, define $w_\tau : X \times \mathcal{M} \to \mathbb{R}$ by $w_\tau(x,m) = \min_{p\in\mathcal{P}(x)} E_p[\sum_{t=0}^{\tau-1}\beta^t R(x) + \beta^\tau m]$ as the value of stopping time τ given x and m. The objective function $f_\tau : \Delta(X)\times X \times \mathcal{M} \to \mathbb{R}$ is given by $f_\tau(p;x,m) = E_p[\sum_{t=0}^{\tau-1}\beta^t R(x) + \beta^\tau m]$. The following lemma characterizes the right-hand derivative of w_τ.

Lemma A.4. *For any τ, define a compact-valued and convex-valued correspondence $Q : X \rightsquigarrow \Delta(X)$ by*

$$Q(x) = \{p \in \mathcal{P}(x) | w_\tau(x,m) = f_\tau(p;x,m)\}$$

Then the right-hand derivative of $w_\tau(x,\cdot)$ at m in the direction h is given by

$$Dw_\tau(x,m)(h) = \min_{p \in Q(x)} E_p(\beta^\tau)$$

Proof of Lemma A.4

Note that (i) $\mathcal{P}(x)$ is compact-valued for any $x \in X$. (ii) $p \longmapsto f_\tau(p;x,m)$ is continuous everywhere. (iii) $\forall x$, $\forall p \in \mathcal{P}(x)$, $m \longmapsto f_\tau(p;x,m)$ is linear in m (hence concave), and differentiable. The lemma is true by proposition 6, p.118 in Aubin (2007).■

Lemma A.5. *For any $x \in X$, $m \longmapsto V(x,m)$ is nondecreasing and differentiable almost everywhere. Moreover, let $\tau(x,m)$ be the optimal stopping time with retirement option m, the derivative of $V(x,m)$ is*

$$\frac{\partial}{\partial m} V(x,m) = \min_{p \in Q(x)} E_p(\beta^{\tau(x,m)}) \text{ a.e.}$$

Proof of Lemma A.5.

Monotonicity is obvious. We show that $m \longmapsto V(x,m)$ is absolutely continuous. First, note that for any x and τ, the mapping $m \longmapsto w_\tau(x,m)$ is absolutely continuous. Since $f_\tau(p;x,m)$ is continuous in m, by Maximum Theorem, $w_\tau(x,m)$ is continuous in m. Moreover, $w_\tau(x,m)$ is concave in m as $w_\tau(x,m)$ is defined via a minimum over a set of measures. But any continuous and concave real function is absolutely continuous.

Second, since $w_\tau(x,m)$ is absolutely continuous, it is almost everywhere differentiable in m. When it exists, the derivative $Dw_\tau(x,m)$ is equal to the right-hand derivative of $w_\tau(x,m)$ at m. Moreover, for almost all $m \in \mathcal{M}$, $0 \leq Dw_\tau(x,m) \leq 1$ as $E_p(\beta^\tau)$ is bounded between 0 and 1. Define $b : \mathcal{M} \rightarrow \mathbb{R}$ by $b(m) = 1\ \forall m$. b is Lebesgue integrable and $|Dw_\tau(x,m)| \leq b(m)$ for almost all m and for all x and τ. By theorem 2 in Milgrom and Segal (2002), $m \longmapsto V(x,m)$ is absolutely continuous and therefore differentiable almost everywhere.

Finally, let $\tau(x,m)$ denotes the optimal stopping time with terminal reward m and initial state x. For any $m \in int(\mathcal{M})$, when V is differentiable at m, $w_{\tau(x,m)}(x,m)$ must exist as $V(x,m) = w_{\tau(x,m)}(x,m)$. Therefore, by theorem 1 in Milgrom and Segal (2002), $\frac{\partial}{\partial m}V(x,m) = Dw_{\tau(x,m)}(x,m)$. But from lemma A.4, we know that $w_{\tau(x,m)}(x,m)$ equals to the right-hand derivative of w. Therefore, $\frac{\partial}{\partial m}V(x,m) = \min_{p \in Q(x)} E_p(\beta^{\tau(x,m)})$ almost everywhere.■

Before proceeding, I need one more lemma.

Lemma A.6. *The mapping* $m \longmapsto \min_{p \in Q(x)} E_p(\beta^{\tau(x,m)})$ *is increasing for all* $x \in X$.

Proof of Lemma A.6.

By lemma 2.1 in Karoui and Karatzas (1993), $m \longmapsto \tau(x,m)$ is decreasing. Therefore, the objective function $E_p(\beta^{\tau(x,m)})$ is increasing as $\beta \in [0,1)$. The claim follows from the Monotone Maximum Theorem.■

The Bellman equation for the $(\mathcal{K}, (A_i)_{i \in \mathcal{K}}, \beta)$ bandit with retirement option $m \in \mathcal{M}$ is

$$V(x,m) = \max\left\{m, \max_{i \in \mathcal{K}}\left\{R_i(x_i) + \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int V(\tilde{x}_i, x_{-i}, m) dp_i(\tilde{x}_i, x_i)\right\}\right\} \tag{A.0.6}$$

and for the stopping problem with a single arm i

$$V_i(x_i,m) = \max\left\{m, R_i(x_i) + \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i, m) dp_i(\tilde{x}_i, x_i)\right\} \tag{A.0.7}$$

Under the index policy specified in (a) and (b), one can conjecture the form of $V(x,m)$. Let $\tau(x,m)$ denote the stopping time for the entire bandit problem. Denote by $\tau_i(x_i,m)$ the retirement time for the single bandit i with retirement option m. (a) and (b) imply

$$\tau(x,m) = \sum_{i=1}^{K} \tau_i(x_i,m) \tag{A.0.8}$$

Lemma A.7. *For almost all m,*

$$\frac{\partial}{\partial m} V(x,m) = \prod_{i=1}^{K} \frac{\partial}{\partial m} V_i(x_i,m)$$

Proof of Lemma A.7.

We have the following chain of equalities:

$$\begin{aligned}
\tfrac{\partial}{\partial m} V(x,m) &= \min_{p\in Q(x)} E_p(\beta^{\tau(x,m)}) \\
&= \min_{p\in Q(x)} E_p(\beta^{\sum_i \tau_i(x_i,m)}) \\
&= \min_{p\in Q(x)} \prod_{i=1}^{K} E_{p_i}(\beta^{\tau_i(x_i,m)}) \\
&= \prod_{i=1}^{K} \min_{p_i\in Q_i(x_i)} E_{p_i}(\beta^{\tau_i(x_i,m)}) \\
&= \prod_{i=1}^{K} \tfrac{\partial}{\partial m} V_i(x_i,m)
\end{aligned}$$

The first equality follows from lemma A.5. The second equality uses (**??**). For any given $p \in Q(x)$, the random variables $\{\tau_i(x_i,m)\}$ are independent, hence the third equality. The forth equality requires also the stochastic independence of arms. In particular, for any $p \in Q(x)$, there exists $p_i \in \mathcal{P}_i(x_i)$ such that $p = \times_{i=1}^{K} p_i$ and $p_i \in Q_i(x_i)$.The last equality uses again lemma A.5, applied to arm i's stopping problem.

Integrating $\frac{\partial}{\partial m}V(x,m)$ from m to $\frac{C}{1-\beta}$, we obtain $\int_m^{\frac{C}{1-\beta}} \frac{\partial}{\partial m}V(x,\tilde{m})d\tilde{m} = V(x,\frac{C}{1-\beta}) - V(x,m)$. Notice for $m \geq \frac{C}{1-\beta}$, immediate retirement is optimal, $\Phi(x,\frac{C}{1-\beta}) = \frac{C}{1-\beta}$. Hence

$$V(x,m) = \frac{C}{1-\beta} - \int_m^{\frac{C}{1-\beta}} \frac{\partial}{\partial m}V(x,\tilde{m})d\tilde{m} \qquad \text{(A.0.9)}$$

that is, the value function necessarily satisfies (**??**) under the index strategy. Therefore, we define another value function $\Phi : X \times \mathcal{M} \to \mathbb{R}$ by

$$\Phi(x,m) = \frac{C}{1-\beta} - \int_m^{\frac{C}{1-\beta}} \prod_{j \neq i} \frac{\partial}{\partial m}V_j(x_j,\tilde{m})d\tilde{m} \qquad \text{(A.0.10)}$$

and show that $\Phi(x,m) = V(x,m)$ for any x and m under the index strategy. For this purpose, define for each i, $H_i(x,m) = \prod_{j \neq i} \frac{\partial}{\partial m}V_j(x_j,m)$. Note that $H_i(x,m)$ is (i) nonnegative as $E_{p_j}(\beta^{\tau_j(x_j,m)}) \geq 0$ for all $p_j \in Q_j(x_j)$. (ii) nondecreasing in m by lemma A.6. (iii) Ranging from 0 to 1. Because for $m \leq -\frac{C}{1-\beta}$, 'never stop' is optimal, so for all j,$\tau_j(x_j,m) = \infty$ a.e. and $E_{p_j}(\beta^{\tau_j(x_j,m)}) = 0$. On the other hand, $m \geq \frac{C}{1-\beta}$, 'immediate retirement' is optimal, so for all j, $\tau_j(x_j,m) = 0$ a.e. and $E_{p_j}(\beta^{\tau_j(x_j,m)}) = 1$. Therefore, H_i is a distribution function on $\mathcal{M}$. Moreover, $H_i(x,m)$ does not depend on x_i.

Using the definition and integrate (**??**) by parts,

$$\Phi(x,m) = V_i(x_i,m)H_i(x,m) + \int_m^{\frac{C}{1-\beta}} V_i(x_i,\tilde{m})dH_i(x,\tilde{m}) \qquad \text{(A.0.11)}$$

We show that Φ satisfies the Bellman equation (**??**)

Step 1: Show that $\Phi(x,m) \geq m$. From lemma A.5, $V_i(x_i,m)$ is nondecreasing in m. This implies

$$\begin{aligned}
\Phi(x,m) \quad & \geq V_i(x_i,m)H_i(x_i,m) + V_i(x_i,m)\int_m^{\frac{C}{1-\beta}} dH_i(x_i,\tilde{m}) \\
& = V_i(x_i,m)H_i(x,m) + V_i(x_i,m)\left[H_i(x_i,\tfrac{C}{1-\beta}) - H_i(x_i,m)\right] \\
& = V_i(x_i,m)H_i(x_i,\tfrac{C}{1-\beta}) \\
& = V_i(x_i,m) \geq m
\end{aligned}$$

where the last inequality follows from (**??**) for all i.

Step 2: We show that $\Delta_i \geq 0$ for all i, where

$$\Delta_i = \Phi(x,m) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} \Phi(\tilde{x}_i, x_{-i}, m) dp_i(\tilde{x}_i, x_i)$$

Substituting out $\Phi(x,m)$ in Δ_i using (**??**),

$$\begin{aligned}
&\Delta_i \\
&= V_i(x_i,m)H_i(x,m) + \int_m^{\frac{C}{1-\beta}} V_i(x_i,m)dH_i(x_i,\tilde{m}) - R_i(x_i) \\
&-\beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} \left[V_i(\tilde{x}_i,m)H_i(\tilde{x}_i,x_{-i},m) + \int_m^{\frac{C}{1-\beta}} V_i(\tilde{x}_i,\tilde{m})dH_i(\tilde{x}_i,x_{-i},m) \right] dp_i(\tilde{x}_i,x_i) \\
&= H_i(x,m) \left[V_i(x_i,m) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i,m)dp_i(\tilde{x}_i,x_i) \right] \\
&+ \int_m^{\frac{C}{1-\beta}} \left[V_i(x_i,\tilde{m}) - R_i(x_i) \right] dF_i(x,\tilde{m}) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} \int_m^{\frac{C}{1-\beta}} V_i(\tilde{x}_i,\tilde{m})dH_i(x,\tilde{m})dp_i(\tilde{x}_i,x_i) \\
&= H_i(x,m) \left[V_i(x_i,m) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i,m)dp_i(\tilde{x}_i,x_i) \right] \\
&+ \int_m^{\frac{C}{1-\beta}} \left[V_i(x_i,\tilde{m}) - R_i(x_i) \right] dH_i(x,\tilde{m}) \\
&-\beta \int_m^{\frac{C}{1-\beta}} \left[\min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i,\tilde{m})dp_i(\tilde{x}_i,x_i) \right] dH_i(x,\tilde{m}) \\
&= H_i(x,m) \left[V_i(x_i,m) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i,m)dp_i(\tilde{x}_i,x_i) \right] \\
&+ \int_m^{\frac{C}{1-\beta}} \left[V_i(x_i,\tilde{m}) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i,\tilde{m})dp_i(\tilde{x}_i,x_i) \right] dH_i(x,\tilde{m})
\end{aligned}$$

Therefore, by the Bellman equation (**??**), $\Delta_i \geq 0$ for all i.

Next, we show that $\Phi(x,m) = m$ and $\Delta_i = 0$ for all i under the index strategy. In other words, the index strategy is uniquely optimal for the multi-armed bandit.

Step 3: Show that $\Phi(x,m) = m$ on the event $\left\{ m \geq \max_{i \in \mathcal{K}} \{G_i(x_i)\} \right\}$ for all $x \in X$.

For any $\tilde{m} \geq m \geq \max_{i \in \mathcal{K}} \{G_i(x_i)\}$, the index strategy (a) requires the retirement for all bandit processes. This implies that for any i, $\tau_i(x_i,m) = 0$ for all $x \in X$. Hence

$\frac{\partial}{\partial m} V_i(x_i, m) = 1$ and $H_i(x, m) = 1$ and $dH_i(x, m) = 0$. By (**??**) and Proposition 3.2, $\Phi(x, m) = V_i(x_i, m) = m$ for all i.

Step 4: Show that $\Phi(x, m) = R_k(x_k) + \beta \min_{p_k \in \mathcal{P}_k^{+1}(x_k)} \int \Phi_k(\tilde{x}_k, m) dp_k(\tilde{x}_k, x_k)$ on the event $\left\{ m < G_k(x_k) = \max_{i \in \mathcal{K}} \{G_i(x_i)\} \right\}$ for all $x \in X$.

If i is such that $m \leq \tilde{m} < G_i(x_i)$, the index strategy (b) requires the continuation with arm i. Hence, by proposition 3.2,

$$V_i(x_i, \tilde{m}) = R_i(x_i) + \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i, \tilde{m}) dp_i(\tilde{x}_i, x_i)$$

Then Δ_i reduces to

$$\Delta_i = \int_{G_i(x_i)}^{\frac{C}{1-\beta}} \left[V_i(x_i, \tilde{m}) - R_i(x_i) - \beta \min_{p_i \in \mathcal{P}_i^{+1}(x_i)} \int_{X_i} V_i(\tilde{x}_i, \tilde{m}) dp_i(\tilde{x}_i, x_i) \right] dF_i(x, \tilde{m})$$

If k is such that $\tilde{m} \geq G_k(x_k) = \max_{i \in \mathcal{K}} \{G_i(x_i)\}$, then $dH_i(x, \tilde{m}) = 0$ and $\Delta_k = 0$. Then

$$\Phi(x, m) = R_k(x_k) + \beta \min_{p_k \in \mathcal{P}_k^{+1}(x_k)} \int_{X_k} V_k(\tilde{x}_k, \tilde{m}) dp_k(\tilde{x}_k, x_k)$$

Therefore, $\Phi(x, m)$ satisfies the Bellman equation (**??**). In addition, for $m \geq G_k(x_k)$, $V(x, m) = m$ so retirement is optimal. For $m < G_k(x_k)$, we have

$$V(x, m) = R_k(x_k) + \beta \min_{p_k \in \mathcal{P}_k^{+1}(x_k)} \int_{X_k} V_k(\tilde{x}_k, \tilde{m}) dp_k(\tilde{x}_k, x_k),$$

so continuation with arm k is optimal.■

Bibliography

[1] Applebaum, D., 2009, *Levy Processes and Stochastic Calculus,* Second edition, Cambridge studies in advanced mathematics.

[2] Aubin, J.P. 2007. *Mathematical Methods of Game and Economic Theory: Revised Edition.* Dover Publications.

[3] Banks, J.S. and Sundaram, R. 1992. "Denumerable-Armed Bandits," *Econometrica*, 60, 1071-1096.

[4] Bergemann, D., and Valimaki, 2006. "Bandit Problems," Cowles Foundation Discussion Paper, no.1551.

[5] Biais, B., and Casamtta, C., 1999, "Optimal Leverage and Aggregate Investment," *Journal of Finance*, 1291-1323.

[6] Biais, B., Mariotti, T., Plantin, G., and Rochet, J-C. 2007 "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economics Studies*, 74, 345-390.

[7] Biais, B., Mariotti, T., Rochet, J-C., 2013. "Dynamic Financial Contracting," *Advances in Economics and Econometrics*, Tenth World Congress, vol. 1, Cambridge University Press.

[8] Biais, B., Mariotti, T., Rochet, J-C., and Villeneuve, S., 2010, "Large Risks, Limited Liability, and Dynamic Moral Hazard," *Econometrica*, 78(1), 73-118.

[9] Bolton, P., Chen, H., and Wang, N., 2011. "A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management," *Journal of Finance*, 1545-1578.

[10] Bonatti, A., and Horner, J., 2011. "Collaborating," *American Economic Review*, 101, 632-663.

[11] Cadenilla, A., Cvitanic, J., and Zapatero, F. 2007. "Optimal Risk-Sharing with Effort and Project Choice," *Journal of Economic Theory*, 133, 403-440.

[12] Che, Y.K., and Yoo, S.W. 2001, "Optimal Incentives for Teams," *American Economic Review*, 12, 525-541.

[13] Cvitanic, J. and Zhang, J. 2012, *Contract Theory in Continuous-Time Models.* Springer Finance.

[14] DeMarzo, P., Fishman, M., He, Z., and Wang, N. 2012. "Dynamic Agency and the q Theory of Investment," *Journal of Finance*, 67, 2295-2340.

[15] DeMarzo, P., Livdan, D., and Tchistyi, A., 2011, "Risking Other People's Money: Gambling, Limited Liability, and Optimal Incentives," working paper.

[16] DeMarzo, P., and Sannikov, Y. 2006. "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 61, 2681-2724.

[17] Epstein, L. and Schneider, M. 2003. "Recursive Multiple-Priors," *Journal of Economic Theory*, 113, 1-31.

[18] Georgiadis, G. 2014. "Projects and Team Dynamics ," *Review of Economic Studies*, forthcoming.

[19] Gilboa,I. and Schmeidler, D. 1989. "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141-153.

[20] Gittins, J., Glazerbrook, K., and Weber, R. (2010): *Multi-armed Bandit Allocation Indices*. Wiley; second edition.

[21] Gittins, J. and Jones, D. 1974. "A Dynamic Allocation Index for the Sequential Allocation of Experiments," In *Progress in Statistics*. 241-266.

[22] Hansen, L.P., and Sargent, T.J. 2001. "Acknowledging Misspecification in Macroeconomic Theory," *Review of Economic Dynamics*, 4, 519-535.

[23] He, Z. 2009. "Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion," *Review of Financial Studies*, 22, 859-892.

[24] He, Z. 2011. "A Model of Dynamic Compensation and Capital Structure," *Journal of Financial Economics*, 100, 351-366.

[25] Hoffmann, F. and Pfeil, S. 2010. "Reward for Luck in a Dynamic Agency Model," *Review of Financial Studies*, 23, 3329-3345.

[26] Holmstrom, B., 1979. "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74-91.

[27] Holmstrom, B., 1982. "Moral Hazard in Teams," *The Bell Journal of Economics*, 13, 324-340.

[28] Karatzas, I., and Sherve, S., 1991 *Brownian Motion and Stochastic Calculus.*, Springer-Verlag, New York.

[29] Karoui, N.E. and Karatzas, I. 1993. "General Gittins Index Processes in Discrete Time," *Proceedings of the National Academy of Sciences of the USA*, 90, 1232-1236.

[30] Kolokolova, O., and Mattes, A., 2013. "Recovering Risk-Taking from Daily Hedge Funds Returns: Incentives at Work?" *working paper*.

[31] Leland, H., 1994. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, 49, 1213-1252.

[32] Leland, H., 1998. "Agency Costs, Risk Management, and Capital Structure," *Journal of Finance*, 1213-1243.

[33] Maccheroni, F., Marinacci, M., and Rustichini, A., 2006. "Ambiguity Aversion, Robustness and the Variational Representation of Preferences," *Econometrica*, 74, 1447-1498.

[34] Makarov, I., and Plantin, G., 2012, "Rewarding Trading Skills without Inducing Gambling," working paper.

[35] Miao, J. and Rivera, A. 2013, "Robust Contracts in Continuous Time," working paper.

[36] Miao, J. and Wang, N. 2011. "Risk, Uncertainty, and Option Exercise," *Journal of Economic Dynamics and Control*, 35, 442-461.

[37] Milgrom, P. and Segal, I. 2002. "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583-601.

[38] Mookherjee, D. 1984. "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies*, 51, 433-446.

[39] Oksendal, B. and Sulem, A. 2005, *Applied Stochastic Control of Jump Diffusions.* Springer Finance.

[40] Palomino, F., and Prat, A. 2003 "Risk-Taking and Optimal Contracts for Money Managers," *Rand Journal of Economics*, 34, 113-137

[41] Panageas, S., and Westerfield, M. 2009. "High-Water Marks: High Risk Appetities? Convex Compensation, Long Horizons, and Portfolio Choice," *Journal of Finance*, 64, 1-36.

[42] Piskorski, T. and Tchistyi, A. 2010. "Optimal Mortgage Design," *Review of Financial Studies*, 23, 3098-3140.

[43] Rampini, A., and Viswanathan, S., 2010. "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance,* 65, 2293-2322.

[44] Riedel, F. 2009. "Optimal Stopping With Multiple Priors," *Econometrica*, 77, 857-908.

[45] Sannikov, Y., 2008, "A Continuous-Time Version of the Principal-Agent Problem," *Review of Economic Studies*, 75, 957-984.

[46] Sannikov, Y., 2012, "Dynamic Security Design and Corporate Financing," *Handbook of Economics and Finance, Vol. 2*, Elsevier-Northholland.

[47] Shelef, O. 2012. "Incentive Thresholds, Risk-Taking, and Performance. Evidence from Hedge Funds," working paper.

[48] Sufi, A. 2009. "Bank Lines of Credit in Corporate Finance: An Empirical Analysis," *Review of Financial Studies*, 22, 1057-1088.

[49] Sung, J., 1995, "Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems," *RAND Journal of Economics*, 26(4), 720-743.

[50] Szydlowski, M. 2012, "Incentives, Project Choice, and Dynamic Multitasking," *working paper.*

[51] Whittle, P. 1980. "Multi-armed Bandits and the Gittins Index," *Journal of Royal Statistical Society Series B*, 42, 143-149.

[52] Williams, N. 2009. "On Dynamic Principal-Agent Problems in Continuous Time," *working paper.*

[53] Williams, N. 2011. "Persistent Private Information," *Econometrica*, 79, 1233-1274.

[54] Wong, T.Y. 2013. "Dynamic Team Incentives," *working paper*, Boston University.

[55] Zhang, Y. 2009. "Dynamic Contracting with Persistent Shocks," *Journal of Economic Theory*, 144, 635-675.

[56] Zhu, J., 2012, "Optimal Contracts with Shirking," *Review of Economics Studies*, forthcoming.

Curriculum Vitae

Contact	Tak-Yuen Wong Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215, USA etywong110@gmail.com
Education	**University of Hong Kong**, B.Economics and Finance (First Class Honors), August 2003 – June 2006. **Chinese University of Hong Kong**, M.Phil, Economics, August 2006 – August 2008. **Boston University**, Ph.D. candidate, August 2008 – present. Thesis advisor: Jianjun Miao.
Fields of Interest	Microeconomic Theory, Information Economics, Corporate Finance
Fellowships and Awards	Special Fellowship, Boston University, Fall 2013 – Spring 2014 Graduate Fellowship, Boston University, Fall 2008 – Spring 2013 Highest GPA Award, Chinese University of Hong Kong, 2008 Dean's Honors List, University of Hong Kong, 2005 and 2006